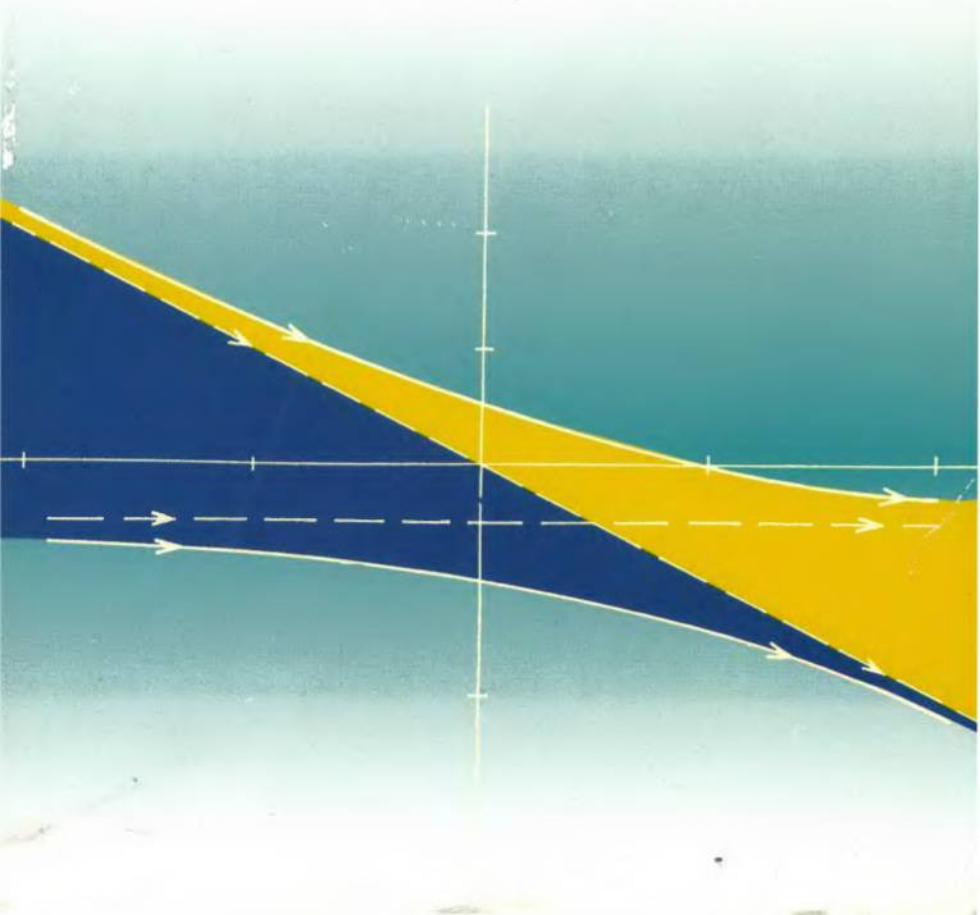
ALBERT MESSIAH

QUANTUM MECHANICS

Two Volumes Bound as One



DOVER BOOKS ON PHYSICS

METHODS OF QUANTUM FIELD THEORY IN STATISTICAL PHYSICS, A.A. Abrikosov et al. (0-486-63228-8)

DYNAMIC LIGHT SCATTERING: WITH APPLICATIONS TO CHEMISTRY, BIOLOGY, AND PHYSICS, Bruce J. Berne and Robert Pecora. (0-486-41155-9)

QUANTUM THEORY, David Bohm. (0-486-65969-0)

ATOMIC PHYSICS (8TH EDITION), Max Born. (0-486-65984-4)

EINSTEIN'S THEORY OF RELATIVITY, Max Born. (0-486-60769-0)

INTRODUCTION TO HAMILTONIAN OPTICS, H.A. Buchdahl. (0-486-67597-1)

THE METAPHYSICAL FOUNDATIONS OF MODERN SCIENCE, E. A. Burtt. (0-486-42551-7)

MATHEMATICS OF CLASSICAL AND QUANTUM PHYSICS, Frederick W. Byron, Jr. and Robert W. Fuller. (0-486-67164-X)

PRIMER OF QUANTUM MECHANICS, Marvin Chester. (0-486-42878-8)

THE THEORY OF THE ELECTROMAGNETIC FIELD, David M. Cook. (0-486-42567-3)

MECHANICS, J.P. Den Hartog. (0-486-60754-2)

MATHEMATICAL METHODS IN PHYSICS AND ENGINEERING, John W. Dettman. (0-486-65649-7)

INVESTIGATIONS ON THE THEORY OF THE BROWNIAN MOVEMENT, Albert Einstein. (0-486-60304-0)

THE PRINCIPLE OF RELATIVITY, Albert Einstein, et al. (0-486-60081-5)

THE PHYSICS OF WAVES, William C. Elmore and Mark A. Heald. (0-486-64926-1)

THERMODYNAMICS, Enrico Fermi. (0-486-60361-X)

QUANTUM THEORY OF MANY-PARTICLE SYSTEMS, Alexander L. Fetter and John Dirk Walecka. (0-486-42827-3)

INTRODUCTION TO MODERN OPTICS, Grant R. Fowles. (0-486-65957-7)

DIALOGUES CONCERNING TWO NEW SCIENCES, Galileo Galilei. (0-486-60099-8)

GROUP THEORY AND ITS APPLICATION TO PHYSICAL PROBLEMS, Morton Hamermesh. (0-486-66181-4)

ELECTRONIC STRUCTURE AND THE PROPERTIES OF SOLIDS: THE PHYSICS OF THE CHEMICAL BOND, Walter A. Harrison. (0-486-66021-4)

SOLID STATE THEORY, Walter A. Harrison. (0-486-63948-7)

PHYSICAL PRINCIPLES OF THE QUANTUM THEORY, Werner Heisenberg. (0-486-60113-7)

ATOMIC SPECTRA AND ATOMIC STRUCTURE, Gerhard Herzberg. (0-486-60115-3)

AN INTRODUCTION TO STATISTICAL THERMODYNAMICS, Terrell L. Hill. (0-486-65242-4)

EQUILIBRIUM STATISTICAL MECHANICS, E. Atlee Jackson. (0-486-41185-0)

OPTICS AND OPTICAL INSTRUMENTS: AN INTRODUCTION, B.K. Johnson. (0-486-60642-2)

THEORETICAL SOLID STATE PHYSICS, VOLS. I & II, William Jones and Norman H. March. (0-486-65015-4, 0-486-65016-2)

THEORETICAL PHYSICS, Georg Joos, with Ira M. Freeman. (0-486-65227-0)

THE VARIATIONAL PRINCIPLES OF MECHANICS, Cornelius Lanczos. (0-486-65067-7)

WHAT IS RELATIVITY? L. D. Landau and G. B. Rumer. (0-486-42806-0)

(continued on back flap)

QUANTUM MECHANICS

TWO VOLUMES BOUND AS ONE

Albert Messiah

DOVER PUBLICATIONS, INC.
Mineola, New York

Bibliographical Note

This Dover edition, first published in 1999, is an unabridged reprint in one volume of the English translation of *Mécanique Quantique*, published in two volumes in the United States as *Quantum Mechanics* by John Wiley & Sons c.1958. Volume I was translated from the French by G. H. Temmer, Vol. II by J. Potter. The combined index for both volumes has been placed at the end of this new edition.

Library of Congress Cataloging-in-Publication Data

Messiah, Albert.
[Mécanique quantique. English]
Quantum mechanics / Albert Messiah.
p. cm.
Translation of: Mécanique quantique.
Previously published in English: Amsterdam : North-Holland Pub. Co. ; New York : Interscience Publishers, 1961–1962.
Includes index.
ISBN-13: 978-0-486-40924-5
ISBN-10: 0-486-40924-4
1. Quantum theory. I. Title.

QC174.12 .M4813 1999
530.12—dc21

99-055362

Manufactured in the United States by Courier Corporation
40924409
www.doverpublications.com

PREFACE

Nowadays, there hardly exists a branch of physics which one can seriously approach without a thorough knowledge of Quantum Mechanics. Its presentation, which is given in this work is, I hope, simple enough to be accessible to the student, and yet sufficiently complete to serve as a reference book for the working physicist.

This book resulted from a course given at the Center of Nuclear Studies at Saclay since 1953. Numerous discussions with students as well as with my colleagues, have helped me considerably in clarifying its presentation. Several people to whom I had transmitted certain parts of the manuscript, have kindly given me their criticism; among them I should like to mention Messrs. Edmond Bauer and Jean Ullmo, to whom I am indebted for interesting remarks concerning the presentation of principles. I am more particularly grateful to Mr. Roger Balian for having critically examined a large portion of the manuscript, and for having suggested to me a large number of improvements. Finally, I wish to thank those of my students who were kind enough to check over the text and the calculations of the various chapters, and to help me with the correction of the proofs.

The problems which occur at the end of each chapter were chosen not only for their educational value, but also to point out certain properties worthy of interest; this may explain the relative difficulty of certain ones among them.

The several works or articles cited as references have the purpose of aiding the reader to complete or round out certain passages. It was out of the question to give a complete bibliography of the various subjects treated here. An entire volume would not have sufficed for that.

October, 1958 ALBERT MESSIAH

CONTENTS

OF VOLUME I

PART ONE

THE FORMALISM AND ITS INTERPRETATION

CHAPTER I

THE ORIGINS OF THE QUANTUM THEORY

1. Introduction 3

I. THE END OF THE CLASSICAL PERIOD 4
2. Classical Theoretical Physics. 3. Progress in the knowledge of microscopic phenomena and the appearance of quanta in physics.

II. LIGHT QUANTA OR PHOTONS 11
4. The photoelectric effect. 5. The Compton effect. 6. Light quanta and interference phenomena. 7. Conclusions.

III. QUANTIZATION OF MATERIAL SYSTEMS 21
8. Atomic spectroscopy and difficulties of Rutherford's classical model. 9. Quantization of atomic energy levels. 10. Other examples of quantization: space quantization.

IV. CORRESPONDENCE PRINCIPLE AND THE OLD QUANTUM THEORY . 27
11. Inadequacy of classical corpuscular theory. 12. Correspondence principle. 13. Application of the correspondence principle to the calculation of the Rydberg constant. 14. Lagrange's and Hamilton's forms of the equations of classical mechanics. 15. Bohr–Sommerfeld quantization rules. 16. Successes and limitations of the Old Quantum Theory. 17. Conclusions.

CHAPTER II

MATTER WAVES AND THE SCHRÖDINGER EQUATION

1. Historical survey and general plan of the succeeding chapters 45

I. MATTER WAVES 49
2. Introduction. 3. Free wave packet. Phase velocity and group velocity. 4. Wave packet in a slowly varying field. 5. Quantization of atomic energy levels. 6. Diffraction of matter waves. 7. Corpuscular structure of matter. 8. Universal character of the wave-corpuscle duality.

II. THE SCHRÖDINGER EQUATION 59
9. Conservation law of the number of particles of matter. 10. Necessity for a wave equation and conditions imposed upon this equation. 11. The operator concept. 12. Wave equation of a free particle. 13. Particle in a scalar potential. 14. Charged particle in an electromagnetic field. 15. General rule for forming the Schrödinger equation by correspondence.

III. THE TIME-INDEPENDENT SCHRÖDINGER EQUATION 71
16. Search for stationary solutions. 17. General properties of the equation. Nature of the energy spectrum.

CHAPTER III

ONE-DIMENSIONAL QUANTIZED SYSTEMS

1. Introduction 77

I. SQUARE POTENTIALS 78
2. General remarks. 3. Potential step. Reflection and transmission of waves. 4. Infinitely high potential barrier. 5. Infinitely deep square potential well. Discrete spectrum. 6. Study of a finite square well. Resonances. 7. Penetration of a square potential barrier. The "tunnel" effect.

II. GENERAL PROPERTIES OF THE ONE-DIMENSIONAL SCHRÖDINGER EQUATION 98
8. Property of the Wronskian. 9. Asymptotic behavior of the solutions. 10. Nature of the eigenvalue spectrum. 11. Unbound states: reflection and transmission of waves. 12. Number of nodes of bound states. 13. Orthogonality relations. 14. Remark on parity.

CHAPTER IV

STATISTICAL INTERPRETATION OF THE WAVE-CORPUSCLE DUALITY AND THE UNCERTAINTY RELATIONS

1. Introduction 115

I. STATISTICAL INTERPRETATION OF THE WAVE FUNCTIONS OF WAVE MECHANICS 116
2. Probabilities of the results of measurement of the position and the momentum of a particle. 3. Conservation in time of the norm. 4. Concept of current. 5. Mean values of functions of $\boldsymbol{r}$ or of $\boldsymbol{p}$. 6. Generalization to systems of several particles.

II. HEISENBERG'S UNCERTAINTY RELATIONS 129
7. Position-momentum uncertainty relations of a quantized particle. 8. Precise statement of the position-momentum un-

certainty relations. 9. Generalization: uncertainty relations between conjugate variables. 10. Time-energy uncertainty relation. 11. Uncertainty relations for photons.

III. UNCERTAINTY RELATIONS AND THE MEASUREMENT PROCESS . . . 139
12. Uncontrollable disturbance during the operation of measurement. 13. Position measurements. 14. Momentum measurements.

IV. DESCRIPTION OF PHENOMENA IN QUANTUM THEORY. COMPLEMENTARITY AND CAUSALITY 149
15. Problems raised by the statistical interpretation. 16. Description of microscopic phenomena and complementarity. 17. Complementary variables. Compatible variables. 18. Wave-corpuscle duality and complementarity. 19. Complementarity and causality.

CHAPTER V

DEVELOPMENT OF THE FORMALISM OF WAVE MECHANICS AND ITS INTERPRETATION

1. Introduction 162

I. HERMITEAN OPERATORS AND PHYSICAL QUANTITIES 163
2. Wave-function space. 3. Definition of mean values. 4. Absence of fluctuations and the eigenvalue problem.

II. STUDY OF THE DISCRETE SPECTRUM 171
5. Eigenvalues and eigenfunctions of a Hermitean operator. 6. Expansion of a wave function in a series of orthonormal eigenfunctions. 7. Statistical distribution of the results of measurement of a quantity associated with an operator having a complete set of eigenfunctions with finite norm.

III. STATISTICS OF MEASUREMENT IN THE GENERAL CASE 179
8. Difficulties of the continuous spectrum. Introduction of the Dirac δ-functions. 9. Expansion in a series of eigenfunctions in the general case. Closure relation. 10. Statistical distribution of the results of measurement in the general case. 11. Other ways of treating the continuous spectrum. 12. Comments and examples.

IV. DETERMINATION OF THE WAVE FUNCTION 196
13. Measuring process and "filtering" of the wave packet. Ideal measurements. 14. Commuting observables and compatible variables. 15. Complete sets of commuting observables. 16. Pure states and mixtures.

V. COMMUTATOR ALGEBRA AND ITS APPLICATIONS 206
17. Commutator algebra and properties of basic commutators. 18. Commutation relations of angular momentum. 19. Time dependence of the statistical distribution. Constants of the motion. 20. Examples of constants of the motion. Energy. Parity.

CHAPTER VI

CLASSICAL APPROXIMATION AND THE WKB METHOD

I. THE CLASSICAL LIMIT OF WAVE MECHANICS 214
1. General remarks. 2. Ehrenfest's theorem. 3. Motion and spreading of wave packets. 4. Classical limit of the Schrödinger equation. 5. Application to Coulomb scattering. The Rutherford formula.

II. THE WKB METHOD 231
6. Principle of the method. 7. One-dimensional WKB solutions. 8. Conditions for the validity of the WKB approximation. 9. Turning points and connection formulae. 10. Penetration of a potential barrier. 11. Energy levels of a potential well.

CHAPTER VII

GENERAL FORMALISM OF THE QUANTUM THEORY

(A) MATHEMATICAL FRAMEWORK

1. Superposition principle and representation of dynamical states by vectors 243

I. VECTORS AND OPERATORS 245
2. Vector space. "Ket" vectors. 3. Dual space. "Bra" vectors. 4. Scalar product. 5. Linear operators. 6. Tensor product of two vector spaces.

II. HERMITEAN OPERATORS, PROJECTORS, AND OBSERVABLES . . . 254
7. Adjoint operators and conjugation relations. 8. Hermitean (or self-adjoint) operators, positive definite Hermitean operators, unitary operators. 9. Eigenvalue problem and observables. 10. Projectors (Projection operators). 11. Projector algebra. 12. Observables possessing an entirely discrete spectrum. 13. Observables in the general case. Generalized closure relation. 14. Functions of an observable. 15. Operators which commute with an observable. Commuting observables.

III. REPRESENTATION THEORY 273
16. General remarks on finite matrices. 17. Square matrices. 18. Extension to infinite matrices. 19. Representation of vectors and operators by matrices. 20. Matrix transformations. 21. Change of representation. 22. Unitary transformations of operators and vectors.

CHAPTER VIII

GENERAL FORMALISM

(B) DESCRIPTION OF PHYSICAL PHENOMENA

1. Introduction 294

I. DYNAMICAL STATES AND PHYSICAL QUANTITIES 296
2. Definition of probabilities. Postulates concerning measurement. 3. Observables of a quantized system and their commutation relations. 4. Heisenberg's uncertainty relations. 5. Definition of the dynamical states and construction of the space $\mathscr{E}$. 6. One-dimensional quantum system having a classical analogue. 7. Construction of the $\mathscr{E}$-space of a system by tensor product of simpler spaces.

II. THE EQUATIONS OF MOTION 310
8. Evolution operator and the Schrödinger equation. 9. Schrödinger "representation". 10. Heisenberg "representation". 11. Heisenberg "representation" and correspondence principle. 12. Constants of the motion. 13. Equations of motion for the mean values. Time-energy uncertainty relation. 14. Intermediate representations.

III. VARIOUS REPRESENTATIONS OF THE THEORY 323
15. Definition of a representation. 16. Wave mechanics. 17. Momentum representation ($\{p\}$-representation). 18. An example: motion of a free wave packet. 19. Other representations. Representations in which the energy is diagonal.

IV. QUANTUM STATISTICS 331
20. Incompletely known systems and statistical mixtures. 21. The density operator. 22. Evolution in time of a statistical mixture. 23. Characteristic properties of the density operator. 24. Pure states. 25. Classical and quantum statistics.

PART TWO

SIMPLE SYSTEMS

CHAPTER IX

SOLUTION OF THE SCHRÖDINGER EQUATION BY SEPARATION OF VARIABLES. CENTRAL POTENTIAL

1. Introduction 343

I. PARTICLE IN A CENTRAL POTENTIAL. GENERAL TREATMENT . . 344
2. Expression of the Hamiltonian in spherical polar coordinates.

3. Separation of the angular variables. Spherical harmonics. 4. The radial equation. 5. Eigensolutions of the radial equation. Nature of the spectrum. 6. Conclusions.

II. CENTRAL SQUARE-WELL POTENTIAL. FREE PARTICLE 355
7. Spherical Bessel functions. 8. Free particle. Plane waves and free spherical waves. 9. Expansion of a plane wave in spherical harmonics. 10. Study of a spherical square well.

III. TWO-BODY PROBLEMS. SEPARATION OF THE CENTER-OF-MASS MOTION 361
11. Separation of the center-of-mass motion in classical mechanics. 12. Separation of the center-of-mass motion of a quantized two-particle system. 13. Extension to systems of more than two particles.

CHAPTER X

SCATTERING PROBLEMS

CENTRAL POTENTIAL AND PHASE-SHIFT METHOD

1. Introduction . 369

I. CROSS SECTIONS AND SCATTERING AMPLITUDES 369
2. Definition of cross sections. 3. Stationary wave of scattering. 4. Representation of the scattering phenomenon by a bundle of wave packets. 5. Scattering of a wave packet by a potential. 6. Calculation of cross sections. 7. Collision of two particles. Laboratory system and center-of-mass system.

II. SCATTERING BY A CENTRAL POTENTIAL. PHASE SHIFTS . . . 385
8. Decomposition into partial waves. Phase-shift method. 9. Semi-classical representation of the collision. Impact parameters.

III. POTENTIAL OF FINITE RANGE 389
10. Relation between phase shift and logarithmic derivative. 11. Behavior of the phase shift at low energies. 12. Partial waves of higher order. Convergence of the series. 13. Scattering by a hard sphere.

IV. SCATTERING RESONANCES 396
14. Scattering by a deep square well. 15. Study of a scattering resonance. Metastable states. 16. Observation of the lifetime of metastable states.

V. VARIOUS FORMULAE AND PROPERTIES 404
17. Integral representations of phase shifts. 18. Dependence upon the potential. Sign of the phase shifts. 19. The Born approximation. 20. Effective range theory. The Bethe formula.

CHAPTER XI

THE COULOMB INTERACTION

1. Introduction 411

I. THE HYDROGEN ATOM 412
2. Schrödinger equation of the hydrogen atom. 3. Order of magnitude of the binding energy of the ground state. 4. Solution of the Schrödinger equation in spherical coordinates. 5. Energy spectrum. Degeneracy. 6. The eigenfunctions of the bound states.

II. COULOMB SCATTERING 421
7. The Coulomb scattering wave. 8. The Rutherford formula. 9. Decomposition into partial waves. 10. Expansion of the wave ψ_c in spherical harmonics. 11. Modifications of the Coulomb potential by a short-range interaction.

CHAPTER XII

THE HARMONIC OSCILLATOR

1. Introduction 432

I. EIGENSTATES AND EIGENVECTORS OF THE HAMILTONIAN . . . 433
2. The eigenvalue problem. 3. Introduction of the operators a, $a^\dagger$ and N. 4. Spectrum and basis of N. 5. The $\{N\}$ representation. 6. Creation and destruction operators. 7. $\{Q\}$ representation. Hermite polynomials.

II. APPLICATIONS AND VARIOUS PROPERTIES 441
8. Generating function for the eigenfunctions $u_n(Q)$. 9. Integration of the Heisenberg equations. 10. Classical and quantized oscillator. 11. Motion of the minimum wave packet and classical limit. 12. Harmonic oscillators in thermodynamic equilibrium.

III. ISOTROPIC HARMONIC OSCILLATORS IN SEVERAL DIMENSIONS . . 451
13. General treatment of the isotropic oscillator in p dimensions. 14. Two-dimensional isotropic oscillator. 15. Three-dimensional isotropic oscillator.

APPENDIX A. DISTRIBUTIONS, δ-"FUNCTION" AND FOURIER TRANSFORMATION 462

APPENDIX B. SPECIAL FUNCTIONS AND ASSOCIATED FORMULAE 479

CONTENTS

OF VOLUME II

PART THREE

SYMMETRIES AND INVARIANCE

CHAPTER XIII

ANGULAR MOMENTUM IN QUANTUM MECHANICS

1. Introduction 507

I. EIGENVALUES AND EIGENFUNCTIONS OF ANGULAR MOMENTUM 508
2. Definition of angular momentum. 3. Characteristic algebraic relations. 4. Spectrum of J^2 and J_z. 5. Eigenvectors of J^2 and J_z. Construction of the invariant subspaces $\mathcal{E}^{(j)}$. 6. Standard representation $\{J^2 J_z\}$. 7. Conclusion.

II. ORBITAL ANGULAR MOMENTUM AND THE SPHERICAL HARMONICS 519
8. The spectrum of l^2 and l_z. 9. Definition and construction of the spherical harmonics.

III. ANGULAR MOMENTUM AND ROTATIONS 523
10. Definition of rotation. Euler angles. 11. Rotation of a physical system. Rotation operator. 12. Rotation of observables. 13. Angular momentum and infinitesimal rotations. 14. Construction of the operator $R(\alpha\ \beta\ \gamma)$. 15. Rotation through an angle 2π and half-integral angular momenta. 16. Irreducible invariant subspaces. Rotation matrices $R^{(j)}$. 17. Rotational invariance and conservation of angular momentum. Rotational degeneracy.

IV. SPIN 540
18. The hypothesis of electron spin. 19. Spin $\frac{1}{2}$ and the Pauli matrices. 20. Observables and wave functions of a spin $\frac{1}{2}$ particle. Spinor fields. 21. Vector fields and particles of spin 1. 22. Spin-dependent interactions in atoms. 23. Spin-dependent nucleon-nucleon interactions.

V. ADDITION OF ANGULAR MOMENTA 555
24. The addition problem. 25. Addition theorem for two angular momenta. 26. Applications and examples. 27. Eigenvectors of the total angular momentum. Clebsch–Gordon coefficients. 28. Application: two-nucleon system. 29. Addition of three or more angular momenta. Racah coefficients. "3sj" symbols.

VI. IRREDUCIBLE TENSOR OPERATORS 569
30. Representation of scalar operators. 31. Irreducible tensor operators. Definition. 32. Representation of irreducible tensor operators. Wigner–Eckhart theorem. 33. Applications.

CHAPTER XIV

SYSTEMS OF IDENTICAL PARTICLES. PAULI EXCLUSION PRINCIPLE

1. Identical particles in quantum theory . 582

I. SYMMETRIZATION POSTULATE 586
2. Similar particles and the symmetrical representation. 3. Permutation operators. 4. Algebra of permutation operators. Symmetrizers and antisymmetrizers. 5. Identical particles and the symmetrization postulate. 6. Bosons and Bose–Einstein statistics. 7. Fermions and Fermi–Dirac statistics. Exclusion principle. 8. It is always necessary to symmetrize the wave-function.

II. APPLICATIONS 603
9. Collision of two spinless identical particles. 10. Collision of two protons. 11. Statistics of atomic nuclei. 12. Complex atoms. Central field approximation. 13. The Thomas–Fermi model of the atom. 14. Nucleon systems and isotopic spin. 15. Utility of isotopic spin. Charge independence.

CHAPTER XV

INVARIANCE AND CONSERVATION THEOREMS. TIME REVERSAL

1. Introduction 632

I. MATHEMATICAL COMPLEMENTS. ANTILINEAR OPERATORS . . . 633
2. Three useful theorems. 3. Antilinear operators in Hilbert space. 4. Antilinear transformations. 5. Antilinear operators and representations.

II. TRANSFORMATIONS AND GROUPS OF TRANSFORMATIONS 643
6. Transformations of the dynamical variables and dynamical states of a system. 7. Groups of transformations. 8. Groups of transformation operators. 9. Continuous groups and infinitesimal transformations. Translations. Rotations. 10. Finite groups. Reflections.

III. INVARIANCE OF THE EQUATIONS OF MOTION AND CONSERVATION LAWS 655
11. Invariant observables. 12. Symmetry of the Hamiltonian and conservation laws. 13. Invariance properties and the evolution

of dynamical states. 14. Symmetries of the Stark and Zeeman effects.

IV. TIME REVERSAL AND THE PRINCIPLE OF MICROREVERSIBILITY 664
15. Time translation and conservation of energy. 16. Time reversal in classical mechanics and in quantum mechanics. 17. The time-reversal operation. Spinless particle. 18. General definition of time reversal. 19. Time reversal and complex conjugation. 20. Principle of microreversibility. 21. Consequence: Kramers degeneracy. 22. Real rotation-invariant Hamiltonian.

PART FOUR

METHODS OF APPROXIMATION

CHAPTER XVI

STATIONARY PERTURBATIONS

1. General introduction to Part Four 685

I. PERTURBATION OF A NON-DEGENERATE LEVEL 686
2. Expansion in powers of the perturbation. 3. First-order perturbations. 4. Ground state of the helium atom. 5. Coulomb energy of atomic nuclei. 6. Higher-order corrections. 7. Stark effect for a rigid rotator.

II. PERTURBATION OF A DEGENERATE LEVEL 698
8. Elementary theory. 9. Atomic levels in the absence of spin-orbit forces. 10. Spin-orbit forces. LS and jj coupling. 11. The atom in LS coupling. Splitting due to spin-orbital coupling. 12. The Zeeman and Paschen–Back effects. 13. Symmetry of H and removal of degeneracy. 14. Quasi-degeneracy.

III. EXPLICIT FORMS FOR THE PERTURBATION EXPANSION IN ALL ORDERS 712
15. The Hamiltonian H and its resolvent $G(z)$. 16. Expansion of $G(z)$, P and HP into power series in V. 17. Calculation of eigenvalues and eigenstates.

CHAPTER XVII

APPROXIMATE SOLUTIONS OF THE TIME-DEPENDENT SCHRÖDINGER EQUATION

1. Change of "representation" and perturbation treatment of a part of the Hamiltonian 722

I. TIME DEPENDENT PERTURBATION THEORY 724
2. Definition and perturbation calculation of transition probabilities. 3. Semi-classical theory of Coulomb excitation of nuclei.

4. Case when V is independent of time. Conservation of unperturbed energy. 5. Application to the calculation of cross-sections in the Born approximation. 6. Periodic perturbation. Resonances.

II. SUDDEN OR ADIABATIC CHANGE OF THE HAMILTONIAN . . . 739
7. The problem and the results. 8. Rapid passage and the sudden approximation. 9. Sudden reversal of a magnetic field. 10. Adiabatic passage. Generalities. Trivial case. 11. "Rotating axis representation". 12. Proof of the adiabatic theorem. 13. Adiabatic approximation. 14. Adiabatic reversal of a magnetic field.

CHAPTER XVIII

THE VARIATIONAL METHOD AND ASSOCIATED PROBLEMS

1. The Ritz variational method 762

I. VARIATIONAL METHOD FOR BOUND STATES 763
2. Variational form of the eigenvalue problem. 3. Variational calculation of discrete levels. 4. A simple example: the hydrogen atom. 5. Discussion. Application to the calculation of excited levels. 6. Ground state of the helium atom.

II. THE HARTREE AND FOCK–DIRAC ATOMS 773
7. The self-consistent field method. 8. Calculation of $E[\Phi]$. 9. The Fock–Dirac equations. 10. Discussion. 11. The Hartree equations.

III. THE STRUCTURE OF MOLECULES 781
12. Generalities. Separation of the electronic and nuclear motions. 13. Motion of the electrons in the presence of fixed nuclei. 14. The adiabatic approximation. 15. Hamiltonian for the nuclei in the adiabatic approximation. 16. The Born–Oppenheimer method. 17. Notions on diatomic molecules.

CHAPTER XIX

COLLISION THEORY

1. Introduction 801

I. FREE WAVE GREEN'S FUNCTION AND THE BORN APPROXIMATION 802
2. Integral representations of the scattering amplitude. 3. Cross sections and the T matrix. Microreversibility. 4. The Born approximation. 5. Integral equation for scattering. 6. The Born expansion. 7. Validity criterion for the Born approximation. 8. Elastic scattering of electrons by an atom. 9. Central potential. Calculation of phase shifts. 10. Green's function as an operator. Relation to the resolvent of H_0.

II. GENERALIZATION TO DISTORTED WAVES 822
11. Generalized Born approximation. 12. Generalization of the

Born expansion. 13. Green's functions for distorted waves. 14. Applications. Definition and formal properties of T. 15. Note on the $1/r$ potentials.

III. COMPLEX COLLISIONS AND THE BORN APPROXIMATION 832
16. Generalities. Cross sections. 17. Channels. 18. Calculation of cross sections. T matrices. 19. Integral representations of the transition amplitude. 20. The Born approximation and its generalizations. 21. Scattering of fast electrons by an atom. 22. Coulomb excitation of nuclei. 23. Green's functions and integral equations for stationary scattering waves. 24. Scattering of a particle by two scattering centers. 25. Simple scattering. Interference. 26. Multiple scattering.

IV. VARIATIONAL CALCULATIONS OF TRANSITION AMPLITUDES . . 856
27. Stationary expressions for the phase shifts. 28. The variational calculation of phase shifts. Discussion. 29. Extension to complex collisions.

V. GENERAL PROPERTIES OF THE TRANSITION MATRIX 863
30. Conservation of flux. Unitarity of the S matrix. 31. The Bohr–Peierls–Placzek relation (optical theorem). 32. Microreversibility. 33. Invariance properties of the T matrix.

PART FIVE

ELEMENTS OF RELATIVISTIC QUANTUM MECHANICS

CHAPTER XX

THE DIRAC EQUATION

I. GENERAL INTRODUCTION 875
1. Relativistic quantum mechanics. 2. Notation, various conventions and definitions. 3. The Lorentz group. 4. Classical relativistic dynamics.

II. THE DIRAC AND KLEIN–GORDON EQUATIONS 884
5. The Klein–Gordon equation. 6. The Dirac equation. 7. Construction of the space $\mathscr{E}^{(s)}$. Dirac representation. 8. Covariant form of the Dirac equation. 9. Adjoint equation. Definition of the current.

III. INVARIANCE PROPERTIES OF THE DIRAC EQUATION 896
10. Properties of the Dirac matrices. 11. Invariance of the form of the Dirac equation in an orthochronous change of referential. 12. Transformation of the proper group. 13. Spatial reflection and the orthochronous group. 14. Construction of covariant quantities. 15. A second formulation of the invariance of form: transformation of states. 16. Invariance of the law of motion.

17. Transformation operators. Momentum, angular momentum, parity. 18. Conservation laws and constants of the motion. 19. Time reversal and charge conjugation. 20. Gauge invariance.

IV. INTERPRETATION OF THE OPERATORS AND SIMPLE SOLUTIONS . . 919
21. The Dirac equation and the correspondence principle. 22. Dynamical variables of a Dirac particle. 23. The free electron. Plane waves. 24. Construction of the plane waves by a Lorentz transformation. 25. Central potential. 26. Free spherical waves. 27. The hydrogen atom.

V. NON-RELATIVISTIC LIMIT OF THE DIRAC EQUATION 933
28. Large and small components. 29. The Pauli theory as the non-relativistic limit of the Dirac theory. 30. Application: hyperfine structure and dipole-dipole coupling. 31. Higher-order corrections and the Foldy–Wouthuysen transformation. 32. FW transformation for a free particle. 33. FW transformation for a particle in a field. 34. Electron in a central electrostatic potential. 35. Discussions and conclusions.

VI. NEGATIVE ENERGY SOLUTIONS AND POSITRON THEORY . . . 949
36. Properties of charge conjugate solutions. 37. Abnormal behavior of the negative energy solutions. 38. Reinterpretation of the negative energy states. Theory of "holes" and positrons. 39. Difficulties with the "hole" theory.

CHAPTER XXI

FIELD QUANTIZATION. RADIATION THEORY

1. Introduction 959

I. QUANTIZATION OF A REAL SCALAR FIELD 960
2. Classical free field. Normal vibrations. 3. Quantization of the free field. 4. Lagrangian of the field. Momentum conjugate to $\Phi(r)$. 5. Complex basis functions. 6. Plane waves. Definition of the momentum. 7. Spherical waves. Definition of the angular momentum. 8. Space and time reflections.

II. COUPLING WITH AN ATOMIC SYSTEM 979
9. Coupling to a system of particles. 10. Weak coupling and perturbation treatment. 11. Level shifts. 12. Emission of a corpuscle. 13. Quantum theory of decaying states. Line width. 14. Elastic scattering. Dispersion formula. 15. Resonance scattering. Formation of a metastable state. 16. Absorption of a corpuscle (photo-electric effect). Radiative capture.

III. CLASSICAL THEORY OF ELECTROMAGNETIC RADIATION 1009
17. The equations of the classical Maxwell–Lorentz theory. 18. Symmetries and conservation laws of the classical theory. 19.

Self-energy and classical radius of the electron. 20. Electromagnetic potential. Choice of the gauge. 21. Longitudinal and transverse parts of a vector field. 22. Elimination of the longitudinal field. 23. Energy, momentum, angular momentum. 24. Hamiltonian for free radiation. 25. Hamiltonian for radiation coupled to a set of particles.

IV. QUANTUM THEORY OF RADIATION 1029
26. Quantization of free radiation. Photons. 27. Plane waves. Radiation momentum. 28. Polarization. 29. Multipole expansion. Photons of determined angular momentum and parity. 30. Coupling with an atomic system. 31. Emission of a photon by an atom. Dipole emission. 32. Low energy Compton scattering. The Thomson formula.

APPENDIX C. VECTOR ADDITION COEFFICIENTS AND ROTATION MATRICES 1053

APPENDIX D. ELEMENTS OF GROUP THEORY . 1079

GENERAL INDEX 1125

Fluctuat nec mergitur

. . . "Il lui proposa de faire le voyage de Copenhague, et lui en facilita les moyens" (*Candide*)

PART ONE

THE FORMALISM AND ITS INTERPRETATION

CHAPTER I

THE ORIGINS OF QUANTUM THEORY

1. Introduction

According to the *classical doctrine* – generally adopted by physicists until the beginning of the 20th century – one associates with physical systems whose evolution one wishes to describe, a certain number of quantities or dynamical variables; each of these variables possesses at each instant a *well-defined value*, and the specification of this set of values defines the dynamical state of the system at that instant. One further postulates that the evolution in time of the physical system is entirely determined if one knows its state at a given initial instant. Mathematically this fundamental axiom is expressed more precisely by the fact that the dynamical variables satisfy a system of differential equations of the first order, as a function of time. The program of Classical Theoretical Physics thus consists in enumerating the dynamical variables of the system under study, and then in discovering the equations of motion which predict its evolution in accord with experimental observation.

From the formulation of Rational Mechanics by Newton until the end of the 19th century, this program was carried out with considerable success, each new experimental discovery being carried over to the theoretical plane either by introducing new variables and new equations, or by modifying the old equations, thereby allowing the newly observed phenomenon to be incorporated into the general scheme. During that entire period no experimental fact, no discovery led to any doubt concerning the soundness of the program itself. On the contrary, Classical Physics constantly progressed toward greater simplicity and greater unity. This happy evolution continued until about 1900; subsequently, as our knowledge of phenomena on the microscopic scale [1]) becomes more precise, Classical Theory runs

[1]) It is important to define the terms "microscopic" and "macroscopic" of which we shall make frequent use throughout this book. We define the "microscopic" scale as the one of atomic or subatomic phenomena, where the lengths which enter into consideration are at most of the order of several ångstroms (1 Å = 10^{-8} cm). The "macroscopic" scale is the one of phenomena observable with the naked eye or with the ordinary microscope, i.e. a resolution of the order of one micron (10^{-4} cm) at best.

into more and more difficulties and contradictions. It rapidly becomes evident that phenomena on the atomic and subatomic scale do not fit into the framework of classical doctrine itself, and that their explanation must be based upon entirely new principles. The discovery of these new principles will occur in stages, at the expense of numerous groping attempts; only around 1925, with the founding of Quantum Mechanics, will we have at our disposal a coherent theory of microscopic phenomena. The origins of this theory constitute the subject of the present chapter.

After sketching an overall picture of Classical Theoretical Physics, we shall discuss the main phenomena which justify the abandonment of the classical ideas. The phenomena are supposed familiar to the reader [1]); we shall therefore merely recall their essential features, emphasizing above all the points of contradiction with Classical Theory. The end of the chapter is devoted to a brief discussion of the first attempts at explaining these phenomena, known as the Old Quantum Theory.

I. THE END OF THE CLASSICAL PERIOD

2. Classical Theoretical Physics

At the end of the classical period, the various branches of physics are integrated in a general and coherent theoretical construct whose main features are as follows. In the universe, one distinguishes two categories of objects, *matter* and *radiation*. Matter is made up of perfectly localizable corpuscles subject to Newton's laws of Rational Mechanics; the state of each corpuscle is defined at any instant by its position and its velocity (or its momentum), that is six dynamical variables in all. Radiation obeys Maxwell's laws of electromagnetism; its dynamical variables — infinite in number — are the components of the electric and magnetic fields at each point of space. In contrast to matter, it is not possible to split radiation into corpuscles which can be localized in space and maintain this localized character during their evolution in the course of time; quite to the contrary, it exhibits a wave-like behavior which manifests itself particularly in the well-known phenomena of interference and diffraction.

[1]) One may find a detailed discussion of these phenomena in the works dealing with Atomic Physics, for instance: M. Born, *Atomic Physics*, 6th ed. (Blackie, Glasgow, 1957).

The *corpuscular theory of matter* continues to develop during the course of the 19th century. While limited at first to the mechanics of heavenly bodies and of solid bodies of macroscopic dimensions, it emerges more and more as the basic theory governing the evolution of matter on the microscopic scale to the extent that the atomic hypothesis, proposed by the chemists, is confirmed. Without being able to verify this hypothesis directly by isolating the molecules and studying their mutual interactions, one can justify it indirectly by showing that the macroscopic properties of material bodies derive from the laws of motion of the molecules of which they are composed. Mathematically, we are dealing with a very complex problem. Under this hypothesis, in fact, macroscopic quantities appear as the mean values of certain dynamical variables of a system having a very large number of degrees of freedom [1]); there is no hope of solving the equations of evolution of such a system exactly, and one must have recourse to statistical methods of investigation. Thus a new discipline originated and developed, Statistical Mechanics, whose results, particularly in the study of gases (Kinetic Theory of Gases) and in Thermodynamics (Statistical Thermodynamics) enable us to verify qualitatively, and within the limits set by the possibilities of calculation, quantitatively, the foundation of a corpuscular theory of matter [2]).

At the same time, the *wave theory of radiation* becomes solidly established. In the field of optics, the old controversy on the wave nature or corpuscle nature of light is cut short in the first half of the 19th century, when decisive progress in the handling of problems of wave propagation (Fresnel) permits the exploration of all the consequences of the wave hypothesis. All the known light phenomena, including geometrical optics can now be based on this hypothesis. Meanwhile, the study of electric and magnetic phenomena develops rapidly. The decisive step forward is taken by Maxwell when he establishes, in 1855, the fundamental electromagnetic equations.

[1]) We recall that the number N of molecules per mole (Avogadro's number) is $N = 6.02 \times 10^{23}$. The first precise determination of N, due to Loschmidt (1865), was based on the kinetic theory of gases.

[2]) It is well to note that in all reasoning of Statistical Mechanics, there underlies a hypothesis of a statistical nature, the hypothesis of molecular chaos, from which one cannot escape without renouncing the statistical method itself. Although this hypothesis seems intuitively correct, its rigorous justification (ergodic theorem) turned out to be particularly delicate and is still the subject of controversy.

On the basis of these equations, he foresees the existence of electromagnetic waves – a prediction ultimately confirmed in spectacular fashion by the discovery of radio waves (Hertz) – and likens the light wave to a particular kind of electromagnetic wave, thus achieving the synthesis of optics and electricity.

Toward the end of the 19th century, the success of the classical program is impressive. All known physical phenomena, it seems, find their explanation in a general theory of matter and radiation; in all the cases where this explanation could not be found, one may reasonably attribute the failure to mathematical difficulties in the solution of the problem, without jeopardizing the form of the basic equations. What is most striking in this theory is its remarkable degree of unity. The desire to unify the various branches of their science has always been one of the most fruitful preoccupations of the physicists. In fact, the physicists of that era attribute to the classical theory more unity than it actually possesses. Indeed, wave propagation is not a phenomenon peculiar to electromagnetism. The study of vibrations was first made in connection with vibrations of matter (vibrating strings, surface waves of a liquid, etc.) and the wave nature of acoustical phenomena was detected before that of light phenomena. Moreover, the existence of waves within matter by no means contradicts the corpuscular theory; in fact, one deals here with a macroscopic phenomenon which can easily be deduced from the microscopic laws of motion with a suitable law of force. By analogy, the classical physicists endow electromagnetic waves with a supporting structure, a kind of material fluid which they call "ether" and whose structure and mechanical properties remain to be specified. Thus matter appears as the fundamental entity, subject to the principles of Newton's Rational Mechanics, and obeying force laws such that it can, under suitable conditions, experience various wave phenomena, of which electromagnetic vibration is one example.

This conception, which is to be completely abandoned later, suggests, at the period we are considering, a whole series of experiments which do not reveal many facts concerning the nature of the ether; however, one of these is to provoke a rather profound upset of classical physics. This is the famous Michelson-Morley experiment (1887), designed to reveal the motion of the earth relative to the ether by seeking to detect how the velocity of propagation of light with respect to the earth varies with the direction of propagation. The negative result

of this experiment is well known. After several more or less artificial attempts at an explanation, this apparent paradox is definitely explained by Einstein in 1905 within the framework of the *Theory of Relativity*, following a critical analysis of the concepts of space and time which calls for the rejection of the notion of absolute time, and of some of the axioms of Newtonian Mechanics. In fact, the latter is but an approximation of Relativistic Mechanics, which is valid only in the limit where the velocities of the particles are negligible compared to the velocity of light c. We shall not elaborate on the principle of relativity here; we shall have occasion to return to it at the end of this book, when we study Relativistic Quantum Mechanics. The essential point to note is that this principle puts in doubt neither the doctrine nor the classical program such as they were defined earlier.

3. Progress in the Knowledge of Microscopic Phenomena and the Appearance of Quanta in Physics

At the turn of the century, the efforts of experimenters follow along two closely related lines; first, to make a precise analysis of the microscopic structure of matter; second, to determine the mutual interaction of material corpuscles, and their interactions with the electromagnetic field.

The first facts concerning the structure of matter are furnished by the study of the rays obtained from the discharge in rarefied gases, cathode rays and canal rays, which are correctly interpreted as beams of electrically charged particles travelling more or less swiftly. Thus the *electron* is discovered (J. J. Thomson, 1897), the particle of the cathode rays; its behavior in the presence of an electromagnetic field is determined experimentally, and a complete theory of the interaction between electrons and electromagnetic waves is established (electron theory of Lorentz) [1]).

Gradually, the very existence of atoms and molecules, for a long time considered to be a fruitful working hypothesis, becomes accepted reality. Its most convincing proof is furnished by the study of Brownian motion, the disordered motion of tiny particles suspended in a liquid or a gas; this motion is attributed to the frequent collisions which these particles undergo with surrounding molecules. It is to some

---

[1]) Cf. L. Rosenfeld, *Theory of Electrons* (North-Holland Publishing Co., Amsterdam, 1951).

extent a large-scale reproduction of molecular agitation and can be quantitatively related (Einstein, Smoluchowski, 1905) to the statistical laws of motion of the molecules themselves. The systematic measurements of Perrin (1908) confirm this hypothesis and furnish several precise and concordant determinations of Avogadro's number [1]). After this decisive progress, physicists no longer doubt the existence of atomic or subatomic particles, and we witness the perfection of more and more experimental techniques permitting observation of individual microscopic phenomena, or to count microscopic particles one by one (measurement of the elementary electric charge by Millikan in 1910; first observation of trajectories of charged particles with the Wilson cloud chamber in 1912; first Geiger counter in 1913). These techniques of "direct" observation have continued to develop, and constitute today nearly the entire arsenal at the disposal of experimenters in the exploration of microscopic phenomena.

However, a new chapter of physics is opened with the discovery of radioactivity (1896), the first manifestation of the properties of atomic nuclei. Important in itself, this discovery puts into the hands of the physicists a potent means for the investigation of atomic structure, namely alpha radiation, consisting of helium nuclei moving at high velocity. By exposing various targets to the alpha rays, Rutherford (1911) studies systematically the scattering of alpha particles by atoms and thus succeeds in extracting the first modern picture of the atom.

The *Rutherford atom* is formed by a central nucleus of small dimensions ($10^{-13} - 10^{-12}$ cm) around which gravitate a certain number Z of electrons. Almost all the mass of the atom is concentrated in the nucleus. The latter carries a positive electric charge Ze which exactly offsets the total charge $-Ze$ of the electrons, so as to form an electrically neutral entity. The Rutherford atom thus resembles a miniature solar system where the forces of gravity are replaced by electrical forces. Under the action of the latter, namely Coulomb attraction of the nucleus and mutual Coulomb repulsions, the electrons describe stable orbits around the nucleus, orbits whose extension is of the order of atomic dimensions, namely 10^{-8} cm.

While the corpuscular character of matter seems to be confirmed as our knowledge of atomic phenomena progresses, the spectrum of

[1]) Cf. J. Perrin, *Les Atomes* (Presses Universitaires, Paris, 1948).

known electromagnetic waves nears completion and extends toward the short wavelengths with the discovery of X-rays (Röntgen, 1895) whose wave nature is established by the experiments on diffraction by crystals (von Laue, 1912). For the sake of completeness one should mention gamma radiation from radioactive bodies as well, whose electromagnetic nature is to be recognized only much later. Figure 1

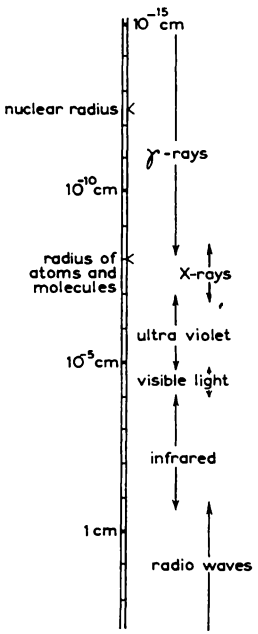


Fig. I.1. Wavelength scale of electromagnetic radiation.

gives the scale of wavelengths of electromagnetic radiation which was thus identified. While at the same time enlarging its domain, spectral analysis of radiation becomes more and more precise and allows the accumulation of a considerable body of information on the problems of emission, scattering and absorption of light by matter, in other words on the interaction of matter and radiation on the microscopic scale. The Lorentz theory of the electron which was already mentioned, a theory of charged particles interacting with the electromagnetic field, makes definite predictions concerning all these phenomena. It is precisely in the comparison of the predictions of this theory with this body of experimental results that the first

disagreements between classical theory and experiment became apparent.

The first difficulties appear when one studies the spectral distribution of electromagnetic radiation in thermodynamic equilibrium with matter. The typical case is the one of the *black body*; by definition, it is a body which absorbs all the radiation it receives. Some very general thermodynamic reasoning shows that the radiation emitted by a black body is a function of temperature only. The spectral distribution of intensity of the radiation emitted by the black body is consequently a fundamental expression which must be derivable, by the methods of Statistical Thermodynamics, from the general laws of interaction between matter and radiation. The expression deduced from the Classical Theory is in violent disagreement with experiment. In 1900, Planck succeeds in removing the difficulty by renouncing the classical law of interaction between matter and radiation [1]). He postulates that the energy exchanges between matter and radiation do not take place in a continuous manner, but by discrete and indivisible quantities or *quanta* of energy; he then shows that the quantum of energy must be proportional to the frequency ν of the radiation:

$$\varepsilon_\nu = h\nu,$$

and he obtains an expression for the spectrum in accord with the experimental distribution by properly adjusting the constant of proportionality. This constant h is known henceforth under the name of *Planck's constant*. It has the dimensions of an action (energy $\times$ time, or momentum $\times$ length). In what follows, we shall rather use the constant

$$\hbar = \frac{h}{2\pi} = 1.054 \times 10^{-27} \text{ erg-sec.}$$

Upon its publication, Planck's hypothesis seemed unacceptable; physicists almost unanimously refused to see therein more than a lucky mathematical artifice which could, some day, be explained within the framework of classical doctrine. The very success of Planck's theory could not be considered as irrefutable proof that the energy

---

[1]) For a detailed study of the theory of black-body radiation, cf. M. Born, *loc. cit.*

exchanges between matter and radiation on the microscopic scale actually take place by quanta; Planck's distribution law is a macroscopic law deduced from this hypothesis by statistical methods. It constitutes but an indirect confirmation. One could likewise doubt the validity of the quantum hypothesis itself, in the same manner as one had doubted for a long time the validity of the atomic hypothesis for lack of being able to verify it directly on the microscopic level. Planck's hypothesis, however, was to be confirmed by a whole array of experimental facts which allowed the direct analysis of the elementary processes, and the direct revelation of the existence of discontinuities in the evolution of physical systems on the microscopic scale, where classical theory predicts a continuous evolution.

II. LIGHT QUANTA OR PHOTONS

A first series of experimental facts forces a radical revision of the radiation theory of Maxwell-Lorentz, and a partial return to the old corpuscular theory. This is mainly the case with the *photoelectric effect* and the *Compton effect.*

4. The Photoelectric Effect

The first step in this direction was taken by Einstein in his celebrated note of 1905 on the photoelectric effect. The general attitude toward Planck's theory was to state that "everything behaves as if" the energy exchanges between radiation and the black body occur by quanta, and to try to reconcile this *ad hoc* hypothesis with the wave theory. Taking the opposite view, and going even further than Planck who limits himself to the introduction of the discontinuity in the absorption or emission mechanism, Einstein postulates that light radiation itself consists of a beam of corpuscles, the photons, of energy $h\nu$ and velocity c (= velocity of light in vacuo $= 3 \times 10^{10}$ cm/sec). He then shows how this surprising hypothesis may account for a certain number of hitherto unexplained phenomena, the photoelectric effect in particular.

One designates by this name the emission of electrons observed when one irradiates an alkali metal under vacuum with ultraviolet light. The electric current intensity thus produced is proportional to the intensity of the radiation striking the metal. On the other hand, the speed of the electrons does not depend upon the radiation intensity,

but only upon its frequency (Lenard, 1902), no matter how large the distance from the light source; only the number of electrons emitted per second is proportional to the intensity, hence inversely proportional to the square of the distance from the source.

Einstein's explanation is very simple. Whatever the distance covered by the light since its emission, it occurs in the form of corpuscles of energy $h\nu$. When one of these photons encounters an electron of the metal, it is entirely absorbed and the electron receives the energy $h\nu$; in leaving the metal, the electron must do an amount of work equal to its binding energy in the metal W, so that the observed electrons have a well-defined kinetic energy:

$$\tfrac{1}{2}mv^2 = h\nu - W. \tag{I.1}$$

This quantitative theory is completely verified by experiment. As expected, the constant W is a characteristic constant of the irradiated metal. As for the constant h, it has the same numerical value as the constant which occurs in the expression for the black-body spectrum of radiation.

In view of the success of the corpuscular theory, one must examine whether the classical wave theory is also capable of explaining the photoelectric effect. This is not *a priori* inconceivable. In fact, a light wave transports a certain quantity of energy proportional to its intensity and can give up all or part of this energy as it penetrates into the metal: the energy which is gradually accumulated in the metal is eventually concentrated on certain electrons who thus succeed in escaping: one can imagine that by some mechanism yet to be specified, an electron could not escape before having received a quantity of energy equal to $h\nu$. The essential difference between this type of explanation and the corpuscular theory lies in the continuous and progressive character of the energy accumulation in the metal; consequently, photoelectric emission, instead of being instantaneous, cannot take place before the metal has received the energy $h\nu$. If one operates with sufficiently fine metallic particles, this minimum delay between the start of the irradiation and the onset of the emission can be made sufficiently long to be detectable experimentally.

Experiments for this purpose were conducted by Meyer and Gerlach (1914) on metallic dusts. Knowing the intensity of the radiation and the dimensions of the dust particles, they could determine the minimum

irradiation time for a dust particle to absorb the energy $h\nu$ necessary for the emission of an electron; this amounted to several seconds in the conditions under which they were operating. In every case they observed the emission of electrons right at the start of irradiation. One must therefore conclude that the wave theory of light, at least in its classical form, is entirely incapable of accounting for the photoelectric effect.

5. The Compton Effect

The Compton effect represents another confirmation of the photon theory, and a refutation of the wave theory. One observes it (Compton, 1924) in the scattering of X-rays by free (or weakly bound) electrons. The wavelength of the scattered radiation exceeds that of the incident radiation. The difference $\Delta\lambda$ varies as a function of the angle θ between the direction of propagation of the incident radiation and the direction along which one observes the scattered light, according to Compton's formula:

$$\Delta\lambda = 2\frac{h}{mc}\sin^2\frac{\theta}{2}, \tag{I.2}$$

where m is the rest mass of the electron [1]). One notes that $\Delta\lambda$ is independent of the incident wavelength. Compton and Debye have shown that the Compton effect is a simple elastic collision between a photon of the incident light and one of the electrons of the irradiated target.

In order to discuss this corpuscular interpretation it is convenient to state a few properties of photons which derive directly from Einstein's hypothesis. Since they possess the velocity c, photons are particles of zero mass [2]). The momentum $\boldsymbol{p}$ and the energy ε of a photon are thus connected by the relation

$$\varepsilon = pc \tag{I.3}$$

---

[1]) The length $\hbar/mc$, intermediate between the mean radius of atoms and that of atomic nuclei ($\hbar/mc = 3.86 \times 10^{-11}$ cm), plays some role in the quantum theory of the electron. It is called the Compton wavelength of the electron.

[2]) According to the principle of relativity, the (rest) mass m, the energy ε and the momentum p of a particle are connected by the relation: $\varepsilon^2 - p^2c^2 = m^2c^4$; its velocity $v = \partial\varepsilon/\partial p = pc^2/\varepsilon$. If $v = c$, $\varepsilon = pc$ and $m = 0$.

Consider a plane, monochromatic light wave

$$\exp\left[2\pi i\left(\frac{\boldsymbol{u}\cdot\boldsymbol{r}}{\lambda}-\nu t\right)\right].$$

$\boldsymbol{u}$ is a unit vector in the direction of propagation, λ is the wavelength, ν the frequency: $\lambda\nu=c$. In accordance with Einstein's hypothesis, this wave represents a stream of photons of energy $h\nu$. The momentum of these photons is evidently directed along $\boldsymbol{u}$ and its absolute value, according to (I.3), is equal to

$$p=\frac{h\nu}{c}=\frac{h}{\lambda}.$$

This relation is a special case of the relation of L. de Broglie which we shall study in Ch. II. It is often convenient to introduce the angular frequency $\omega=2\pi\nu$ and the wave vector $\boldsymbol{k}=(2\pi/\lambda)\boldsymbol{u}$ of the plane wave. The connecting relations are then written:

$$\varepsilon=\hbar\omega, \qquad \boldsymbol{p}=\hbar\boldsymbol{k}. \tag{I.4}$$

The corpuscular theory of the Compton effect consists in writing down that the total energy and momentum are conserved in the elastic collision between the incident photon and the electron. Let $\boldsymbol{p}$, $\boldsymbol{p}'$ be the initial and final momenta of the photon, respectively, $\boldsymbol{P}'$ the recoil momentum of the electron after collision (Fig. I.2).

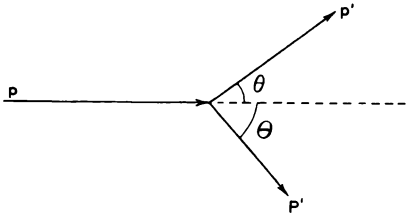


Fig. I.2. Compton collision of a photon with an electron at rest.

The conservation equations are written

$$\begin{aligned}\boldsymbol{p}&=\boldsymbol{p}'+\boldsymbol{P}',\\ mc^2+pc&=\sqrt{P'^2c^2+m^2c^4}+p'c.\end{aligned} \tag{I}$$

According to these equations the collision is completely defined once the initial conditions and the direction of emission of the scattered photon are known. Taking into account the relations (I.4), one can easily deduce the Compton formula which is thus explained theoretically (cf. Problem I.1). Since the first work of Compton, all the other predictions of this theory have been confirmed experimentally. The recoil electrons have been observed and the law of their energy variation as a function of the angle of emission Θ is just the one which one derives from equations (I). Coincidence experiments have shown that the scattered photon and electron are emitted simultaneously, and that the correlation between the emission angles θ and Θ agrees with the theory.

It is instructive to compare these results with the predictions of Classical Theory. The Maxwell-Lorentz theory predicts that part of the incident electromagnetic energy is absorbed by each irradiated electron and then re-emitted in the form of radiation of the same frequency. In contrast to the momentum of the absorbed radiation, the total momentum of the emitted radiation is zero. The light-scattering process is thus accompanied by a continuous transfer of momentum (radiation pressure) from the electromagnetic radiation to the irradiated electron; the electron is thereby continuously accelerated in the direction of propagation of the incident wave. The law of absorption and emission at the same frequency holds in the reference system where the electron is at rest. As soon as the electron is set in motion, the frequencies observed in the laboratory system are different because of the Doppler effect. The wavelength shift $\Delta\lambda$ depends upon the angle of observation θ of the scattered radiation. A simple calculation yields

$$\Delta\lambda = 2\lambda \frac{P_{\text{cl.}}\, c}{E_{\text{cl.}} - P_{\text{cl.}}\, c} \sin^2 \frac{\theta}{2}, \tag{I.5}$$

where λ is the incident wavelength, $P_{\text{cl.}}$ is the momentum of the electron (assumed in motion along the direction of propagation), $E_{\text{cl.}} = \sqrt{m^2c^4 + P_{\text{cl.}}{}^2 c^2}$ is its energy. $\Delta\lambda$ is an increasing function of $P_{\text{cl.}}$ and increases regularly during the irradiation.

Thus the classical predictions do not agree with the experimental facts. The main defect of the classical theory of the Compton effect is to predict a *continuous* transfer of momentum and energy of the

radiation to *all* the electrons exposed to the radiation, while the experimentally observed effect is a *discontinuous* and instantaneous transfer to *certain* electrons among them. Here is a difficulty of the same type as the one we encounter in the case of the photoelectric effect. The two effects are in fact related: Compton scattering may be considered as light absorption followed by re-emission, while the photoelectric effect is a mere absorption.

The introduction of light quanta is unavoidable if one wants to account for the discontinuous transfer of momentum and energy to the electrons. Nevertheless, the similarity between the classical formula (I.5) and the correct formula (I.2) of the Compton effect suggests that the classical theory retains some features of the real phenomenon. This matter merits closer examination.

The Compton effect was calculated above assuming the electron to be initially at rest. Of course, the theory remains valid when its initial velocity is not zero. One may easily generalize the equations (I) and the Compton formula. In the particular case where the electron travels parallel to the incident light wave with momentum $\boldsymbol{P}$ and energy $E=\sqrt{m^2c^4+P^2c^2}$ we obtain (Problem I.1)

$$\Delta\lambda = 2\lambda \frac{(P+p)c}{E-Pc} \sin^2 \frac{\theta}{2}. \tag{I.6}$$

Note the great similarity between this expression and the classical expression (I.5) for the shift $\Delta\lambda$. To pass from one to the other, one must replace the momentum $P_{\text{cl.}}$ in the numerator by the quantity $P+p$ (of the order of the momentum after the photon-electron collision), and in the denominator (in the expression $E_{\text{cl.}}-P_{\text{cl.}}c$) by the momentum before impact P. The mechanism to which expression (I.6) refers, however, is very different from the classical mechanism. Under the effect of prolonged irradiation, each electron receives a first momentum transfer which sets it in motion, then a second, and so on. The momentum transfer varies from one impact to the next around an average value approximately equal to the momentum $\boldsymbol{p}$ of the incident photons. We wish to compare this variation of the momentum by *discontinuous* quanta of the order $\boldsymbol{p}$ and the resulting variation of the shift $\Delta\lambda$ with the *continuous* variations predicted by the classical theory (Fig. I.3).

Such a comparison makes sense only in the limit where the quanta can be considered to be infinitely small and infinite in number, and where one considers the average effect of a very large number of successive impacts. Since the electron gains, on the average, a momentum of the order p per collision and since, after a very large

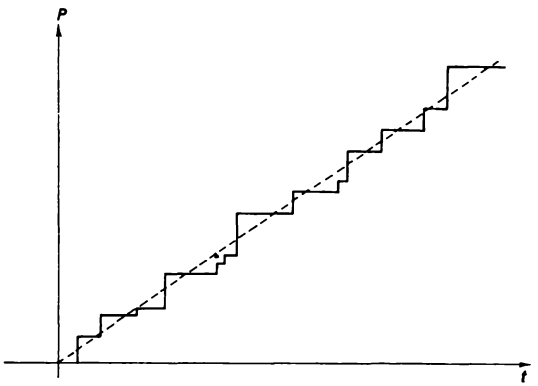


Fig. I.3. Variation as function of time of the momentum P of an electron exposed to monochromatic radiation, by reason of successive Compton collisions (this is a very schematic picture of the phenomenon whose limitations will be discussed in Ch. IV in connection with the uncertainty relations). Dashed line represents the variation $P_{cl}(t)$ predicted by classical theory.

number of impacts, the fluctuations compensate each other, the net resulting effect is practically the same as if the electron received at each impact just this average momentum. The momentum $\mathbf{P}$ of the electron therefore grows by successive jumps, in the direction of the incident radiation. These jumps are of the order of magnitude of the quantum $p=h\nu/c$ and the momentum gain may be likened to a continuous gain to the extent that the quantity p may be regarded as infinitely small. Within these limits of approximation one may define an average momentum $\langle \mathbf{P} \rangle$ growing in continuous fashion in the course of time. A thorough experimental study — which we shall not elaborate here — shows that the variation of this average momentum as a function of time is precisely the one predicted by classical theory; in other words, the vectors $\langle \mathbf{P} \rangle$ and $\mathbf{P}_{cl.}$ remain equal. Furthermore, since the classical value $P_{cl.}$, defined to within p, is equal at each instant to the average value of P, the Compton shift predicted by the

classical theory [eq. (I.5)] is equal at each instant to the average value of the actually observed Compton shift [eq. (I.6)].

6. Light Quanta and Interference Phenomena

If we have some indications that classical wave theory is macroscopically correct, it is nevertheless clear that on the microscopic scale only the corpuscular theory of light is able to account for typical absorption and scattering phenomena such as the photoelectric effect and the Compton effect, respectively. One must still ascertain how the photon hypothesis may be reconciled with the essentially wave-like phenomena of interference and diffraction.

To be definite, let us consider the scattering of a beam of monochromatic light by a parallel grating (Fig. I.4). A screen conveniently placed allows us to view the interference pattern. The quantitative observation of the phenomenon may occur in various ways, for instance by replacing the screen by a photographic plate which one develops after a given irradiation time. The interference pattern appears as a negative upon development, the blackening of each element of surface of the plate being proportional to the quantity of light received. In fact, the absorption of light by the plate occurs in quanta: each photon penetrating into the plate excites a photosensitive microcrystal which yields a black spot upon development [1]). It is not possible to resolve with the naked eye the spots from each other, and one observes in practice a continuous blackening which is more or less pronounced according to whether the density of impacts is high or low. However the existence of separate impacts may be effectively detected if one observes the plate with a sufficiently powerful microscope. Under normal experimental conditions the number of photons received by the plate is very great, and the nearly continuous distribution of impacts forms the interference pattern predicted by the wave theory.

---

[1]) The description of the phenomenon given here is greatly oversimplified. In reality, the impact of a single photon suffices to sensitize the microcrystal only if the energy of the photon is high enough (far-ultraviolet or X-rays). Even in that case one needs special experimental conditions to be able to assert with near certainty that each photon received by the plate excites one and only one microcrystal (microcrystals of appropriate size, sufficiently thick photographic plate, etc.). All these complications may be passed over without affecting the essential features of the argument given here.

Now, just on the basis of experimental observations, we can *a priori* eliminate all explanations of the phenomenon within the framework of a purely corpuscular theory. Note first of all that photons travel independently of each other, and that their mutual interaction is entirely negligible. Indeed, the interference pattern remains unchanged when one reduces the intensity of the source and increases the time of irradiation, keeping constant the quantity of light received by the grating. In other words, when one sends a certain (very large) number N of photons onto the grating, the distribution of impacts over the photographic plate is the same whether the incident photons be bunched or not; this remains true even in the limit of very weak intensities where the photons "fall upon the grating one by one". One would still obtain the same distribution upon sending a *single photon* onto the grating and repeating the experiment N times.

Let us therefore examine the problem of the scattering of one photon by the grating. Under the conditions of the experiment the initial state of the system (photon + grating) is not known exactly; consequently, the corpuscular picture does not lead to a unique trajectory for the photon, nor, *a fortiori*, to a unique impact of the scattered photon upon the plate, but only to a statistical distribution of possible trajectories and a statistical distribution of impacts. Experimentally, one actually observes a statistical distribution of the

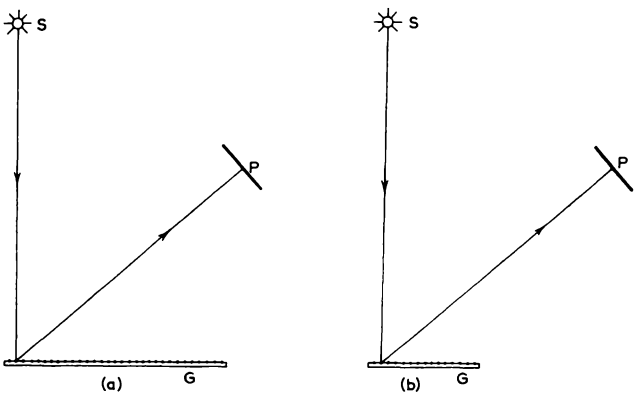


Fig. I.4. Scattering of light by a grating: S = light source; G = grating; P = photographic plate. In the case (*b*), the right-hand half of G has been omitted.

impacts: the latter is, to within a constant, the distribution law of the light intensities making up the interference pattern on the plate. Now, as is well known, the resolving power of a grating depends upon the number of rulings (assumed to be equidistant) and the light fringes are the narrower the larger this number; thus one modifies the interference pattern noticeably when one eliminates one half of the grating. On this very point the corpuscular theory is in clear contradiction with experiment. Indeed, no matter what the equations of motion of each photon interacting with the particles constituting the grating might be, the distribution of trajectories of the photons scattered by the left-hand half of the grating cannot depend upon the presence (Fig. I.4*a*) or the absence (Fig. I.4*b*) of the right-hand half unless one were to attribute to the photon dimensions of the order of those of the grating. If, furthermore, the light source and photographic plate are sufficiently far from each other, the distributions of impacts of the photons scattered by each half are essentially the same. Upon suppressing the right-hand half of the grating, the light intensity received at each point of the plate must thus be reduced without any other modification of the interference pattern. Experiment contradicts these predictions and forces us to assume that the entire grating takes part in the scattering process.

The hypothesis of light particles encounters difficulties of the same kind in all cases of interference or diffraction (Problem I.2). One can effectively detect the arrival of photons, one by one, at the detecting apparatus (screen, photographic plate, or any other more elaborate arrangement), but it is not possible to attribute a precise trajectory to each photon without running into contradictions. The classical doctrine, according to which a corpuscle travels through space in a continuous manner as a function of time, is found to fail. On its way up to the detecting apparatus, everything happens as if light were propagated as a wave; its corpuscular aspect manifests itself only at the instant of detection.

7. Conclusions

We can thus draw a certain number of preliminary conclusions from the experimental results bearing on the interaction between matter and light on the microscopic scale.

Even though the observed discontinuities can properly be explained only with the picture of light corpuscles, it is nevertheless incon-

ceivable to renounce the wave concept. Light presents itself in two forms: wave and corpuscle, each of these aspects appearing in more or less clear-cut fashion depending upon the phenomenon under consideration. The connecting relations (I.4) allow us to pass from one mode of description to the other. The very close bond between these two modes is a *statistical bond* as the discussion of the experiment on scattering by a grating shows: the probability of localizing the photon at a given point is equal to the intensity of the light wave at that point, calculated by the methods of wave optics. *The existence of the wave-corpuscle duality is incompatible with classical doctrine.* It is impossible to consider light either as a stream of classical corpuscles, or as a superposition of classical waves, without entering into contradiction with the experimental facts.

Faced by the prospect of a revision of the classical doctrine, it is particularly important to stress the results of classical wave theory which remain valid. In the first place, *the laws of conservation of momentum and of energy remain rigorously verified.* Furthermore, as we have seen in connection with the Compton effect, *classical theory correctly predicts the average evolution of phenomena in the "macroscopic limit"* where the quantum discontinuities can be treated as infinitely small.

III. QUANTIZATION OF MATERIAL SYSTEMS

8. Atomic Spectroscopy and Difficulties of Rutherford's Classical Model

We have seen in the preceding section how the existence of discontinuities in the interaction mechanism between matter and radiation upsets the Classical Theory of Light. But the upheaval is not limited to light, and equally affects the classical corpuscular theory of matter. This becomes clear when one seeks to reconcile the facts of atomic spectroscopy with the results bearing on the structure of atoms obtained by Rutherford [1]).

[1]) Historically, the first argument showing the necessity of "quantizing" material systems was presented by Einstein in the theory of the specific heat of solids (1907). This theory contains rather crude approximations, which are inevitable in the treatment of a material system which is as complex as a solid body. Furthermore, this theory brings into play results of statistical thermodynamics, similarly to the theory of black-body radiation to which it is closely related. For all these reasons, we shall not enter into detail here, and refer the reader to works such as M. Born, *Atomic Physics, loc. cit.*

One of the most salient features brought to light by the refinement of the study of emission and absorption spectra of light by matter, is the existence of narrow spectral lines. The frequencies of emitted or absorbed radiation vary from one atom to the next; the absorption and emission spectra of a given atom, however, are the same; every line of one spectrum can be observed in the other if one operates under suitable conditions. Each atom may be identified by this spectrum, which therefore constitutes an essential element of information on the structure of the atom and its interaction mechanism with radiation.

The case of hydrogen deserves special mention here since it is the simplest of all atoms (one proton + one electron); all the observed frequencies obey Balmer's empirical formula:

$$\nu = R\left(\frac{1}{n^2} - \frac{1}{m^2}\right),$$

where n and m are positive integers ($m > n$) and R is a numerical constant characteristic of hydrogen (the Rydberg constant).

For more complex atoms there are no equally simple formulae, but there is always a certain correlation among the various observed frequencies: when two frequencies belong to the same spectrum, it often happens that their sum or their difference also occurs in the spectrum. More precisely, one can set up for each atom a table of numbers or spectral terms chosen in such a manner that any frequency of its spectrum is equal to the difference of two of the term values. This rule, of which Balmer's formula is a special case, is named the *Rydberg-Ritz combination principle* (1905). Conversely, all the differences thus formed are not necessarily frequencies of the spectrum, but it is possible to formulate relatively simple *selection rules* which allow us to distinguish those differences which do occur in the spectrum from those which do not.

These experimental facts are in clear-cut disagreement with the classical radiation theory of the Rutherford atom; indeed, the Rutherford model itself encounters serious contradictions when, instead of limiting oneself to the Coulomb interaction, one takes rigorously into account the interaction of the atomic electrons with the electromagnetic field in accordance with the electron theory of Lorentz. Indeed, the electrons radiate while describing their orbits, progressively lose energy, and finally spiral into the nucleus. At each instant, the observed frequencies in the emitted radiation are equal to the frequency of the

orbital motion or to one of its harmonics; as the latter varies during the slowing-down process, a *continuous* light spectrum is emitted. The classical theory of the Rutherford atom therefore explains neither the stability of atoms, nor the existence of line spectra. We are confronted with a new manifestation of discontinuities in the interaction of matter and light, where the classical theory predicts a continuous variation.

9. Quantization of Atomic Energy Levels

In 1913, Bohr obtains a general scheme for the explanation of spectra by completing the quantum hypothesis of light by a new postulate incompatible with classical notions: the quantization of the energy levels of atoms.

According to Bohr, the atom does not behave as a classical system capable of exchanging energy in a continuous manner. It can exist only in a certain number of *stationary states* or quantum states each having a well-defined energy. One says that the energy of the atom is quantized. It can vary only by jumps, each jump corresponding to a transition from one state to another.

This postulate allows us to specify the mechanism of absorption or emission of light through quanta. In the presence of light an atom of energy E_i may undergo a transition to a state of higher energy $E_j (> E_i)$ by absorbing a photon $h\nu$ provided that the total energy is conserved, namely

$$h\nu = E_j - E_i.$$

Similarly, the atom can undergo a transition to a state of lower energy $E_k (< E_i)$ by emitting a photon $h\nu$ whose frequency satisfies the relation

$$h\nu = E_i - E_k.$$

If the atom finds itself in its lowest energy state (ground state) it cannot radiate and remains stable.

In this way an explanation is found for the existence of spectral lines characteristic of each atom and satisfying the Rydberg-Ritz combination principle: the spectral terms are equal, to within a factor of h, to the energies of the quantum states of the atom. In particular, for the case of the hydrogen atom, one rediscovers the

Balmer formula by assuming that the energy levels are given by the formula

$$E_n = -h\frac{R}{n^2} \qquad (n = 1, 2, 3, \ldots, \infty). \tag{I.7}$$

Another confirmation of the quantization of atomic energy levels is furnished by the experiment of Franck and Hertz on the inelastic collisions between electrons and atoms (1914). The experiment consists in bombarding atoms by monoërgic electrons and in measuring the kinetic energy of the scattered electrons. From this one deduces by subtraction the quantity of energy absorbed in the collision by the atoms. Let $E_0, E_1, E_2, \ldots$ be the sequence of quantized energy levels of the atoms, T the kinetic energy of the incident electrons. Under the conditions of the experiment, the atoms of the target are practically all in their ground state. As long as T lies below the difference $E_1 - E_0$ between the energy of the ground state and that of the first excited state, the atom cannot absorb energy and all collisions are elastic. As soon as $T > E_1 - E_0$, inelastic collisions can occur in which the electron loses a quantity of energy equal to $E_1 - E_0$ and the atom goes into its first excited state. This is exactly what is found experimentally. One similarly observes collisions with excitation of the second excited state as soon as $T > E_2 - E_0$, and so on.

Hence the quantization of atomic energy levels appears as experimental fact. This property is not peculiar to atoms. Progress of experimentation, especially in the field of spectroscopy, has shown that quantization is found in the case of molecules and of more complex systems of particles as well. We thus face a very general property of matter which classical corpuscular theory is unable to explain.

10. Other Examples of Quantization: Space Quantization

Another type of experimentally observed quantization is that of the quantization of orientation, or "space quantization" of atomic systems. One observes it every time the atom is situated in an external field possessing a preferred direction. The orientation of the atomic system is not arbitrary, but limited to certain discrete values.

The most direct confirmation of this type of quantization is furnished by the *Stern-Gerlach experiment* (1922) on the deviation of paramagnetic atomic beams (or molecular beams) in an inhomogeneous

magnetic field. Paramagnetic atoms are by hypothesis endowed with a permanent magnetic moment $\boldsymbol{\mu}$, and can be considered as little elementary gyroscopes of angular momentum $\boldsymbol{l}$ proportional to $\boldsymbol{\mu}$:

$$\boldsymbol{\mu} = M\boldsymbol{l}.$$

The orientation of $\boldsymbol{\mu}$ and $\boldsymbol{l}$ defines the orientation of the atom itself. In a magnetic field $\boldsymbol{H}$, the angular momentum executes a precessional motion about $\boldsymbol{H}$ (Larmor precession, cf. Problem I.3). If $\boldsymbol{H}$ is constant, the magnetic energy $-\boldsymbol{\mu}\cdot\boldsymbol{H}$ remains constant and independent of the position of the center of mass of the atom, and the latter remains in uniform rectilinear motion. If $\boldsymbol{H}$ is not constant, the center of mass of the atom is subject to the force $\boldsymbol{F} = \text{grad}\,(\boldsymbol{\mu}\cdot\boldsymbol{H})$ and suffers a certain deflection. This is exactly what one observes in the Stern-Gerlach experiment which is shown schematically in Fig. I.5. Because of the precessional motion about the field $\boldsymbol{H}$, the component μ_z of $\boldsymbol{\mu}$ along the field remains constant while the other components oscillate about zero. Everything takes place as if the atom were subject to the value of the force averaged over several oscillations: $\mu_z\,\text{grad}\,\boldsymbol{H}_z$. Under normal experimental circumstances this average force is directed along Oz and equals $\mu_z(\partial\boldsymbol{H}_z/\partial z)$. Let $2l$ be the distance in the magnetic field traversed by the atom, T the kinetic energy of the atoms in the incident beam; a simple calculation shows that the velocity of each atom is deflected from its initial direction Ox through an angle $\approx\mu_z(\partial\boldsymbol{H}_z/\partial z)(l/T)$. The deflection is thus proportional to the component of $\boldsymbol{\mu}$ in the direction of the field. If the atoms were randomly oriented, μ_z could take on all values between $-\mu$ and $+\mu$ and the deflection angle all values between the corresponding two extreme values. The impacts of the atoms on the screen should then form a single spot elongated in the direction Oz. What one actually observes is a series of small spots, equidistant and aligned parallel to Oz. If one varies the field [and hence $(\partial\boldsymbol{H}_z/\partial z)$] the mutual distance of the spots varies correspondingly, without any other modification of the pattern; in particular the number of spots λ remains constant. Each of the spots corresponds to a definite value of μ_z. Consequently, μ_z is a quantized quantity taking on λ different values. The component l_z of the angular momentum evidently possesses the same property.

Against this interpretation of the Stern-Gerlach experiment one may raise the objection that it is based on a very particular hypothesis

concerning the origin of atomic paramagnetism, namely the existence of a permanent magnetic moment proportional to the angular momentum. We shall not elaborate here on the facts and arguments

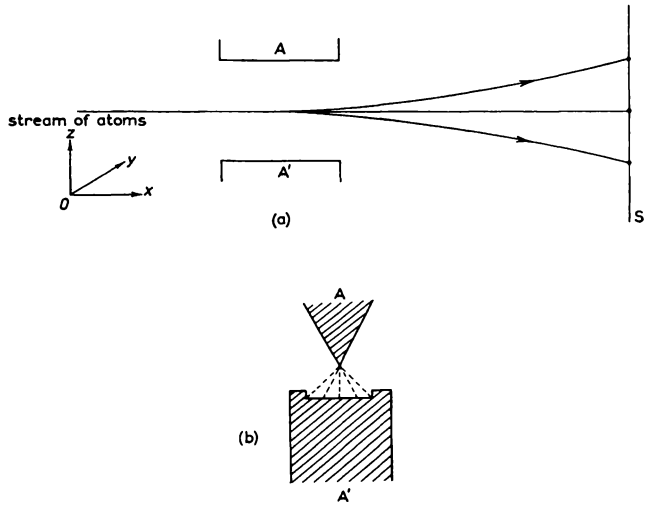


Fig. I.5. The Stern–Gerlach experiment.

(*a*) General scheme of the experiment: the atomic beam passes between the pole pieces AA′ of the magnet where an inhomogeneous magnetic field exists (directed vertically in the figure); the impacts of the atoms are observed on the screen S.

(*b*) Section through the pole pieces of the magnet; dashed lines represent the lines of force of the magnetic field.

which justify such a hypothesis (gyromagnetic effects, Langevin's theory of paramagnetic susceptibility, etc.) which later developments of Quantum Mechanics have amply confirmed. Even if one doubts the details of the explanation given above, one can hardly account for the existence of λ distinct spots on the screen without assuming that certain quantities characteristic of the internal motions are quantized. Indeed, to the extent that the center-of-mass motion follows the laws of Classical Mechanics, its trajectory is entirely determined by the dynamical state of the atom at the entrance of the magnet. The appearance on the screen of a more or less spread-out distribution of impacts indicates that the atoms are not all in the

same initial conditions and that the dynamical variables defining the initial state are statistically distributed over a somewhat extended domain. The existence of λ distinct spots attests to the fact that this statistical distribution presents at least **e** discontinuities; in other words, certain dynamical variables of the atom are quantized. Since the atoms are practically all in their ground state (otherwise they would radiate) one cannot be dealing here with a quantization of the energy and, since the effect observed on the screen is directional, the dynamical variable of the atom whose quantization is thus made evident depends upon the orientation of the atom.

In addition to the Stern-Gerlach experiment, there exist quite a few other less direct manifestations of space quantization. In particular we mention the effect of a constant magnetic field on the structure of spectra, or Zeeman effect (1896), to which we shall return later on. All these phenomena have a common origin, namely the quantization of angular momentum; later developments of Quantum Mechanics will bring this out quite clearly.

IV. CORRESPONDENCE PRINCIPLE AND THE OLD QUANTUM THEORY

11. Inadequacy of Classical Corpuscular Theory

The quantization of certain physical quantities – and we have to insist on that point – is an experimental fact incompatible with the classical corpuscular theory of matter. Thus the energy of a system of classical particles is an essentially continuous quantity; no matter how one modifies the force law, not even by introducing additional dynamical variables, can one change this situation: the fact that the energy of a system of particles is limited to a discrete set of allowed values is a result which falls outside Classical Mechanics. The same remarks apply to all other quantized quantities.

Correlated with this, *the evolution in time of a quantized quantity is impossible to describe in strictly classical terms.* Let us take the example of an atom finding itself initially in its first excited state E_1 and falling back to the ground state with emission of a photon. If we adopt the language of classical physics and try to define the evolution in time of the energy of the atom, one must assume that

this energy undergoes a discontinuous jump from E_1 to E_0 at a certain instant, since any continuous evolution of the energy between these two values is excluded. Yet it is not possible to predict at what precise instant this jump will occur. Indeed, if *the dynamical state of the atom remains identically the same* during the entire period of time preceding the jump, there is no reason why the latter should occur at any one given instant rather than at another. At best one can speak of the probability per unit time of the occurrence of the jump. In reality, classical physics is unable to describe this situation, and the very picture of a jump occurring at a precise instant is incorrect. One must give up imagining an exact evolution of the energy as a function of time. The only thing one may define is the probability that an initially excited atom be found in its ground state at a given later time. We shall see later on that the probability of de-excitation – as well as the law of disintegration of a radioactive nucleus – follows a decreasing exponential law whose characteristic constant is the de-excitation probability per unit time, otherwise stated, the reciprocal of the mean life of the excited state.

We must therefore, at the price of renouncing certain classical concepts, incorporate this phenomenon of quantization into a coherent theory of matter, from which one might deduce the precise numerical values of the quantized quantities, and the quantities pertaining to the various quantum transitions which are possible, as for instance, the mean lives of excited states which were mentioned above. This program will not be fully realized until the establishment of Quantum Mechanics in its modern form. Previously, Bohr and his school (Kramers, Sommerfeld) formulated a first draft of a quantum theory from which the spectral terms of hydrogen-like systems could be predicted correctly. In spite of the difficulties of principle, and the limitations of this Old Quantum Theory, it is useful to know its main features in order to properly appreciate the later development of the theory. Furthermore, this older theory represents a first example of the application of a heuristic principle which played an essential role in the development of Quantum Mechanics: the correspondence principle. It is on this point that we shall especially focus our attention in the following presentation of the Old Quantum Theory. The latter was completed by a semi-classical theory of the interaction between matter and radiation also based on the correspondence principle; we shall not discuss it in this book.

12. Correspondence Principle

The correspondence principle was not clearly formulated by Bohr until 1923 [1]), but it inspired all previous work. It consists in stating precisely to what extent the notions and the results of Classical Mechanics can serve as guides in the elaboration and interpretation of the correct theory.

We have already discussed the domain of validity of the Classical Theory of radiation in connection with light quanta. What has been said is valid for Classical Theory in general. The latter predicts correctly a vast array of phenomena, from the macroscopic scale down to and including certain phenomena of the microscopic domain; among the latter let us mention the motion of electrons in a static electromagnetic field, the thermal motion of the atoms or molecules of a gas, etc. The major difficulty of Classical Theory in explaining phenomena on the microscopic scale stems from the appearance of discontinuities on that scale.

One may therefore assert that *Classical Theory is "macroscopically correct"*, that is to say, it accounts for phenomena in the limit where quantum discontinuities may be considered infinitely small; in all these limiting cases, the predictions of the exact theory must coincide with those of Classical Theory. We have here a very restrictive condition imposed upon the Quantum Theory. One often expresses it in abbreviated form by saying that: *Quantum Theory must approach Classical Theory asymptotically in the limit of large quantum numbers.*

In order that this condition might be fulfilled, one establishes in principle *that there exists a formal analogy between Quantum Theory and Classical Theory*; this "correspondence" between the two theories persists down to the smallest details and must serve as guide in the interpretation of the results of the new theory.

13. Application of the Correspondence Principle to the Calculation of the Rydberg Constant

Let us verify that expression (I.7), giving the energy levels of the hydrogen atom as a function of the quantum number n, is compatible with the correspondence principle, and let us show that the application of this principle unambiguously determines the numerical value of the constant R which occurs in that expression.

---

[1]) N. Bohr, Z. Physik 13 (1923) 117.

According to the classical theory of Rutherford, the hydrogen atom is composed of an electron and a proton in Coulomb interaction [potential $-(e^2/r)$]. In accordance with Kepler's laws, which we shall assume to be known to the reader, the electron describes an elliptical orbit about the proton, the latter (assumed infinitely heavy) being at its focus. To each orbit there corresponds a certain value of the energy $E(<0)$ and a frequency $\nu_{\text{cl.}}$ of the electron motion along that orbit. These quantities in fact depend only upon the length of the major axis of the ellipse; they are related by the expression:

$$\nu_{\text{cl.}}(E) = \frac{1}{\pi e^2}\left(\frac{2|E|^3}{m}\right)^{\frac{1}{2}} \tag{I.8}$$

(m = mass of the electron).

During the course of this motion, the electron emits a certain radiation made up of a superposition of monochromatic waves whose frequencies are equal to $\nu_{\text{cl.}}$ or one of its harmonics; this radiation is the richer in high harmonics the more eccentric the elliptical orbit. It is emitted in continuous fashion and is accompanied by a steady decrease of the energy E.

This should be compared with the energy degradation by discrete jumps in the Bohr theory. When n is very large, the distance of the level E_n to each of its nearest neighbors is a certain integral multiple of $\mathrm{d}E/\mathrm{d}n = 2Rh/n^3$; for all optical transitions where the relative variation $\Delta n/n$ of the quantum number is very small, the emitted frequency is, just as in classical theory, the harmonic (of order $\Delta n - 1$) of a certain fundamental frequency

$$\nu_{\text{qu.}} \approx 2\frac{R}{n^3} = 2\left(\frac{|E_n|^3}{Rh^3}\right)^{\frac{1}{2}}. \tag{I.9}$$

In the limit where n is very large, the energy E_n is degraded on the average by a succession of small and numerous quantum jumps, and the spectrum of emitted frequencies (more precisely, the low-frequency portion of that spectrum which is associated with quanta of the lowest energy) must be identical to the classical spectrum, in accordance with the correspondence principle. In other words,

$$\nu_{\text{qu.}} \underset{n\to\infty}{\sim} \nu_{\text{cl.}}(E). \tag{I.10}$$

Upon examining expressions (I.8) and (I.9) one sees that this condition

may be actually fulfilled if one takes

$$R = \frac{2\pi^2 m e^4}{h^3}. \tag{I.11}$$

The experimental value of R is known with extreme precision ($\approx 10^{-6}$). The theoretical value (I.11) agrees with the former to within less than one part in 10^4 [1]). We have here one of the most spectacular successes of the Bohr theory.

The latter is easily extended to hydrogen-like atoms formed by an electron and a nucleus of charge Ze, in particular to the singly ionized helium atom ($Z = 2$). One has merely to replace e^2 by Ze^2 in all formulae. The spectral terms of He^+ thus obtained coincide, with the same extraordinary precision of 1 in 10^4, with those observed experimentally.

14. Lagrange's and Hamilton's Forms of the Equations of Classical Mechanics

In view of later discussions of the formal correspondence between Quantum Theory and Classical Theory, it is well to recall some points of classical analytical mechanics.

Quite generally, the dynamical state of a classical system is defined by its position – in terms of its coordinates $q_1, q_2, \ldots, q_R$, – and its velocity – defined by the derivatives $\dot{q}_1, \dot{q}_2, \ldots, \dot{q}_R$ of its coordinates with respect to time. R is the number of degrees of freedom of the system [2]). If we are dealing with a system of n particles, one may choose as position coordinates the $3n$ cartesian coordinates of these particles, but the considerations which follow apply equally well to other choices of coordinates. The position of the system can be represented at each instant in a space with R dimensions, *configuration space*, by a point M having $q_1, q_2, \ldots, q_R$ as coordinates in that space.

---

[1]) In order to claim such a precision in the determination of R, it is necessary to take into account the fact that the mass of the proton M is finite. To do this one merely replaces in formula (I.11) the mass m by the reduced mass $m' = mM/(m + M)$. Having taken this correction into account ($\approx 5 \times 10^{-4}$), the theoretical value of R lies slightly below the experimental value. The difference is due essentially to relativistic effects, which amount to a slight increase of the mass m'.

[2]) We only consider here systems without constraints; in other words, the q's may vary independently of each other without any limitation.

Classical Mechanics has as its objective the formulation of the laws of evolution of the system as a function of time or, if one prefers, the laws of motion of its representative point M in configuration space.

For a large number of dynamical systems – the only ones we shall have to consider here – one can write down the laws of motion by introducing a certain characteristic function of the system, the Lagrange function (or Lagrangian):

$$L \equiv L(q_1, q_2, \ldots, q_R;\ \dot{q}_1, \dot{q}_2, \ldots, \dot{q}_R;\ t).$$

The coordinates q will satisfy R second-order differential equations (Lagrange's equations):

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_r}\right) - \frac{\partial L}{\partial q_r} = 0 \qquad (r = 1, 2, \ldots, R).$$

The quantities

$$p_r \equiv \frac{\partial L}{\partial \dot{q}_r} \qquad (r = 1, 2, \ldots, R)$$

which occur in these equations, are called the Lagrange *conjugate momenta*. In the case where q_r is one of the cartesian coordinates of a particle of mass m and the forces are derivable from a static potential, p_r is the corresponding coordinate for the momentum of that particle: $p_r = m\dot{q}_r$.

The laws of motion may be expressed equally well in the form of a variational principle. The system of Lagrange's equations is, in fact, equivalent to the *principle of least action* (Maupertuis-Hamilton):

$$\delta \int_{t_1}^{t_2} L \,\mathrm{d}t = 0, \qquad \delta M(t_1) = \delta M(t_2) = 0 \tag{I.12}$$

whose significance is as follows: of all the laws $M(t)$ allowing the system to pass from position M_1 at time t_1, to position M_2 at time t_2, the law of motion which is actually realized is the one making the integral $\int_{t_1}^{t_2} L \,\mathrm{d}t$ stationary.

Another, particularly useful form of the laws of Classical Mechanics is *Hamilton's canonical form*. We note that the dynamical state of a classical system at a given instant is completely defined by giving its R position coordinates $q_1, q_2, \ldots, q_R$ and the R corresponding conjugate momenta $p_1, p_2, \ldots, p_R$. It is convenient to introduce a space with

$2R$ dimensions, the *phase space*, where such a dynamical state is represented by a point P having the q's and p's as coordinates. Now let us define the classical Hamiltonian:

$$H \equiv H(q_1, \ldots, q_R;\ p_1, \ldots, p_R;\ t) = \sum_{r=1}^{R} \dot{q}_r \frac{\partial L}{\partial \dot{q}_r} - L. \qquad \text{(I.13)}$$

The equations of motion may be written in the following canonical form

$$\dot{q}_r \equiv \frac{\partial H}{\partial p_r}, \qquad \dot{p}_r \equiv -\frac{\partial H}{\partial q_r} \qquad (r = 1, 2, \ldots, R). \qquad \text{(I.14)}$$

These are first-order differential equations. Knowledge of the coordinates and momenta at the initial instant suffices to determine their values at all later instants. Hence when H does not depend upon time, one and only one trajectory passes through each point of phase space, representing a possible motion of the system.

In the most common cases, L is the difference between the kinetic energy T, a homogeneous quadratic function of the $\dot{q}$, and the potential energy V; $H = T + V$ is the total energy of the system expressed as a function of the q's and p's. However, the formalisms of Lagrange and of Hamilton apply to more general dynamical systems (cf. Problem I.4). By extension, one agrees in all cases to consider H as the total energy of the system. From Hamilton's equations one deduces that $\dot{H} \equiv (\mathrm{d}H/\mathrm{d}t) = (\partial H/\partial t)$; in other words, if the Hamiltonian does not explicitly depend upon the time, the total energy of the system is a constant of the motion. One says that the system is conservative.

In order to illustrate these points, let us consider an electron in the Coulomb field of a proton (considered infinitely heavy). Let $\boldsymbol{r}(x, y, z)$ be its position in a system of cartesian axes centered at the proton, $\boldsymbol{v} = \mathrm{d}\boldsymbol{r}/\mathrm{d}t$ its velocity, and $\boldsymbol{p}(p_x, p_y, p_z)$ its momentum. The Lagrangian is

$$L = \tfrac{1}{2}mv^2 + \frac{e^2}{r}.$$

The conjugate momenta of the coordinates are the respective components of the momentum ($p_x = \partial L/\partial v_x$, $p_y = \partial L/\partial v_y$, $p_z = \partial L/\partial v_z$). From the explicit form of the Hamiltonian function:

$$H \equiv \frac{p^2}{2m} - \frac{e^2}{r},$$

one obtains Hamilton's equations:

$$\frac{d\boldsymbol{r}}{dt} = \frac{\boldsymbol{p}}{m}, \qquad \frac{d\boldsymbol{p}}{dt} = \operatorname{grad} \frac{e^2}{r} = -e^2 \frac{\boldsymbol{r}}{r^3}.$$

With these equations one verifies directly that the angular momentum, $\boldsymbol{l} = \boldsymbol{r} \times \boldsymbol{p}$, is a constant of the motion: $d\boldsymbol{l}/dt = 0$, a consequence of the central character of the potential $-e^2/r$, and that the electron trajectory is entirely situated in the plane containing the origin and perpendicular to this constant vector $\boldsymbol{l}$.

In the same way one obtains the equations of motion using any other system of coordinates. For the trajectories located in the xy plane ($z = \dot{z} = 0$), one obtains in polar coordinates (r, φ):

$$x = r \cos \varphi, \qquad y = r \sin \varphi,$$

$$L = \tfrac{1}{2}m[\dot{r}^2 + (r\dot{\varphi})^2] + \frac{e^2}{r}, \qquad p_r = m\dot{r}, \qquad p_\varphi = mr^2\dot{\varphi},$$

$$H = \frac{1}{2m}\left(p_r^2 + \frac{p_\varphi^2}{r^2}\right) - \frac{e^2}{r},$$

whence Hamilton's equations:

$$\begin{aligned} \dot{p}_\varphi &= 0, & \dot{\varphi} &= \frac{p_\varphi}{mr^2}, \\ \dot{p}_r &= \frac{p_\varphi^2}{mr^3} - \frac{e^2}{r^2}, & \dot{r} &= \frac{p_r}{m}. \end{aligned} \tag{I.15}$$

p_φ is the magnitude of the angular momentum: it is, indeed, a constant of the motion.

15. Bohr-Sommerfeld Quantization Rules

The Old Quantum Theory is essentially a general method of calculating quantized quantities, based upon the hypotheses of Bohr and the correspondence principle. The procedure is the following. One assumes that the systems of material particles follow the laws of Classical Mechanics. One further postulates that, of all the solutions of the equations of motion, one must only retain those which satisfy certain *ad hoc* quantization rules. One thus selects a discontinuous family of motions; these are, by hypothesis, the only motions which can actually be realized. To each of these motions there corresponds a certain value of the energy. The discontinuous sequence of energy

values thus obtained is the spectrum of quantized energy levels. One likewise obtains a discrete spectrum of allowed values for all other constants of the motion.

The determination of the "quantization rules" constitutes the central problem of this Old Quantum Theory. It is above all a matter of intuition: one postulates certain rules and compares the spectra of quantized quantities derived therefrom with the experimental results. In this search, the correspondence principle serves as a valuable guide.

There exists a very simple case where this principle leads to the result in a very natural way, namely the case where the classical motions are periodic motions whose frequency is a function of the energy alone,

$$\nu_{\mathrm{cl.}} = \nu_{\mathrm{cl.}}(E).$$

In particular, this is the case in the hydrogen atom [cf. eq. (I.8)]. Let $E_1, E_2, \ldots, E_n, \ldots$, be the sequence of quantized energies. One may always consider the energy of the system to be a continuous function $E(n)$ of the quantum number n, the quantization of energy then resulting from the fact that n can take on only integral values. Upon repeating the reasoning of § 13 for the calculation of the Rydberg constant, we obtain the relation for the correspondence between classical and quantum frequency [cf. eq. (I.10)]:

$$\frac{1}{h}\frac{\mathrm{d}E}{\mathrm{d}n} \underset{n\to\infty}{\sim} \nu_{\mathrm{cl.}}(E),$$

from which we obtain the quantization rule

$$\int^{E} \frac{\mathrm{d}E}{\nu_{\mathrm{cl.}}(E)} = nh + \text{constant},$$

for large values of the integer n. It is natural to extend this rule to all values of n and to put

$$\int_{E_{\min}}^{E} \frac{\mathrm{d}E}{\nu_{\mathrm{cl.}}(E)} = nh \qquad (n = 1, 2, \ldots, \infty). \tag{I.16}$$

($E_{\min}$ is the minimum value of the energy of the classical system.) In the case of the hydrogen atom, this quantization rule yields just the terms of the Balmer formula.

This rule applies equally to periodic systems with a single degree of freedom. In that case it is possible to put it into a form more suited to generalizations. Let q be the position coordinate of such a system, p its canonically conjugate momentum, and $H(q, p) = E$ the total energy. The two-dimensional phase space and the periodic motions are represented by closed curves $H(q, p) = \text{const.}$ in this space [1]). Using Hamilton's equations, it is possible to show that

$$\int_{E_{\min}}^{E} \frac{\mathrm{d}E}{\nu(E)} = \oint_{H=E} p \, \mathrm{d}q,$$

where the symbol $\oint_{H=E}$ implies that one must integrate over a complete period of the motion corresponding to the energy E (the integral $\oint p \, \mathrm{d}q$ is called the action integral). One thus obtains the quantization rule equivalent to the rule (I.16):

$$\oint p \, \mathrm{d}q = nh \qquad (n = 1, 2, \ldots, \infty). \tag{I.17}$$

It determines the allowed trajectories of phase space and the corresponding quantized energies. This rule is known as the *Bohr-Sommerfeld quantization rule.*

Wilson and Sommerfeld have generalized this rule to multiply-periodic systems. These are systems with several degrees of freedom whose motions can be represented by a sequence of functions $p_1(q_1), p_2(q_2), \ldots, p_R(q_R)$, for a suitable choice of coordinates $q_1, q_2, \ldots, q_R$ and of conjugate momenta $p_1, p_2, \ldots, p_R$; in other words, the trajectories in phase space are such that each momentum p_r is a function of the coordinate q_r only, to the exclusion of all other variables. Each function $p_r(q_r)$ represents a periodic motion of frequency ν_r; the motion of the system results from the combination of periodic motions of respective frequencies $\nu_1, \nu_2, \ldots, \nu_R$. In that case, the quantization rules are the R relations

$$\oint p_r \, \mathrm{d}q_r = n_r h \qquad (r = 1, 2, \ldots, R). \tag{I.18}$$

The R (integral) quantum numbers $n_1, n_2, \ldots, n_R$ define the quantized trajectories of the system, and the quantized values of the various

[1]) If q is a cyclic variable (for instance an angular variable), in other words if the values of q differing by an integral multiple of a certain period Q represent the same configuration of the system, the periodic motion in phase space is not represented by a closed curve but by a curve of period Q.

constants of the motion such as the energy, angular momentum, etc. In particular, the energy $E(n_1, n_2, \ldots, n_R)$, considered as a function of the variables $n_1, n_2, \ldots, n_R$, fulfills the conditions of correspondence

$$\frac{\partial E}{\partial n_r}_{n_r \to \infty} \sim h\nu_r \qquad (r = 1, 2, \ldots, R).$$

By way of an application we shall briefly consider the quantization of the hydrogen atom. Once we have chosen the plane of the electron orbit, this is a two-dimensional problem whose equations in polar coordinates we have already written down [eq. (I.15)]. The angular momentum and the energy are constants of the motion. If one fixes the respective values $L(>0)$ and $E(<0)$ of these two quantities, one obtains a possible trajectory of the classical motion: it is an ellipse of eccentricity

$$\sqrt{1 + \frac{2L^2E}{me^4}}.$$

The momenta p_φ and p_r are functions of their respective conjugate coordinates. Indeed

$$p_\varphi = L, \qquad \frac{1}{2m}\left(p_r^2 + \frac{L^2}{r^2}\right) - \frac{e^2}{r} = E.$$

We can thus apply the Bohr-Sommerfeld quantization rules:

$$\oint p_\varphi \, \mathrm{d}\varphi = lh, \qquad \oint p_r \, \mathrm{d}r = kh,$$

where l, the azimuthal quantum number, and k, the radial quantum number, are positive (or zero) integers. The first rule yields the quantized value of the angular momentum:

$$L = l\hbar.$$

The second rule gives after a somewhat lengthy but straightforward calculation,

$$\sqrt{\frac{2\pi^2 me^4}{-E}} - 2\pi L = kh,$$

from which, introducing the "principal quantum number" $n = l + k$, one extracts the Balmer formula:

$$E_n = -\frac{me^4}{2\hbar^2 n^2},$$

with the same value for the Rydberg constant as the one we calculated previously [eq. (I.11)].

The quantized energy depends only upon the sum of the two quantum numbers l and k. This property, which is characteristic of the Coulomb potential, derives from the fact that the azimuthal and radial frequencies are equal: $\nu_\varphi = \nu_r$. To the energy E_n there correspond n quantized orbits, defined respectively by the following values of l: $l = 1, 2, \ldots, n$ (one eliminates the value $l = 0$ for reasons we shall not discuss here); they are ellipses of excentricity $\sqrt{1 - (l^2/n^2)}$. The value $l = n$ corresponds to the circular orbit [1]).

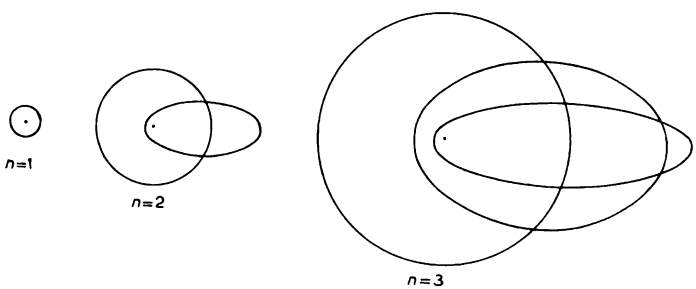


Fig. I.6. The Bohr orbits of the ground state ($n = 1$) and of the first two excited states ($n = 2, 3$) of the hydrogen atom. The relative dimensions of the orbits have been drawn to scale.

One can apply the same quantization rules to the relativistic equations of motion and thus determine the relativistic corrections to this theory of the hydrogen atom. One thus obtains a value of the Rydberg constant in better accord with experiment (cf. footnote p. 31). Furthermore, the "degeneracy" of the energy is lifted: to each value of n there correspond n very closely spaced but distinct values of the quantized energy, each corresponding to a precise value of the angular momentum $l\hbar$. Experimentally, one actually observes a fine structure of the hydrogen atom in fair agreement with the predictions of this theory.

Let us now examine the problem of space quantization. The above treatment, where one quantizes the orbits in the same plane, contains

---

[1]) The quantization of circular orbits led Bohr to find the Balmer formula as early as 1913. The quantization of elliptical orbits is due to Sommerfeld, who extended the theory to the relativistic case (see further on).

no privileged direction and only permits the determination of quantized spectra for scalar quantities such as the energy or the *magnitude* of the angular momentum ($L = l\hbar$). There exists a very general rule for space quantization for systems possessing axial symmetry (atom in constant magnetic field, for instance). In that case, the component L_z of the total angular momentum along the axis of symmetry Oz is a constant of the motion: it can be shown in Classical Mechanics that the latter can always be taken as the momentum conjugate to an angle φ fixing the overall orientation of the system about the axis Oz. Consequently

$$\oint L_z \, d\varphi = mh \qquad (m \text{ integer})$$

and since L_z is constant,

$$L_z = m\hbar. \tag{I.19}$$

The component of the angular momentum along the axis of symmetry is a (positive or negative) integral multiple of $\hbar$. The quantum number m is known as the *magnetic quantum number*.

16. Successes and Limitations of the Old Quantum Theory

We shall not pursue the discussion of the Old Quantum Theory any further. It has led to considerable progress in the knowledge of spectra, by furnishing a correct evaluation of the spectral terms of a large number of atomic and molecular systems. The treatment of the hydrogen atom is easily extended to hydrogen-like atoms (He^+, Li^{++}) and to the alkali atoms. The theory also applies to the vibration and rotation spectra of molecules, to the X-ray spectra of atoms, and to the normal Zeeman effect. Complemented by the semi-classical theory of the interaction between matter and radiation, it yields the various selection rules and probabilities of the different possible quantum transitions. In all cases, the predictions of the theory agree remarkably well with the experimental results, aside from occasional disagreements for very small quantum numbers, disagreements which one can remedy by imposing some empirical modifications on the quantization (the suppression of the value $l = 0$ of the azimuthal quantum number is an example).

Nevertheless it is an incomplete theory. The Bohr-Sommerfeld rules apply only to periodic or multiply-periodic systems. There exists no method of quantization for aperiodic motions. Thus, the

mechanism of an experiment as fundamental as that of Franck and Hertz remains completely unexplained: the Bohr-Sommerfeld theory gives the quantized energy levels of the atomic target, but there is no theoretical treatment whatsoever which would indicate the trajectory actually followed by the electron, and which would describe in detail the inelastic collision between electron and target atom. As a general rule, all collision phenomena lie outside the scope of the theory. Even in the calculation of spectral terms, the success of the theory is limited to the simplest systems; numerous difficulties arise as soon as one tries to set up the problem of the quantization of complex atoms in a rigorous manner. Moreover, there are cases of violent disagreement, such as the one of the (un-ionized) helium atom or the anomalous Zeeman effect.

Furthermore, the theory is by no means free of ambiguity and contradictions. We have an example in the prescriptions concerning space quantization. The quantization rule of the component L_z of the angular momentum of a system possessing axial symmetry about the z axis should apply just as well to systems with spherical symmetry, since the latter possess axial symmetry about any axis passing through the origin. In this last case we are led to the absurd prescription that the component of angular momentum along any axis passing through the origin must be an integral multiple of $\hbar$.

But the difficulties of principle of the Old Quantum Theory are considerably more serious than all these faults and contradictions. The quantization rules are purely formal restrictions imposed upon the solutions of the classical equations of motion; they were determined in an entirely empirical manner. The profound justification of this quantization of classical trajectories is completely absent. In fact, the very notion of trajectory is hard to reconcile with the quantization phenomenon. It implies that the particle possesses at each instant a well-defined position and momentum, and that these quantities vary in a continuous manner in the course of time. Under these conditions what would be the trajectory of a particle such as the electron in the Franck and Hertz experiment? If this electron has a precise trajectory, its energy varies in a continuous manner, and one must consequently give up the idea of an exchange of energy by quanta with the target atom; in other words, one must renounce the quantization of the energy levels of the target atom. Conversely, to postulate this quantization amounts to giving up the (classical) idea of a precise

trajectory of the electron and, quite logically, the idea of trajectory in general. We shall see in Chapters II and IV that this renunciation is fully justified, and we shall analyze its physical significance and its consequences in detail. In any event, we are led to renounce the classical equations of motion; under these conditions, one may well question the physical meaning of the solutions of these equations of motion such as those leading to the quantized trajectories of atoms.

The Old Quantum Theory undoubtedly represented a great step forward. Predicting a considerable body of experimental results from a few simple rules, it constitutes a general scheme for the phenomenological explanation of spectra and has, by virtue of this fact, played a large clarifying role in the history of contemporary physics. But this rather haphazard mixture of classical mechanics and *ad hoc* prescriptions can in no way be considered as a definitive theory.

17. Conclusions

In this chapter we have analyzed the main difficulties encountered by Classical Theory in the domain of microscopic physics. These difficulties appear as soon as one makes a detailed study of the interaction between matter and radiation. The essential fact is the appearance of discontinuity on this scale, connected with the existence of an indivisible quantum of action $\hbar$.

This *atomism of action* seems to be a fundamental characteristic of natural phenomena. On the macroscopic scale, the quantum $\hbar$ may be regarded as an infinitesimal quantity and one may be content with a classical description of phenomena, in which the evolution of physical systems is represented by means of dynamical variables, defined at each instant in a precise manner, and varying in continuous fashion with time. In atomic and subatomic physics, on the other hand, $\hbar$ ceases to be negligible and characteristic quantum effects appear. The upheaval affects the entire classical construct.

The classical wave treatment of electromagnetic radiation is irreconcilable with the experimental fact that the transfers of energy and momentum between matter and radiation take place through discrete and indivisible quanta. The photoelectric effect, and the Compton effect, in particular, cannot be understood unless one imagines light as a stream of corpuscles; in contrast, the photon hypothesis cannot be reconciled with the existence of interference or diffraction phe-

nomena, in which the behavior of light is characteristic of that of a superposition of waves. If one keeps to the language of classical physics, it is not possible to account in a completely coherent manner for the totality of light phenomena. One must appeal, depending upon the case, to one or the other of two contradictory representations – stream of corpuscles or superposition of waves. The connection between these two representations is given by the fundamental relations (I.4), into which the quantum $\hbar$ enters. The simplest way to interpret this wave-corpuscle duality is to postulate a statistical bond between waves and corpuscles, where the intensity of the wave at a point gives the probability of finding the photon which is associated with it, at that point.

As far as material systems are concerned, the quantization of certain quantities sets a limit to the conception according to which matter is made up of corpuscles obeying Newtonian Mechanics. The quantization of atomic energy levels, the quantization of the orientation of atoms or molecules placed in suitable conditions, are experimental facts incompatible with classical corpuscular theory.

In order to have a guide in the search for a coherent theory, it is well to list the elements of Classical Theory which can be maintained. The first realization which imposes itself concerns the fundamental laws of the conservation of energy and momentum: none of the experimental facts discussed in this chapter invalidates them. It therefore seems that these laws remain valid on the level of microscopic phenomena. In the second place, the failure of Classical Theory seems to have as sole origin the atomism of action. This theory retains its entire validity on the macroscopic scale, and more generally, whenever the quantum discontinuities may be considered negligible. This second statement lies at the base of the correspondence principle, stated in the last part of this chapter. The successes of the Old Quantum Theory, surprising for a theory whose foundations are so fragile, illustrate the fruitful character of this fundamental heuristic principle.

EXERCISES AND PROBLEMS

1. Consider the scattering of monochromatic photons by free electrons (Compton effect). Determine the wavelength shift, the magnitude and the direction of the recoil momentum of the electron as a function of the deflection angle of the photon, assuming: a) that the electron is initially at rest; b) that the electron has an initial momentum P in the direction of the incident photon.

What is the maximum value of the momentum transferred to the electron in the two cases?

2. Discuss, in connection with the experiment of Young's interference slits, the double aspect — wave and corpuscular — of light. Show that it is not possible to attribute to each photon a precise trajectory passing through one of the two slits.

3. A gyroscope has a magnetic moment $\boldsymbol{\mu}$ proportional to its angular momentum: $\boldsymbol{\mu} = M\boldsymbol{l}$. From the expression $-\boldsymbol{\mu}\cdot\boldsymbol{H}$ of the magnetic energy, derive the equation of motion of $\boldsymbol{l}$ in a constant magnetic field $\boldsymbol{H}$, and show that the gyroscope carries out a precessional motion of circular frequency $\omega_L = MH$ (Larmor frequency).

4. In the non-relativistic limit, the classical equations of motion of an electron in an electromagnetic field are

$$\frac{\mathrm{d}}{\mathrm{d}t}(m\boldsymbol{v}) = e\left(\boldsymbol{E} + \frac{v}{c}\times\boldsymbol{H}\right). \qquad (a)$$

Let $\boldsymbol{A}$ and φ be the vector and scalar potentials of this field, respectively

$$\left(\boldsymbol{E} = -\operatorname{grad}\varphi - \frac{1}{c}\frac{\partial \boldsymbol{A}}{\partial t}; \qquad \boldsymbol{H} = \operatorname{rot}\boldsymbol{A}\right).$$

Show that these equations can be derived from the Lagrangian

$$L = \tfrac{1}{2}mv^2 + e\left(\frac{v\cdot\boldsymbol{A}}{c} - \varphi\right).$$

Calculate the conjugate Lagrange momenta and form the classical Hamiltonian.

5. The equations (a) of Problem (I.4) remain true in the relativistic domain provided that one replaces the rest mass m by the "relativistic mass" $M = m[1 - (v^2/c^2)]^{-\frac{1}{2}}$. Verify that the formalisms of Lagrange and Hamilton remain valid, the equations being deduced from the Lagrangian

$$L = -mc^2\sqrt{1 - \frac{v^2}{c^2}} + e\left(\frac{\boldsymbol{v}\cdot\boldsymbol{A}}{c} - \varphi\right).$$

Show that the Hamiltonian in that case is written

$$H = \left[m^2c^4 + \left(\boldsymbol{p} - \frac{e}{c}\boldsymbol{A}\right)^2 c^2\right]^{\frac{1}{2}} + e\varphi$$

where $\boldsymbol{p}$ stands for the momentum, i.e. the vector whose three components p_x, p_y, p_z are the respective conjugate momenta of x, y, z.

N.B.: $\boldsymbol{p} = M\boldsymbol{v} + (e/c)\boldsymbol{A}$ and $H = Mc^2 + e\varphi\cdot H$ and $\boldsymbol{p}c$ form a four-vector of space-time, just like φ and $\boldsymbol{A}$, and one has: $(H - e\varphi)^2 - (\boldsymbol{p}c - e\boldsymbol{A})^2 = m^2c^4$.

6. A material point of mass m is constrained to move on the x axis where it is subject to a restoring force $(-Kx)$ proportional to its distance from the origin (***harmonic oscillator***). Apply the Bohr-Sommerfeld quantization rule to this system; calculate the energy, the period, and the amplitude of the quantized trajectories.

7. Quantize the circular electronic orbits of the hydrogen atom by applying the Bohr-Sommerfeld rule. Determine the energy, the period, and the radius of the quantized orbits. Calculate specifically the numerical values of the energy, the period, and the radius of the lowest orbit [$mc^2 = 0.51$ MeV; $(\hbar c/e^2) \approx 137$] and determine in that particular case the relativity correction. Verify that in the limit of large quantum numbers, the Bohr frequencies tend toward those predicted by classical electrodynamics (*correspondence principle*).

MATTER WAVES AND THE SCHRÖDINGER EQUATION

1. Historical Survey and General Plan of the Succeeding Chapters

The founding of Quantum Mechanics can be placed between the years of 1923 and 1927. Two equivalent formulations thereof have been proposed almost simultaneously: Matrix Mechanics and Wave Mechanics.

The starting point of *Matrix Mechanics* [1]) is a critical analysis of the Old Quantum Theory. The point of view developed by Heisenberg is the following. In any physical theory one must distinguish the concepts and quantities which are physically observable from those which are not. The former must of necessity play a role in the theory, the latter can be modified or abandoned without impairment. In establishing a satisfactory theory of microscopic phenomena, one must as far as possible start only from the former. Now the Old Quantum Theory calls upon a whole set of notions without experimental foundation; it is put in jeopardy to the extent that these notions are erroneous.

The notion of electronic orbit is an example of a concept without experimental foundation. Indeed, let us ask if it is conceivable [2]) to follow experimentally the motion of the electron in a Bohr orbit of the hydrogen atom. In order to observe this motion one has to perform successive measurements of the position of the electron, where the tolerated margin of error is very much smaller than the mean radius a of the orbit. Such measurements are conceivable with X-rays of sufficiently short wavelength: $\lambda \ll a$. However, according to the laws of the Compton effect, the collision of each X-ray photon with an electron is accompanied by a momentum transfer of the order of $\hbar/\lambda$ ($\gg \hbar/a$) and consequently by a finite perturbation of the electron motion which one wishes to observe. One can readily show (cf. Problem

[1]) W. Heisenberg, Z. Physik 33 (1925) 879; M. Born and P. Jordan, Z. Physik 34 (1925) 858; M. Born, W. Heisenberg, and P. Jordan, Z. Physik 35 (1926) 557; P. A. M. Dirac, Proc. Roy. Soc. London A 109 (1925) 642.

[2]) The limits of possibility of observation at which we shall arrive are imposed by the very nature of things; the difficulties of practical realization of the measurements are of no concern here.

II.1) that this perturbation is the more appreciable the smaller the quantum number n; in particular, when the electron is on the lowest orbit ($n = 1$), the mean energy transferred per collision is at least equal to the ionization energy. This uncontrollable perturbation of the observed system by the measuring apparatus limits the precision with which one might hope to know the electronic orbits. In the limit of small quantum numbers, this perturbation is so strong that any observation of the orbit, no matter how coarse, is doomed to failure. Since no experiment allows us to assert that the electron actually describes a precise orbit in the hydrogen atom, nothing prevents us from abandoning the very notion of an orbit; in other words, the fact that the atom is in a well-defined energy state does not necessarily imply that the electron has at each instant a well-defined position and momentum [1]).

The Matrix Mechanics of Heisenberg, Born and Jordan abandons the notion of an electron orbit. Starting exclusively from physically observable quantities such as the frequencies and the intensities of the radiation emitted by atoms, the theory associates with each physical quantity a certain matrix; in contrast to the quantities of ordinary algebra, these matrices obey a non-commutative algebra [2]); it is on this essential point that the new mechanics differs from the Classical Mechanics. The equations of motion of the dynamical variables of a quantized system are thus equations between matrices. Following the correspondence principle, one assumes that these equations are formally identical to the equations (between quantities of ordinary algebra) of the corresponding classical system.

The *Wave Mechanics* of Schrödinger [3]) starts from an entirely different point of view. It originates in the works of L. de Broglie

---

[1]) This is to be compared with the discussion at the end of the first chapter on the Franck and Hertz experiment, showing that the existence of a continuous trajectory for the bombarding electron is incompatible with the quantization of the energy levels of the target atom. Postulating such a quantization forces us to renounce the notion of a trajectory.

[2]) The definition of matrices and the presentation of their principal properties will be given in Chapters VII and VIII, where we shall see the close relationship existing between these matrices and the linear operators with which we shall be dealing later on (§ 11).

[3]) E. Schrödinger, Ann. Physik (4) 79 (1925) 361 and 489; 80 (1926) 437; 81 (1926) 109.

on matter waves [1]). Seeking to establish the bases of a unified theory of matter and radiation, de Broglie had stated the hypothesis that the wave-corpuscle duality is a general property of microscopic objects, and that matter, as well as light, exhibits both wave and corpuscular aspects. Having established the correspondence between the dynamical variables of the corpuscle and the characteristic quantities of the associated wave, he was able to deduce the quantization rules of Bohr-Sommerfeld by a semi-quantitative argument. The speculations of L. de Broglie on the wave nature of matter were to be confirmed very directly a few years later by the discovery of diffraction phenomena analogous to those of wave optics. Meanwhile Schrödinger, pursuing and generalizing this notion of matter waves, discovered the equation of propagation of the wave function representing a given quantum system; a very simple correspondence rule enables us to deduce that fundamental equation from the Hamiltonian of the corresponding classical system. The Schrödinger equation constitutes the essential element of Wave Mechanics.

As Schrödinger has shown [2]), Wave Mechanics and Matrix Mechanics are equivalent. They are two particular formulations of a theory which can be presented in very general terms. The setting up of this general formalism of the Quantum Theory is essentially due to Dirac [3]). The quantum theory thus obtained is a non-relativistic theory of material particles. It was complemented by a quantum theory of the electromagnetic field [4]) to form a coherent whole suitable for the treatment of all problems dealing with systems of non-relativistic, material particles interacting with the electromagnetic field. We might add that the interpretation and internal consistency of the theory were not fully understood until after the works of Born, Heisenberg and Bohr appeared [5]). The major part of this book deals with this theoretical framework and its applications. The problems

---

[1]) L. de Broglie, Nature 112 (1923) 540; *Thesis*, Paris, 1924, Ann. Physique (10) 2 (1925).

[2]) E. Schrödinger, Ann. Physik (4) 79 (1926) 734.

[3]) P. A. M. Dirac, *The Principles of Quantum Mechanics*, (Oxford, Clarendon Press), 1st ed. (1930), 4th ed. (1958).

[4]) P. A. M. Dirac, Proc. Roy. Soc. London A 114 (1927) 243 and 710; P. Jordan and W. Pauli, Z. Physik 45 (1928) 151.

[5]) M. Born, Z. Physik 38 (1926) 803; W. Heisenberg, Z. Physik 43 (1927) 172; N. Bohr, Naturwiss. 16 (1928) 245; 17 (1929) 483 and 18 (1930) 73. One

of Relativistic Quantum Mechanics will be attacked only in the last part; we shall essentially restrict ourselves to an outline of Dirac's relativistic theory of the electron [1]), and to an elementary introduction to the theory of quantized fields.

Of the various ways of introducing the Quantum Theory, the one which uses the general formalism is undoubtedly the most elegant and the most satisfactory. However, it requires the handling of a mathematical symbolism whose abstract character runs the risk of masking the underlying physical reality. Wave Mechanics, which utilizes the more familiar language of waves and partial differential equations, lends itself better to a first encounter. Furthermore, it is in that form that the Quantum Theory is most frequently used in elementary applications. That is why we shall begin with a general outline of Wave Mechanics. We start this chapter with a discussion of the matter-wave concept; we then set up the Schrödinger equation and discuss its principal properties; we show in particular how this equation allows us to determine the energy levels of stationary states. In order to acquire a certain familiarity with the Schrödinger equation, we devote Chapter III to the handling of simple problems relating to one-dimensional quantum systems, and to the derivation of several important properties of such systems. We shall then be in a position

---

can also find a thorough discussion of the physical interpretation of the theory in W. Heisenberg, *The Physical Principles of the Quantum Theory* (Chicago, University of Chicago Press, 1930); also (New York, Dover); Niels Bohr, *Atomic Theory and the Description of Nature* (Cambridge University Press, Cambridge, 1934) (translated from the Danish).

The interpretation under discussion here is the statistical interpretation of the Copenhagen school. It is the one we are developing in this book. After violent controversies, it has finally received the support of the great majority of physicists. However, it had (and still has) a number of die-hard opponents, among which one should notably list Einstein, Schrödinger, and de Broglie. The controversy has finally reached a point where it can no longer be decided by any further experimental observations; it henceforth belongs to the philosophy of science rather than to the domain of physical science proper. The main arguments can be found in *Albert Einstein, Philosopher-Scientist*, P. A. Schilpp, editor, (New York, Tudor Publishing Company, 1949 and 1951) (see especially the articles of Bohr and Einstein), and in the book of L. de Broglie, *La Théorie de la Mesure en Mécanique Ondulatoire* (Paris, Gauthier–Villars, 1957).

[1]) P. A. M. Dirac, Proc. Roy. Soc. London A 117 (1928) 610 and A 118 (1928) 351.

to attack the general problems of interpretation of the Quantum Theory; the latter form the subject of Chapter IV. Chapter V is devoted to the development of the formalism of Wave Mechanics and its statistical interpretation in accordance with the principles defined in Chapter IV. Chapter VI deals with the classical approximation of Wave Mechanics. Only after this general survey of the theory expressed in the language of Wave Mechanics shall we present (Chs. VII–VIII) the general formalism of the Quantum Theory.

I. MATTER WAVES

2. Introduction

The double aspect of light – wave-like and corpuscular – is one of the most striking features connected with the appearance of quanta in physics. Let us suppose that matter too possesses this dual character; just as an electromagnetic wave is associated with each photon, so we associate with each material particle a wave whose angular frequency ω is connected with the energy of the particle E by the relation $E=\hbar\omega$. If one adopts this point of view, the atom behaves as a resonant cavity having a discrete series of proper frequencies; in this way the quantization of its energy levels is explained.

At the same time the possibility exists of establishing a unified theory in which matter and radiation are different varieties of the same type of object, having both wave-like and corpuscular character. These suppositions, which have guided de Broglie in his theory of matter waves, were found to be entirely justified, as we shall see.

The main properties of matter waves are obtained by analogy with optics. Just as for photons, we assume that the value at each point of the intensity of the wave associated with a particle gives the probablity of finding the particle at that point. The particle is better localized, the more restricted the domain occupied by the wave. The conditions for the validity of Classical Mechanics are fulfilled when the wave maintains in the course of time a sufficiently small extension so that it may be approximated by a point and one may attribute a precise motion to the particle. An analogous situation is encountered in optics whenever the wavelength λ may be considered negligibly small; this is the approximation used in *geometrical optics*, where no typically wave-like phenomenon can be detected. This approximation is valid when the optical properties of the medium through which the

light travels, remain essentially constant over a distance of several wavelengths ($|\operatorname{grad}\lambda| \ll 1$). This suggests that classical corpuscular theory is correct in the absence of a field or in a slowly varying field, and to the extent that one does not seek to localize the particle with too great precision. These predictions are in accord with the most common observations concerning the motion of the atomic and subatomic particles in the presence of quasi-static and near-uniform fields: trajectories of charged particles in a static electromagnetic field, deflections of paramagnetic atoms in the Stern-Gerlach magnet, etc. In these limiting cases the theory of matter waves must be equivalent to the classical theory (correspondence principle).

3. Free Wave Packet. Phase Velocity and Group Velocity

Consider the propagation of matter waves in a homogeneous, isotropic medium. The simplest type of wave is a plane, monochromatic wave

$$e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} \tag{II.1}$$

which represents a vibration of wavelength $\lambda = 2\pi/k$ travelling in the direction of its wave vector $\mathbf{k}$ with constant velocity. The velocity considered here is the velocity of propagation of planes of equal phase, or *phase velocity:*

$$v_\varphi = \frac{\omega}{k}.$$

The frequency ω is independent of the direction of $\mathbf{k}$ but may eventually depend upon the length of that vector. As any wave may be considered as a superposition of plane monochromatic waves, knowledge of the "dispersion law" $\omega(k)$ is sufficient to determine the behavior of any wave in the course of time.

By hypothesis, each frequency ω corresponds to a well-defined energy E of the particle

$$E = \hbar\omega. \tag{II.2}$$

It is therefore natural to associate the wave (II.1) with uniform rectilinear motion of energy E directed parallel to $\mathbf{k}$.

By examining the classical approximation we shall be able to relate $\mathbf{k}$ to the momentum $\mathbf{p}$ of the particle. In order to realize this approximation, one must associate with the particle a wave of limited

extension. The wave (II.1) evidently does not satisfy this condition; one may realize it, however, by the superposition of plane waves with neighboring wave vectors. One thus forms a *wave packet*:

$$\psi(\mathbf{r}, t) = \int f(\mathbf{k}')\, e^{i(\mathbf{k}'\cdot\mathbf{r} - \omega' t)}\, d\mathbf{k}'.$$

We designate by A and α the absolute value and the phase of f, respectively. By hypothesis, A has appreciable values only in a small region surrounding $\mathbf{k}$. Our purpose is to examine to what extent and under what conditions the "motion" of this wave packet may be likened to the motion of a classical particle.

To simplify matters, we shall first take up the problem of the one-dimensional wave packet:

$$\psi(x, t) = \int_{-\infty}^{+\infty} f(k')\, e^{i(k'x - \omega' t)}\, dk'.$$

Let us set

$$\varphi = k'x - \omega' t + \alpha,$$

$\psi(x, t)$ is the integral of the product of a function A exhibiting a pronounced peak in a region S of extension Δk surrounding the point $k' = k$, by an oscillating function exp (iφ). If the number of oscillations of exp (iφ) in this region is large, the contributions to the integral from the different portions of the region interfere destructively and ψ remains negligible. The largest (absolute) values of ψ obtain when the phase φ remains practically constant in S, i.e. $d\varphi/dk \approx 0$ (the symbol d/dk denotes the value taken by the derivative of the function with respect to k' when $k' = k$). Roughly speaking the only appreciable values of ψ are those for which exp (iφ) carries out only one or a fraction of one oscillation, namely

$$\Delta k \times \left| \frac{d\varphi}{dk} \right| \lesssim 1.$$

Since

$$\frac{d\varphi}{dk} = x - t\frac{d\omega}{dk} + \frac{d\alpha}{dk},$$

the wave $\psi(x, t)$ is practically concentrated in a region of extension

$$\Delta x \simeq \frac{1}{\Delta k}$$

surrounding the "center of the wave packet", defined by the condition $d\varphi/dk = 0$, namely

$$x = t\frac{d\omega}{dk} - \frac{d\alpha}{dk}.$$

This point travels with uniform motion, whose velocity

$$v_g = \frac{d\omega}{dk} \tag{II.3}$$

is called the *group velocity* of the wave $\exp[i(kx - \omega t)]$. It is this velocity v_g, and not the phase velocity v_φ which, in the classical approximation where one considers the extension of the wave packet to be negligible, must be identified with the particle velocity

$$v = \frac{dE}{dp} \qquad (\approx p/m \text{ in non-relativistic approximation}).$$

From the condition $v = v_g$ and from relation (I.2) one obtains [1]) the de Broglie relation:

$$p = \hbar k = \frac{h}{\lambda}. \tag{II.4}$$

The above treatment is easily generalized to the three-dimensional wave packet: the center of the packet travels with uniform motion at the velocity

$$v_g = \text{grad}_k\, \omega. \tag{II.3'}$$

This group velocity must be identified with the particle velocity:

$$v = \text{grad}_p\, E.$$

This condition, combined with relation (II.2), leads to the following relations [2]) between dynamical variables of the particle and characteristic quantities of the associated wave:

$$E = \hbar\omega, \qquad \mathbf{p} = \hbar\mathbf{k}. \tag{II.5}$$

---

[1]) Rigorously, these two conditions define k as a function of p only to within an additive constant. One fixes the constant by requiring the relation between p and k to be independent of the direction of travel chosen along the coordinate axis.

[2]) One obtains the second relation (II.5) to within a constant vector, which one chooses equal to zero by imposing upon that relation the condition to be invariant under a rotation of axes (cf. preceding footnote).

These relations are identical to the relations (I.4) found in the case of photons.

To conclude, we shall examine the preceding results from the point of view of the principle of relativity.

In the non-relativistic approximation, the energy E is defined only to within a constant; a modification of the zero of the energy scale has the effect of adding to the frequency $\omega(k)$ a constant frequency ω_0 [eq. (II.2)] and of multiplying the function $\psi(\boldsymbol{r}, t)$ by the phase factor $\exp(-i\omega_0 t)$. This in no way affects the preceding results concerning the motion of the wave packet, nor the relations (II.5) derived therefrom.

However, the preceding treatment does not depend at all on the non-relativistic approximation. The relativity principle allows us to define unambiguously the energy $E = \sqrt{m^2c^4 + p^2c^2}$, and the frequency ω which corresponds to it. The energy E and the momentum $\boldsymbol{p}$ are respectively the time and space components of the same four-vector (with the convention $c = 1$). The same holds true for the frequency ω and the wave vector $\boldsymbol{k}$. The relations (II.5) satisfy the principle of relativity: they state that the four-vectors $(E, \boldsymbol{p})$ and $(\omega, \boldsymbol{k})$ are proportional.

4. Wave Packet in a Slowly Varying Field

The above results, and especially relations (II.5) may be extended to the motion of particles in a slowly varying field; the conditions for the classical approximation are realized as long as the field variations are negligible on the wavelength scale.

The laws of propagation are those of geometrical optics. In particular, a wave packet of limited extension, analogous to those which were studied in the preceding section, follows the trajectory of a ray with a velocity equal to the group velocity. In order to compare the motion of the wave packet to the motion of a classical particle, it is necessary:

(*a*) that the rays corresponding to the (angular) frequency ω be identical to the classical trajectories of energy $E = \hbar\omega$;

(*b*) that the group velocity along each ray be equal to the velocity of the corresponding classical particle.

The trajectories of the classical particle are given by the principle

of least action (I.12); for a fixed energy E, the Lagrangian is, according to equation (I.13), equal to $\boldsymbol{p}\cdot(\mathrm{d}\boldsymbol{r}/\mathrm{d}t)-E$ and the principle may be written:

$$\delta I_{12} \equiv \delta \int_{M_1}^{M_2} \boldsymbol{p}\cdot\mathrm{d}\boldsymbol{r} = 0.$$

It expresses the fact that the integral I_{12} calculated along a curve joining the points M_1 and M_2 is stationary when this curve is the actual trajectory of the particle from M_1 to M_2. The momentum $\boldsymbol{p}$ in the most general case is a function of the position $\boldsymbol{r}$ and the velocity $\boldsymbol{v}=(\mathrm{d}\boldsymbol{r}/\mathrm{d}t)$, in other words, a function of the position on the curve and of the direction of its tangent. In the case of a non-relativistic particle in a scalar potential $V(\boldsymbol{r})$

$$E=\frac{p^2}{2m}+V(\boldsymbol{r}) \quad \text{and} \quad \boldsymbol{p}=m\boldsymbol{v}, \tag{II.6}$$

momenta and velocities are parallel ($\boldsymbol{p} \parallel \mathrm{d}\boldsymbol{r}$). However, the principle is equally valid in more general cases, such as that of a particle in a magnetic field where this condition does not obtain [1]).

At a given fixed frequency ω, the rays of geometrical optics are determined by another variational principle, Fermat's principle, which may be written

$$\delta J_{12} \equiv \delta \int_{M_1}^{M_2} \boldsymbol{k}\cdot\mathrm{d}\boldsymbol{r} = 0;$$

$\boldsymbol{k}$ is the wave vector. The integral J_{12} calculated along a given curve joining the points M_1 and M_2 represents the optical path along that curve. Fermat's principle expresses the fact that the ray joining M_1 and M_2 is the curve along which the optical path is stationary. As a general rule $\boldsymbol{k}$ (perpendicular to the surfaces of equal phase) is a function of the position on the curve and of the slope of that curve. In an isotropic medium, where the phase velocity is independent of the direction, the vector $\boldsymbol{k}$ is directed along the tangent to the curve and its length $k=2\pi/\lambda$ depends only upon the position along the curve and not upon the direction of propagation. However, Fermat's principle applies equally well to anisotropic media.

[1]) We recall that the momentum (i.e. linear momentum) of a particle is the vector whose three components are the Lagrange-conjugate momenta of the three position coordinates. It is sometimes, but not always, equal to the product of the mass and the velocity of the particle.

Note the perfect formal analogy of the two variational principles. In order that the rays relative to the frequency ω merge with the classical trajectories of energy E [condition (a)], it is sufficient that $\boldsymbol{k}$ and $\boldsymbol{p}$ be proportional:

$$\boldsymbol{p} = \alpha \boldsymbol{k}.$$

The constant of proportionality α is determined by condition (b). The group velocity $\boldsymbol{v}_g$ is the gradient with respect to $\boldsymbol{k}$ of the frequency ω; therefore

$$\boldsymbol{v}_g = \frac{1}{\hbar} \operatorname{grad}_{\boldsymbol{k}} E = \frac{\alpha}{\hbar} \operatorname{grad}_{\boldsymbol{p}} E .$$

As for the velocity of the particle, it is given by the formula

$$\boldsymbol{v} = \operatorname{grad}_{\boldsymbol{p}} E.$$

These two velocities are equal if $\alpha = \hbar$.

We thus come back to the relations (II.5) in a completely general way.

In the case of a non-relativistic particle in a scalar potential $V(\boldsymbol{r})$ which is slowly changing, the wave travels in an isotropic medium and the wavelength λ is given [cf. eq. (II.6)] by the relation

$$\lambda = \frac{h}{p} = \frac{h}{\sqrt{2m(E - V(\boldsymbol{r}))}} . \tag{II.7}$$

5. Quantization of Atomic Energy Levels

The theory of matter waves leads very simply to the quantization of the energy levels of atoms.

To be specific, let us return to the hydrogen atom. Consider an elliptical orbit of energy E. The value of the wave vector $\boldsymbol{k}$ is given at each point of the orbit by relations (II.5). The phase of the wave at a given point of the orbit increases by $\oint \boldsymbol{k} \cdot \mathrm{d}\boldsymbol{r}$ in each revolution. In order that a standing wave be set up, it is necessary that this phase be an integral multiple of 2π. This yields the quantization condition

$$\oint \boldsymbol{p} \cdot \mathrm{d}\boldsymbol{r} = \hbar \oint \boldsymbol{k} \cdot \mathrm{d}\boldsymbol{r} = nh \qquad (n \text{ integer} > 0)$$

which may also be written, adopting the notation of the first chapter

$$\oint p_r \, \mathrm{d}r + \int p_\varphi \, \mathrm{d}\varphi = nh.$$

Similar arguments lead to the Bohr-Sommerfeld quantization rules in all cases of periodic and multiply-periodic motions.

Of course, the preceding results are valid only in the geometrical optics approximation, where the notions of wavelength and of wave vector retain their meaning. In particular, we cannot assert that the quantization conditions remain the same for small quantum numbers. The only certain fact is the energy quantization which is tied to the establishment of a standing wave.

To deal with the more general cases, one must extend this approximate theory as one does classical optics when one goes from geometrical optics to wave optics [1]: having postulated the existence of matter waves, one has yet to discover their equation of propagation. Before attacking this problem, we shall examine how the wave nature of matter may be – and actually is – revealed experimentally, although the concept of corpuscles of matter can by no means be abandoned.

6. Diffraction of Matter Waves

The practical possibilities of observation evidently depend on the wavelength of the object under study. With macroscopic objects, the commonly realized wavelengths are so tiny that typical wave effects cannot be detected in practice. With objects of atomic dimensions, on the other hand, it is possible to form beams of wavelength comparable to that of X-rays and capable of giving rise to similar diffraction effects by crystals [2].

The first diffraction experiments with matter waves were performed

---

[1] Hence the name Wave Mechanics given to the theory.

[2] To be specific, let us consider a particle used to demonstrate Brownian motion. The smallest particles of this kind have a diameter of the order of 1 micron and a mass

$$M \approx 10^{-12} \text{ g}.$$

In thermal equilibrium at ordinary temperatures, their mean kinetic energy $\frac{3}{2}kT$ is about 0.4×10^{-13} erg, from which we obtain a mean wavelength

$$\lambda = \frac{h}{p} = \frac{h}{\sqrt{3MkT}} \simeq 5 \times 10^{-6} \text{ Å}.$$

At the same energy a helium atom has a wavelength $\lambda \approx 0.9$ Å, a neutron $\lambda \approx 1.9$ Å, and an electron $\lambda \approx 77$ Å.

with electrons by Davisson and Germer (1927), G. P. Thomson (1928) and Rupp (1928).

Davisson and Germer worked with reflection by a single crystal and observed spots of the von Laue type; G. P. Thomson and Rupp studied the Debye-Scherrer rings obtained by diffraction through a thin, polycrystalline target.

In these experiments the incident beam is obtained by the acceleration of electrons in an electrical potential. If E is the energy of the electrons in electron-volts, the de Broglie wavelength of the electrons measured in angstrom units is

$$\lambda = \frac{12.2}{\sqrt{E_{(\mathrm{eV})}}} \text{ Å}.$$

With energies in the range from 1 to 100 keV, one finds oneself in the domain of ordinary crystal spectrography. Knowing the parameters of the crystal lattice, it is possible to deduce an experimental value for the electron wavelength from the interference pattern; the theoretical value of de Broglie is in perfect accord with this experimental value.

Analogous crystal-diffraction experiments were performed with monoërgic beams of helium atoms and hydrogen molecules (Stern, 1932), furnishing a new verification of the de Broglie relation. The wavelength entering here is the wavelength associated with the motion of the center of gravity of each atom or molecule of the beam. The same observations can be made with beams of slow neutrons from nuclear reactors. All these experiments clearly show that the wave-like structure is not peculiar to electrons, but that one is actually dealing with a very general property of material objects.

7. Corpuscular Structure of Matter

Pursuing the analogy between matter waves and classical wave optics, one may ask if it is possible to renounce the notion of corpuscles of matter once and for all, and to replace the classical theory by a wave theory where the wave $\psi(\boldsymbol{r}, t)$ would play the role played by the electromagnetic field in the theory of radiation.

The picture of corpuscles of matter, localized grains of energy and momentum, would be replaced by that of a continuous wave with a continuous distribution of energy and momentum. The particles of classical mechanics would actually be wave packets of finite but

negligibly small size. As shown above such packets obey the laws of motion of classical particles under certain limiting conditions which are precisely those where classical mechanics turns out to be correct. Nevertheless, even in the absence of a field, a wave packet cannot maintain this appearance of a particle indefinitely because it spreads out little by little in the course of its displacement, and finally, after a sufficiently long time, it occupies as large a region of space as one wishes (cf. Problem II.6) [1]). It would be difficult to explain under these conditions why matter appears so often in the form of well-localized particles.

However, the inadequacy of such a pure wave theory is clearly exhibited by a careful analysis of any diffraction experiment with matter waves. Consider, for instance, a beam of monoërgic electrons traversing a polycrystalline foil; on a screen suitably placed on the other side of the foil, one observes a central spot due to the transmitted wave, surrounded by concentric rings due to the diffracted wave. Suppose, to be specific, that the incident wave is a wave packet $\psi(\boldsymbol{r}, t)$ restricted in space; it is obtained by placing (Fig. II.1) an

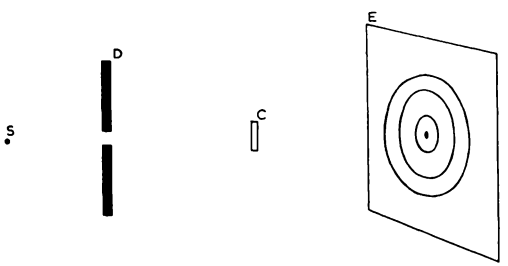


Fig. II.1. Diffraction of electrons by a polycrystalline foil. The electron beam, originating from the source S is collimated by the diaphragm D and then diffracted by the polycrystalline foil C. One observes the diffraction pattern on screen E.

intense source S of monoërgic cathode rays behind a diaphragm D equipped with a shutter whose opening is fixed once and for all. When this wave propagates through the foil C, it splits into a transmitted and a diffracted wave and forms the interference pattern described above, on the screen. Since it is assumed to be a continuous, classical

[1]) The case of the harmonic oscillator forms an exception to this very general rule concerning the indefinite spreading of a wave packet (cf. Ch. VI).

wave, the observed interference pattern must be continuous. If one diminishes the intensity of the incident wave, everything else being equal (for instance, by increasing the distance between the source S and the diaphragm D), the intensity of the interference spots decreases accordingly, but the interference pattern must remain continuous. Experiment invalidates these predictions. The observed pattern is actually made up of a succession of well-localized impacts. If one decreases the intensity of the wave, the number of these impacts decreases proportionately. In the limit of very low intensity, one eventually observes just a single impact, either on the central spot, or on one of the diffraction rings. The simplest explanation is to attribute each impact to the passage of a corpuscle of matter: an electron.

Note the perfect parallelism between the situation described here and the diffraction experiment of light by a grating, discussed in the first chapter. One may pursue the parallelism to the end and conclude that the simplest interpretation one might give of the wave-corpuscle duality is a statistical interpretation: namely that the intensity of the wave at each point of the screen gives the probability of occurrence of an impact at that point.

8. Universal Character of the Wave-Corpuscle Duality

We conclude from all this that microscopic objects have a very general property: they appear under two apparently irreconcilable aspects, the wave aspect on the one hand, exhibiting the superposition property characteristic of waves, and the corpuscular aspect on the other hand, namely localized grains of energy and momentum, There exists a universal relationship between these two aspects, given by equations (II.5). Furthermore, the bond between corpuscles and associated waves is a statistical one, on which we shall have more to say later.

II. THE SCHRÖDINGER EQUATION

9. Conservation Law of the Number of Particles of Matter

All we have said thus far reveals a remarkable similarity of properties of light and matter. Nevertheless, an important difference is to be noted. Even in the simplest situations, the number of photons may vary in the course of time through emission or absorption. By

contrast, the number of electrons, and more generally the number of elementary particles of matter remains constant. This is indicated by an analysis of the most common phenomena of atomic physics and the very success of the quantum mechanics of systems of particles will confirm the validity of this important conservation law.

In reality we are not dealing here with an absolute conservation law, and the disparity between matter and light is not as pronounced as we have just stated. Since the discovery of the positron (Anderson, 1932), a particle of the same mass m as that of the electron and of opposite charge, one knows that it is possible, under certain circumstances, to create electron-positron pairs (emission of matter) and conversely, that a positron and an electron undergoing collision can annihilate (absorption of matter) giving off energy in the form of radiation. In accordance with the law of equivalence between mass and energy, the energy necessary to create an electron-positron pair is at least equal to $2mc^2$ ($\approx$ 1 MeV). One has another case of emission of electrons (or positrons) in the beta decay of atomic nuclei. If one restricts oneself to phenomena of atomic physics, the positrons are absent, nuclei are stable and the energy transfers lie below the threshold for electron-positron pair creation; the conservation law stated above is then obeyed. We shall assume this in what follows.

This law greatly facilitates the development and interpretation of the quantum theory of matter. The various quantum systems which we consider are made up of a well-defined number of material particles. The simplest system is that of a particle, for instance an electron, in an external force field. The wave which is associated with it at each instant t is a function $\Psi(\boldsymbol{r}, t)$ of the position coordinates of that particle. The hydrogen atom is a system of two particles, an electron and a proton, interacting with each other. The associated wave $\Psi(\boldsymbol{r}_e, \boldsymbol{r}_p; t)$ depends upon the position coordinates $\boldsymbol{r}_e$ and $\boldsymbol{r}_p$ of these two particles. A complex atom is formed of a nucleus of charge Ze defined by its position $\boldsymbol{R}$, and of Z electrons whose positions are determined by the vectors $\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_Z$, respectively. The associated wave is a certain function $\Psi(\boldsymbol{R}, \boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_Z; t)$. One can similarly define the wave function of more complex systems.

10. Necessity for a Wave Equation and Conditions Imposed upon this Equation

We have seen that the intensity at a given point and at a given

instant of the wave associated with a particle gives the probability of finding the particle at that point and at that instant. More generally, we *postulate* that the wave function Ψ of a quantum system completely defines its dynamical state; otherwise stated, all the predictions which can be made concerning the dynamical properties of the system at a given instant of time t can be deduced from a knowledge of Ψ at that instant. Hence the central problem of the theory is the following: knowing the wave function at a given initial instant t_0, determine that function at all later instants. In order to do this, one must know the equation of propagation of the wave Ψ.

It is quite clear that no deductive reasoning can lead us to that equation. Like all equations of mathematical physics, it must be postulated and its only justification lies in the success of the comparison of its predictions with the experimental results. Nevertheless, the choice of a wave equation is restricted by a certain number of *a priori* conditions if one wishes to maintain the previously defined interpretation for Ψ:

(A) The equation must be *linear* and *homogeneous;* the wave thus possesses the property of superposition, characteristic of waves in general: namely if Ψ_1 and Ψ_2 are solutions of the equation, any linear combination $\lambda_1\Psi_1+\lambda_2\Psi_2$ of these functions is also a solution thereof.

(B) It must be a differential equation of the *first order with respect to time*; thus, specifying Ψ at a given initial instant uniquely defines its entire later evolution, in accord with the hypothesis that the dynamical state of the physical system is entirely determined once Ψ is given.

On the other hand, the predictions of the theory must coincide with those of Classical Mechanics in the domain where the latter is valid. In other words, the equation must lead to the same laws of motion of wave packets as the de Broglie theory, in the limit where the geometrical optics approximation is valid. This suggests that the equation bears a formal analogy to certain equations of Classical Mechanics (*correspondence principle*).

All these considerations will lead us to the Schrödinger equation in a very natural way. Before carrying out this program, we shall introduce a mathematical concept which turns out to be very useful later on, namely the operator concept.

11. The Operator Concept

Consider the function $\partial\Psi/\partial t$, the derivative of Ψ with respect to time; one can say that the operator $\partial/\partial t$ acting on the function Ψ yields the function $\partial\Psi/\partial t$. More generally, if a certain operation allows us to bring into correspondence with each function Ψ of a certain function space, one and only one well-defined function Ψ' of that same space, one says that Ψ' is obtained through the action of a given *operator* A on the function Ψ, and one writes

$$\Psi' = A\Psi.$$

By definition A is a *linear operator* if its action on the function $\lambda_1\Psi_1 + \lambda_2\Psi_2$, a linear combination with constant (complex) coefficients, of two functions of this function space, is given by

$$A(\lambda_1\Psi_1 + \lambda_2\Psi_2) = \lambda_1(A\Psi_1) + \lambda_2(A\Psi_2).$$

Among the linear operators acting on the wave functions

$$\Psi \equiv \Psi(\mathbf{r}, t) \equiv \Psi(x, y, z, t)$$

associated with a particle, let us mention:

1) the differential operators $\partial/\partial x$, $\partial/\partial y$, $\partial/\partial z$, $\partial/\partial t$, such as the one which was considered above;

2) the operators of the form $f(\mathbf{r}, t)$ whose action consists in multiplying the function Ψ by the function $f(\mathbf{r}, t)$.

Starting from certain linear operators, one can form new linear operators by the following algebraic operations:

a) multiplication of an operator A by a constant c:

$$(cA)\Psi \equiv c(A\Psi);$$

b) the sum $S = A + B$ of two operators A and B:

$$S\Psi \equiv A\Psi + B\Psi;$$

c) the product $P = AB$ of an operator B by the operator A:

$$P\Psi \equiv AB\Psi \equiv A(B\Psi).$$

Note that in contrast to the sum, *the product of two operators is not commutative*. Therein lies a very important difference between

the algebra of linear operators and ordinary algebra. The product AB is not necessarily identical to the product BA; in the first case, B first acts on the function Ψ, then A acts upon the function $(B\Psi)$ to give the final result; in the second case, the roles of A and B are inverted. The difference $AB-BA$ of these two quantities is called the *commutator* of A and B; it is represented by the symbol $[A, B]$:

$$[A, B] \equiv AB - BA. \tag{II.8}$$

If this difference vanishes, one says that the two operators commute:

$$AB = BA.$$

As an example of operators which do not commute, we mention the operator $f(x)$, multiplication by function $f(x)$, and the differential operator $\partial/\partial x$. Indeed we have, for any Ψ,

$$\frac{\partial}{\partial x} f(x)\Psi = \frac{\partial}{\partial x}(f\Psi) = \frac{\partial f}{\partial x}\Psi + f\frac{\partial \Psi}{\partial x} = \left(\frac{\partial f}{\partial x} + f\frac{\partial}{\partial x}\right)\Psi.$$

In other words

$$\left[\frac{\partial}{\partial x}, f(x)\right] = \frac{\partial f}{\partial x} \tag{II.9}$$

and, in particular

$$\left[\frac{\partial}{\partial x}, x\right] = 1. \tag{II.10}$$

However, any pair of derivative operators such as $\partial/\partial x$, $\partial/\partial y$, $\partial/\partial z$, $\partial/\partial t$, commute.

A typical example of a linear operator formed by sum and product of linear operators is the Laplacian operator

$$\triangle \equiv \text{div grad} \equiv (\nabla \cdot \nabla) \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

which one may consider as the scalar product of the vector operator gradient $\nabla \equiv (\partial/\partial x, \partial/\partial y, \partial/\partial z)$ by itself.

12. Wave Equation of a Free Particle

The theory of matter waves leads unambiguously to the wave equation of a free particle (in non-relativistic approximation). Indeed, the wave $\Psi(\boldsymbol{r}, t)$ is a superposition:

$$\Psi(\boldsymbol{r}, t) = \int F(\boldsymbol{p})\, e^{i(\boldsymbol{p}\cdot\boldsymbol{r} - Et)/\hbar}\, d\boldsymbol{p} \tag{II.11}$$

of monochromatic plane waves $\exp[i(\boldsymbol{p}\cdot\boldsymbol{r} - Et)/\hbar]$ whose frequency $E/\hbar$ is connected with the wave vector $\boldsymbol{p}/\hbar$ by the relation connecting momentum and energy for a particle of mass m

$$E = \frac{p^2}{2m}. \tag{II.12}$$

Taking the partial derivatives of the two sides of equation (II.11), – we omit questions of convergence since mathematical rigor is of no concern to us in this argument – one obtains successively:

$$i\hbar \frac{\partial}{\partial t} \Psi(\boldsymbol{r}, t) = \int E F(\boldsymbol{p})\, e^{i(\boldsymbol{p}\cdot\boldsymbol{r} - Et)/\hbar}\, d\boldsymbol{p} \tag{II.13}$$

$$\frac{\hbar}{i} \nabla \Psi(\boldsymbol{r}, t) = \int \boldsymbol{p} F(\boldsymbol{p})\, e^{i(\boldsymbol{p}\cdot\boldsymbol{r} - Et)/\hbar}\, d\boldsymbol{p} \tag{II.14}$$

$$-\hbar^2 \triangle \Psi(\boldsymbol{r}, t) = \int p^2 F(\boldsymbol{p})\, e^{i(\boldsymbol{p}\cdot\boldsymbol{r} - Et)/\hbar}\, d\boldsymbol{p}. \tag{II.15}$$

According to relation (II.12), the expressions under the integral signs of equations (II.13) and (II.15) are proportional; therefore the integrals themselves differ by the same proportionality factor. Consequently

$$i\hbar \frac{\partial}{\partial t} \Psi(\boldsymbol{r}, t) = -\frac{\hbar^2}{2m} \triangle \Psi(\boldsymbol{r}, t). \tag{II.16}$$

This is the Schrödinger equation for a free particle; it satisfies conditions (A) and (B); from the very manner in which it was obtained it also satisfies the requirements of the correspondence principle. Indeed the formal analogy with Classical Mechanics is actually realized: equation (II.16) is in a sense the quantum-mechanical translation of the classical equation (II.12), the energy and momentum being represented in this quantum language by differential operators acting on the wave function according to the correspondence rule

$$E \to i\hbar \frac{\partial}{\partial t}, \qquad \boldsymbol{p} \to \frac{\hbar}{i} \nabla. \tag{II.17}$$

Thus the quantity $\boldsymbol{p}^2 = p_x^2 + p_y^2 + p_z^2$ is represented by the operator

$$-\hbar^2 \triangle = \left(\frac{\hbar}{i}\right)^2 \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right).$$

Just like relation (II.12) from which it originated, equation (II.16) obviously does not satisfy the principle of relativity. On the other hand, the de Broglie theory does not suffer from this limitation. To obtain a relativistic equation of the free particle, one may try to repeat the preceding argument, replacing relation (II.12) by a relation between energy and momentum in conformity with the theory of relativity. The correct relation $E=\sqrt{p^2c^2+m^2c^4}$ is not suitable because of the presence of the square root. To avoid that difficulty, one can use the relation

$$E^2=p^2c^2+m^2c^4 \tag{II.18}$$

from which one deduces the equation

$$-\hbar^2\frac{\partial^2}{\partial t^2}\Psi=-\hbar^2c^2\triangle\Psi+m^2c^4\Psi,$$

which may also be written

$$\left[\Box+\left(\frac{mc}{\hbar}\right)^2\right]\Psi(\boldsymbol{r},t)=0 \tag{II.19}$$

making use of the Dalembertian operator

$$\Box\equiv\frac{1}{c^2}\frac{\partial^2}{\partial t^2}-\triangle.$$

One again finds the same formal correspondence between equations (II.18) and (II.19), as the one which exists between equations (II.12) and (II.16).

Equation (II.19), the so-called *Klein-Gordon equation*, plays an important role in Relativistic Quantum Theory. As it does not satisfy criterion (B), it cannot be adopted as wave equation without a physical reinterpretation of the wave Ψ. Actually, the fact that a wave can represent the dynamical state of one and only one particle is fully justified only in the non-relativistic limit, i.e. when the law of conservation of the number of particles is satisfied. Hence we shall restrict ourselves from now on to the search for a non-relativistic wave equation.

13. Particle in a Scalar Potential

In order to form the wave equation of a particle in a potential $V(\boldsymbol{r})$, we operate at first under the conditions of the "geometrical

optics approximation" and seek to form an equation of propagation for a wave packet $\Psi(\mathbf{r}, t)$ moving in accordance with the de Broglie theory.

The center of the packet travels like a classical particle whose position, momentum, and energy we shall designate by $\mathbf{r}_{\text{cl.}}$, $\mathbf{p}_{\text{cl.}}$, and $E_{\text{cl.}}$, respectively. These quantities are connected by the relation

$$E_{\text{cl.}} = H(\mathbf{r}_{\text{cl.}}, \mathbf{p}_{\text{cl.}}) \equiv \frac{p^2_{\text{cl.}}}{2m} + V(\mathbf{r}_{\text{cl.}}) \tag{II.20}$$

$H(\mathbf{r}_{\text{cl.}}, \mathbf{p}_{\text{cl.}})$ is the classical Hamiltonian. We suppose that $V(\mathbf{r})$ does not depend upon the time explicitly (conservative system), although this condition is not absolutely necessary for the present argument to hold. Consequently $E_{\text{cl.}}$ remains constant in time, while $\mathbf{r}_{\text{cl.}}$ and $\mathbf{p}_{\text{cl.}}$ are well-defined functions of t. Under the approximate conditions considered here, $V(\mathbf{r})$ remains practically constant over a region of the order of the size of the wave packet; therefore

$$V(\mathbf{r})\, \Psi(\mathbf{r}, t) \approx V(\mathbf{r}_{\text{cl.}})\, \Psi(\mathbf{r}, t). \tag{II.21}$$

On the other hand, if we restrict ourselves to time intervals sufficiently short so that the relative variation of $\mathbf{p}_{\text{cl.}}$ remains negligible, $\Psi(\mathbf{r}, t)$ may be considered as a superposition of plane waves of the type (II.11), whose frequencies are in the neighborhood of $E_{\text{cl.}}/\hbar$ and whose wave vectors lie close to $\mathbf{p}_{\text{cl.}}/\hbar$. Therefore

$$\begin{aligned} i\hbar \frac{\partial}{\partial t} \Psi(\mathbf{r}, t) &\approx E_{\text{cl.}}\, \Psi(\mathbf{r}, t) \\ \frac{\hbar}{i} \nabla \Psi(\mathbf{r}, t) &\approx \mathbf{p}_{\text{cl.}}(t)\, \Psi(\mathbf{r}, t) \end{aligned} \tag{II.22}$$

and taking the divergence of this last expression, one obtains

$$-\hbar^2 \triangle \Psi(\mathbf{r}, t) \approx p^2_{\text{cl.}}\, \Psi(\mathbf{r}, t). \tag{II.23}$$

Combining the relations (II.21), (II.22), and (II.23) and making use of equation (II.20), we obtain

$$i\hbar \frac{\partial}{\partial t} \Psi + \frac{\hbar^2}{2m} \triangle \Psi - V\Psi \approx \left(E_{\text{cl.}} - \frac{p^2_{\text{cl.}}}{2m} - V(\mathbf{r}_{\text{cl.}}) \right) \Psi \approx 0.$$

The wave packet $\Psi(\mathbf{r}, t)$ satisfies – at least approximately – a wave equation of the type we are looking for. We are very naturally

led to adopt this equation as the wave equation of a particle in a potential, and we postulate that in all generality, even when the conditions for the "geometrical optics" approximation are not fulfilled, the wave Ψ satisfies the equation

$$i\hbar \frac{\partial}{\partial t} \Psi(\mathbf{r}, t) = \left(-\frac{\hbar^2}{2m} \Delta + V(\mathbf{r}) \right) \Psi(\mathbf{r}, t). \tag{II.24}$$

It is the Schrödinger equation for a particle in a potential $V(\mathbf{r})$.

14. Charged Particle in an Electromagnetic Field

The above argument may be repeated in more general cases where the potential V is an explicit function of time, or in the case of a particle with charge e in an electromagnetic field derived from a vector potential $\mathbf{A}(\mathbf{r}, t)$ and a scalar potential $\varphi(\mathbf{r}, t)$. In the latter case, the classical relation (II.20) must be replaced (cf. Problem I.4) by the relation

$$E = \frac{1}{2m} \left(\mathbf{p} - \frac{e}{c} \mathbf{A}(\mathbf{r}, t) \right)^2 + e\varphi(\mathbf{r}, t). \tag{II.25}$$

Considerations of the behavior of wave packets on the "geometrical optics" approximation lead us to the wave equation

$$i\hbar \frac{\partial}{\partial t} \Psi(\mathbf{r}, t) = \left[\frac{1}{2m} \left(\frac{\hbar}{i} \nabla - \frac{e}{c} \mathbf{A} \right)^2 + e\varphi \right] \Psi(\mathbf{r}, t). \tag{II.26}$$

It is the Schrödinger equation of a charged particle in an electromagnetic field [1]).

---

[1]) On the right-hand side of equation (II.26), the operator $\left(\frac{\hbar}{i} \nabla - \frac{e}{c} \mathbf{A} \right)^2$ designates the scalar product of the vector operator $\frac{\hbar}{i} \nabla - \frac{e}{c} \mathbf{A}$ by itself; in other words, the function which results from its action on Ψ is the sum of the expression

$$\left(\frac{\hbar}{i} \frac{\partial}{\partial x} - \frac{e}{c} A_x \right) \left(\frac{\hbar}{i} \frac{\partial}{\partial x} - \frac{e}{c} A_x \right) \Psi \equiv -\hbar^2 \frac{\partial^2 \Psi}{\partial x^2} - \frac{e\hbar}{ic} \left(A_x \frac{\partial \Psi}{\partial x} + \frac{\partial}{\partial x} (A_x \Psi) \right) + \frac{e^2}{c^2} A_x^2 \Psi$$

and of two other expressions which are obtained from it by substituting y and z for x, namely

$$-\hbar^2 \Delta \Psi - 2 \frac{e\hbar}{ic} (\mathbf{A} \cdot \nabla \Psi) + \left(-\frac{e\hbar}{ic} \operatorname{div} \mathbf{A} + \frac{e^2}{c^2} A^2 \right) \Psi.$$

In all of this one must realize that the components of the operator ∇ and those of the operator $\mathbf{A}$ do not in general commute with each other.

Equations (II.24) and (II.26) are the generalizations of equation (II.16) and the same remarks apply to them. They are indeed linear, homogeneous, partial differential equations of the first order in the time [conditions (A) and (B)]. Furthermore, they can be deduced from the classical relations (II.20) and (II.25), respectively, by the correspondence relation defined by (II.17).

15. General Rule for Forming the Schrödinger Equation by Correspondence

Generalizing this correspondence operation, one can formulate a systematic method for constructing the Schrödinger equation, which can be applied to the most general systems.

Consider a classical dynamical system whose Hamiltonian is $H(q_1, \ldots, q_R; p_1, \ldots, p_R; t)$. This function depends on the coordinates $q_1, \ldots, q_R$ of the system in configuration space, on their respective momenta $p_1, \ldots, p_R$, and on the time t. The total energy E of the system is

$$E = H(q_1, \ldots, q_R; p_1, \ldots, p_R; t). \tag{II.27}$$

To this classical system corresponds a quantum system whose dynamical state is represented by a wave function $\Psi(q_1, \ldots, q_R; t)$ *defined in configuration space* and whose wave equation can be obtained by performing on both sides of equation (II.27) the substitutions

$$E \to \mathrm{i}\hbar \frac{\partial}{\partial t} \qquad p_r \to \frac{\hbar}{\mathrm{i}} \frac{\partial}{\partial q_r} \qquad (r = 1, 2, \ldots, R) \tag{II.28}$$

and by writing down that these two quantities, considered as operators, give identical results when acting on Ψ. The equation thus obtained is the *Schrödinger equation* of the corresponding quantum system:

$$\mathrm{i}\hbar \frac{\partial}{\partial t} \Psi(q_1, \ldots, q_R; t) = \\ = H\left(q_1, \ldots, q_R; \frac{\hbar}{\mathrm{i}} \frac{\partial}{\partial q_1}, \ldots, \frac{\hbar}{\mathrm{i}} \frac{\partial}{\partial q_R}; t\right) \Psi(q_1, \ldots, q_R; t). \tag{II.29}$$

The operator $H\left(q_1, \ldots, q_R; \frac{\hbar}{\mathrm{i}} \frac{\partial}{\partial q_1}, \ldots, \frac{\hbar}{\mathrm{i}} \frac{\partial}{\partial q_R}; t\right)$ is called the *Hamiltonian* of the system under consideration.

It is important to note that the correspondence rule stated above does not define the Schrödinger equation uniquely. Indeed, there exist two causes of ambiguity.

The first cause of ambiguity comes from the fact that this rule is not invariant under a change of coordinates of configuration space. To illustrate this point, take the simple case of a free particle in two-dimensional space. From the classical Hamiltonian

$$\frac{p_x^2 + p_y^2}{2m}$$

in cartesian coordinates, one deduces the equation

$$i\hbar \frac{\partial}{\partial t} \Psi(x, y; t) = -\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) \Psi(x, y; t).$$

If by carrying out a change of variables one goes over to polar coordinates (r, φ), one obtains, after a straightforward calculation, the following equation for the wave function Ψ considered as a function of these new coordinates:

$$i\hbar \frac{\partial}{\partial t} \Psi(r, \varphi; t) = -\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \varphi^2} \right) \Psi(r, \varphi; t).$$

If, on the other hand, one applies correspondence rule (II.28) directly to the classical Hamiltonian in polar coordinates

$$\frac{1}{2m} \left(p_r^2 + \frac{p_\varphi^2}{r^2} \right)$$

one obtains a different equation, namely

$$i\hbar \frac{\partial}{\partial t} \Psi(r, \varphi; t) = -\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial r^2} + \frac{1}{r^2} \frac{\partial^2}{\partial \varphi^2} \right) \Psi(r, \varphi; t).$$

In order to avoid this type of ambiguity, we adopt the convention not to apply rule (II.28) unless the coordinates q are cartesian coordinates[1]).

[1]) This convention is not arbitrary. It automatically ensures the invariance of the form of the Schrödinger equation under a rotation of axes. In fact, one can remove this restriction and formulate the correspondence rule in covariant form by adopting a suitable metric in configuration space, and by replacing the operation $\partial/\partial q$ in (II. 28) by the covariant derivative. In this connection, see Brillouin, L., *Les Tenseurs en Mécanique et en Elasticité* (Paris, Masson, 2nd. ed, 1949), p. 200; also (New York, Dover, 1946).

The second cause of ambiguity stems from the fact that rule (II.28) substitutes for quantities obeying the rules of ordinary algebra, operators which do not all commute with each other. As a consequence different Hamiltonians may correspond to equivalent forms of the classical Hamiltonian. Thus, to the two equivalent expressions for the kinetic energy (in a one-dimensional problem)

$$\frac{p^2}{2m} \quad \text{and} \quad \frac{1}{2m}\frac{1}{\sqrt{q}}\,pqp\,\frac{1}{\sqrt{q}},$$

there respectively correspond the operators

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial q^2} \text{ and } -\frac{\hbar^2}{2m}\left(\frac{1}{\sqrt{q}}\frac{\partial}{\partial q}\cdot q\cdot\frac{\partial}{\partial q}\cdot\frac{1}{\sqrt{q}}\right) \equiv -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial q^2}+\frac{1}{4q^2}\right)$$

which differ by the quantity $\hbar^2/8mq^2$.

No rule based on the correspondence with Classical Mechanics can resolve such ambiguities, since the latter arise from the non-commutability of operators, which in turn is tied to the finite and non-zero character of $\hbar$. One must therefore fix empirically the precise form of the Hamiltonian function to which one applies rule (II.28). In all cases of practical interest, one must conform to the following prescriptions:

In *cartesian* coordinates, the classical Hamiltonian is the sum of a quadratic expression in the p's completely independent of the q's, a function depending exclusively upon the q's, and possibly a linear function of the p's of the form $\sum_i p_i f_i(q_1, \ldots, q_R)$. Having put the Hamiltonian function into this form, one replaces the last term of the summation by the "symmetrized" expression

$$\tfrac{1}{2}\sum_i [p_i f_i(q_1, \ldots, q_R) + f_i(q_1, \ldots, q_R) p_i];$$

the correspondence rule (II.28) must be applied to this expression.

The "symmetrization" of the linear terms in p is, as we shall see in Chapter IV, a necessary condition for the statistical interpretation of the wave function to be consistent. The case of a particle in an electromagnetic field [eqs. (II.25) and (II.26)] is an example of a system where this procedure must be carried out.

To conclude this section consider the following application. We

propose to form the Schrödinger equation of a complex atom, formed by a nucleus of charge Ze and of mass M, and by Z electrons of charge $-e$ and mass m. The Hamiltonian function is made up of $Z+1$ kinetic energy terms, Z terms of Coulomb attraction of the Z electrons by the nucleus, and $\frac{1}{2}Z(Z-1)$ terms of mutual Coulomb repulsion between each pair of electrons; thus, adopting the notation of § 9,

$$\frac{P^2}{2M}+\sum_{i=1}^{Z}\frac{p_i^2}{2m}-\sum_{i=1}^{Z}\frac{Ze^2}{|\boldsymbol{R}-\boldsymbol{r}_i|}+\sum_{i<j}\frac{e^2}{|\boldsymbol{r}_i-\boldsymbol{r}_j|}.$$

From it we derive the Schrödinger equation

$$\begin{aligned}i\hbar\frac{\partial}{\partial t}\Psi(\boldsymbol{R},\boldsymbol{r}_1,\ldots,\boldsymbol{r}_Z;t)=\\=\left[-\hbar^2\left(\frac{\triangle_R}{2M}+\sum_{i=1}^{Z}\frac{\triangle_i}{2m}\right)-\sum_{i=1}^{Z}\frac{Ze^2}{|\boldsymbol{R}-\boldsymbol{r}_i|}+\sum_{i<j}\frac{e^2}{|\boldsymbol{r}_i-\boldsymbol{r}_j|}\right]\Psi\end{aligned}\tag{II.30}$$

where the operator $\triangle_R$ designates the Laplacian with respect to the position vector $\boldsymbol{R}$, i.e. $\partial^2/\partial X^2+\partial^2/\partial Y^2+\partial^2/\partial Z^2$, and the operator $\triangle_i$ denotes the Laplacian with respect to the position vector of the ith electron $\boldsymbol{r}_i$.

In the case of the hydrogen atom ($Z=1$) the equation reads

$$i\hbar\frac{\partial}{\partial t}\Psi(\boldsymbol{r}_p,\boldsymbol{r}_e;t)=\left(-\frac{\hbar^2}{2M}\triangle_p-\frac{\hbar^2}{2m}\triangle_e-\frac{e^2}{|\boldsymbol{r}_p-\boldsymbol{r}_e|}\right)\Psi(\boldsymbol{r}_p,\boldsymbol{r}_e;t)\tag{II.31}$$

(here M is the mass of the proton, $\boldsymbol{r}_p$ its position, and $\boldsymbol{r}_e$ the position of the electron). As a first approximation, one can consider the proton to be infinitely heavy, and treat the hydrogen atom as an electron in an attractive Coulomb field $-e^2/r$, $\boldsymbol{r}$ designating the position of the electron in a coordinate system whose origin is located at the proton (assumed to be at rest). The wave function $\Psi(\boldsymbol{r},t)$ of the electron satisfies the Schrödinger equation:

$$i\hbar\frac{\partial}{\partial t}\Psi(\boldsymbol{r},t)=\left(-\frac{\hbar^2}{2m}\triangle-\frac{e^2}{r}\right)\Psi(\boldsymbol{r},t).\tag{II.32}$$

III. THE TIME-INDEPENDENT SCHRÖDINGER EQUATION

16. Search for Stationary Solutions

The Schrödinger equation of a quantum system is formally written

$$i\hbar\frac{\partial}{\partial t}\Psi=H\Psi.\tag{II.33}$$

Let us assume that the Hamiltonian H does not explicitly depend upon the time. This is the case of conservative systems, corresponding to classical systems whose energy is a constant of the motion. We look for a solution Ψ representing a dynamical state of well-defined energy E.

Such a wave Ψ must have a well-defined (angular) frequency ω, namely the one given by the Einstein relation $E=\hbar\omega$. Recall that this relationship between frequency of the wave and energy of the system constitutes the basic postulate of the theory of matter waves. Ψ is thus put into the form

$$\Psi = \psi e^{-iEt/\hbar} \tag{II.34}$$

where ψ depends upon the coordinates of configuration space but not upon the time. Substitution of this expression into equation (II.33) leads to the equation

$$H\psi = E\psi. \tag{II.35}$$

It is the so-called *time-independent Schrödinger equation.*

When the system is in a state represented by a wave of type (II.34) it is said to be in a *stationary* state of energy E; the time-independent wave function ψ is usually called the wave function of the stationary state, although the true wave function differs from the latter by the phase factor $\exp(-iEt/\hbar)$.

17. General Properties of the Equation. Nature of the Energy Spectrum

To simplify our presentation we shall discuss the special case of a particle of mass m in a scalar potential $V(\mathbf{r})$. We further assume that $V(\mathbf{r}) \to 0$ as $r \to \infty$. The function ψ is a function of the vector $\mathbf{r}(x, y, z)$ defining the position of the particle, and the time-independent Schrödinger equation is written

$$H\psi(\mathbf{r}) \equiv \left[-\frac{\hbar^2}{2m} \triangle + V(\mathbf{r}) \right] \psi(\mathbf{r}) = E\psi(\mathbf{r}). \tag{II.36}$$

In the language of the theory of partial differential equations, an equation of the type (II.36) is known as an *eigenvalue equation.* A solution $\psi_E(\mathbf{r})$ of this equation is called an *eigenfunction* (proper function) corresponding to the eigenvalue E of the operator H.

In fact, this eigenvalue problem is not well defined unless one specifies the conditions of "regularity" and the boundary conditions

which the function ψ must satisfy. The conditions to be imposed upon $\psi(\boldsymbol{r})$ must of course be compatible with the interpretation given to the wave function. We shall return to this point in Chapter IV. Let us agree for the time being that $\psi(\boldsymbol{r})$ and its partial derivatives of the first order shall be *continuous*, *uniform*, and *bounded* functions over all space.

One can then show the following results, which we shall accept without proof and verify later on in numerous examples.

a) If $E<0$, equation (II.36) has solutions only for certain particular values of E forming a *discrete spectrum*. The eigenfunction $\psi(\boldsymbol{r})$ corresponding to it – or each of the eigenfunctions when several exist – vanishes at infinity. More precisely, the integral $\int|\psi(\boldsymbol{r})|^2 \mathrm{d}\boldsymbol{r}$ extended over the entire configuration space is convergent. Following the statistical interpretation, there is a vanishing probability of finding the particle at infinity, and the particle remains practically localized in a finite region. The particle is said to be in a *bound state*.

b) If $E>0$, equation (II.36) may be solved for any positive value of E: the positive energies form a *continuous spectrum*. However, the corresponding eigenfunctions do not vanish at infinity; their asymptotic behavior is analogous to that of the plane wave $\exp(i\boldsymbol{k}\cdot\boldsymbol{r})$. More precisely, when $r\to\infty$ the absolute value $|\psi(\boldsymbol{r})|$ approaches a non-zero constant, or oscillates indefinitely between limits, one of which at least is not zero. The particle does not remain localized in a finite region. Wave functions of this type are used in collision problems; one is dealing with a so-called *unbound state*, or stationary state of collision.

We thus obtain a first fundamental result: the quantization of the energy of bound states, one of the most striking facts among those which caused the abandonment of the Classical Theory. The determination of the quantized energies appears here as an eigenvalue problem. To solve this problem as accurately as possible is one of the central problems of Wave Mechanics. For some particularly simple forms of the Hamiltonian, it can be solved rigorously. This is especially the case of the hydrogen atom (we shall treat it in detail in Chapter XI) whose energy levels are the eigenvalues of the operator

$$[-(\hbar^2/2m)\triangle-(e^2/r)].$$

As will be shown in Chapter XI the eigenvalue spectrum is identical to that predicted by the Old Quantum Theory; we have already emphasized its extraordinary agreement with experiment. In more complex situations, one has to have recourse to suitable methods of approximation. In all the cases where the energy spectrum could be calculated with reasonable precision, the agreement with experimental results is as good as might be expected from a non-relativistic theory.

The eigensolution ψ_E itself may be to some extent subject to experimental check. Indeed, the eigenfunctions of the discrete spectrum enter into the calculations of various measurable quantities, such as transition probabilities. As to the eigenfunctions of the continuous spectrum, their asymptotic form is very directly related to the cross sections – characteristic parameters of the collision phenomena, whose precise definition will be given later on (Ch. X).

EXERCISES AND PROBLEMS

1. Let us try to observe the motion of the electron along a circular Bohr orbit of the hydrogen atom by carrying out several successive measurements of the electron position with sufficiently hard X-rays.

Evaluate the order of magnitude of the kinetic energy transfer ΔT to the electron in a collision with an X-ray photon as a function of the wavelength λ of the latter. To observe the motion along an orbit, λ must be much smaller than the radius of this orbit. Compare in that case the perturbation ΔT with the distance between neighboring energy levels. What must one conclude from this concerning the observability of Bohr orbits?

2. In relativistic quantum mechanics, the total energy E and the momentum p of a free particle of (rest) mass m and velocity v are respectively equal to

$$\frac{mc^2}{\sqrt{1-v^2/c^2}} \quad \text{and} \quad \frac{mv}{\sqrt{1-v^2/c^2}}.$$

Verify that the equations of motion may be written in the Hamiltonian form by taking for the Hamiltonian function $H = E = \sqrt{m^2c^4 + p^2c^2}$.

From this, deduce the equality between the velocity of this particle and the group velocity v_g of the associated de Broglie wave. Calculate the phase velocity v_φ of this wave: show that it is superior to the velocity c and that $v_g v_\varphi = c^2$.

3. Examine the soundness of a classical description of the motion of an atom in a diatomic molecule. To do this, assume that the atom executes harmonic oscillations of angular frequency ω with an average kinetic energy of the order of $\frac{1}{2}kT$, and compare the average wavelength of the atom with the amplitude of these oscillations. First, treat the case of the hydrogen molecule at ordinary temperatures: $T = 300°$ K, $kT = 0.025$ eV, $\hbar\omega = 0.5$ eV; then consider the

case of a molecule with heavy atoms of mass 200 times that of hydrogen assuming the restoring force to be the same as for hydrogen, at 300° K, and also at 10° K.

4. An electron follows a circular trajectory in a constant magnetic field $\boldsymbol{H}$. Apply to this rotational motion the de Broglie resonance condition. Show that the kinetic energy of the electron is quantized, that the energy levels are equidistant, and that the distance between levels is equal to $(e\hbar/mc)\boldsymbol{H}$ [this result differs from that of the rigorous quantum theory only by an overall displacement of all the energy levels by an amount $(e\hbar/2mc)\boldsymbol{H}$.] Calculate the radius, the momentum and the kinetic energy of the quantized trajectories for a field of 10^4 gauss. Compare the radius of the orbit of quantum number unity to that of the Bohr orbit of the ground state of the hydrogen atom.

[*N.B.* Be sure to distinguish in this problem between the momentum $\boldsymbol{p}$ and the quantity $m\boldsymbol{v}$. If $\boldsymbol{A}$ is the vector potential from which the magnetic field is derived, $\boldsymbol{p} = m\boldsymbol{v} + (e/c)\boldsymbol{A}$.]

5. Utilizing the fact that any wave can be considered as a superposition of plane waves, show that in the absence of a field, the matter wave $\psi(\boldsymbol{r}_2, t_2)$ at the point $\boldsymbol{r}_2$ at the instant t_2 can be deduced from the values $\psi(\boldsymbol{r}_1, t_1)$ taken by the wave at the instant t_1, by the operation

$$\psi(\boldsymbol{r}_2, t_2) = \int K(\boldsymbol{r}_2 - \boldsymbol{r}_1; t_2 - t_1)\, \psi(\boldsymbol{r}_1, t_1)\, \mathrm{d}\boldsymbol{r}_1 \tag{1}$$

where

$$K(\boldsymbol{\rho}; \tau) = (2\pi\hbar)^{-3} \int \exp\left[\frac{\mathrm{i}}{\hbar}(\boldsymbol{p}\cdot\boldsymbol{\rho} - E\tau)\right] \mathrm{d}\boldsymbol{p},$$

an expression in which E, a function of $\boldsymbol{p}$, is equal to the energy of the particle, corresponding to the momentum $\boldsymbol{p}$. Show that for a non-relativistic particle of mass m,

$$K(\boldsymbol{\rho}; \tau) = \mathrm{e}^{-\frac{1}{2}\pi\mathrm{i}} \left(\frac{m}{2\pi\hbar\tau}\right)^{\frac{3}{2}} \exp\left(\mathrm{i}\frac{m\rho^2}{2\hbar\tau}\right).$$

Deduce from this that the main contribution to integral (1) giving the wave function at $\boldsymbol{r}_2$ at time t_2 comes from a region surrounding the point $\boldsymbol{r}_2$ whose radius is of the order of

$$\left[\frac{2\hbar(t_2 - t_1)}{m}\right]^{\frac{1}{2}}.$$

6. How does the method of the preceding problem have to be modified in order to apply to a particle in one dimension? Using this method, determine the wave function at the instant t of a free, non-relativistic particle of mass m whose wave function at the instant $t = 0$ is

$$\psi(x, 0) = (\pi\xi_0^2)^{-\frac{1}{4}} \exp(\mathrm{i}p_0 x/\hbar) \exp(-x^2/2\xi_0^2).$$

The intensity (modulus squared) of that wave at the instant $t = 0$ is a Gaussian of width ξ_0. Show that the shape remains Gaussian at all later instants, but that its width increases according to the law

$$\xi = \xi_0 \left(1 + \frac{\hbar^2 t^2}{m^2 \xi_0^4}\right)^{\frac{1}{2}}$$

(spreading of the wave packet, cf. Ch. VI).

CHAPTER III

ONE-DIMENSIONAL QUANTIZED SYSTEMS

1. Introduction

In order to acquire a certain amount of practice in dealing with the Schrödinger equation before we attack problems of interpretation of the Quantum Theory, we shall study the wave mechanics of one-dimensional systems. One-dimensional problems are of interest, not only as simple models allowing us to display a certain number of properties which one encounters again in more complex situations, but also because in a number of problems one is led, after some suitable manipulations, to equations of the same type as the one-dimensional Schrödinger equation.

Let us consider the motion of a particle of mass m constrained to move on the x axis in a certain potential $V(x)$. Its Schrödinger equation is:

$$i\hbar \frac{\partial}{\partial t} \Psi(x, t) = \left(- \frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) \right) \Psi(x, t). \tag{III.1}$$

We shall especially be concerned with the search for the stationary states. If E is the energy of the stationary state, we have

$$\Psi(x, t) = \psi(x)\, e^{-iEt/\hbar} \tag{III.2}$$

and the function $\psi(x)$ is a solution of the time-independent Schrödinger equation

$$\left(- \frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) \right) \psi = E\psi. \tag{III.3}$$

Throughout this chapter we adopt the notation

$$V(x) = \frac{\hbar^2}{2m} U(x), \qquad E = \frac{\hbar^2}{2m} \varepsilon \tag{III.4}$$

which allows us to rewrite the preceding equation in the form

$$\psi'' + [\varepsilon - U(x)]\psi = 0. \tag{III.5}$$

It is a differential equation of the Sturm-Liouville type for which we intend to find solutions which are finite, continuous and differentiable over the entire interval $(-\infty, +\infty)$.

If such a solution exists, all multiples of that solution are solutions as well; we shall not consider two solutions to be distinct if they differ only by a constant multiplicative factor. If two linearly independent solutions are acceptable solutions, any linear combination thereof is a solution as well; the eigenvalue is said to be *degenerate* of the second order. The order of degeneracy is by definition the number of linearly independent eigenfunctions.

Equation (III.5) is real [$V(x)$ is a real function of x]. If ψ is an eigenfunction, so are its real part and its imaginary part (the latter are necessarily multiples of one another if the eigenfunction is not degenerate). As a consequence one merely needs to know the real eigenfunctions to construct all the eigenfunctions belonging to a given eigenvalue. This remark simplifies the calculations considerably.

In the first section, we solve the eigenvalue problem exactly for some simple square potentials. We shall dwell upon the principal points of difference between classical and quantized motions, especially the quantization of the energies of bound states, and the phenomena of wave reflection, resonance and potential-barrier penetration in the motion of unbound "particles". In the second section, we shall make a systematic study of equation (III.5) for an arbitrary potential $U(x)$, and show that many results obtained with square potentials are quite general.

I. SQUARE POTENTIALS

2. General Remarks

In order that typical quantum effects may appear, the potential $V(x)$ must possess an appreciable relative variation over a distance of the order of a wavelength. The simplest type of potential fulfilling these conditions is the *square potential*; it is a potential which exhibits discontinuities of the first kind (that is to say, sudden jumps by a finite amount) at certain points, and remains constant everywhere else. The x axis is thus subdivided into n intervals, say, in each of which the potential remains constant.

The existence of discontinuities of the first kind in the potential

$U(x)$ does not modify the conditions of regularity imposed upon the function ψ. Indeed, according to the Schrödinger equation

$$\psi'' = (U - \varepsilon)\psi.$$

At each discontinuity of the potential, U and consequently ψ'' exhibit a sudden jump by a finite amount, but the integral of ψ'' remains continuous at these points: ψ' and *a fortiori* ψ are therefore continuous everywhere.

Now, let U_i be the (constant) value of $U(x)$ in the ith region ($i = 1, 2, \ldots, n$). The general solution in this region is a linear combination of exponentials. Its behavior is in fact very different according to whether $(\varepsilon - U_i)$ is positive or negative.

If $\varepsilon > U_i$, it is a linear combination of imaginary exponentials

$$e^{ik_i x} \quad \text{and} \quad e^{-ik_i x} \qquad (k_i = \sqrt{\varepsilon - U_i})$$

or, equivalently, a combination of sine and cosine: it has an "oscillatory" behavior.

If $\varepsilon < U_i$, we have a combination of real exponentials

$$e^{\varkappa_i x} \quad \text{and} \quad e^{-\varkappa_i x} \qquad (\varkappa_i = \sqrt{U_i - \varepsilon}).$$

In that case, we shall say that the solution has an "exponential" behavior.

To obtain the general solution of the differential equation, one writes it in the form of a linear combination of (real or imaginary) exponentials for each of the n regions where the potential is constant. The parameters of these combinations ($2n$ in number) are fixed by the conditions of continuity of the function and its derivative at the points of discontinuity of the potential. This yields $2(n-1)$ conditions since there are $(n-1)$ points of discontinuity. Therefore the general solution which one thus obtains depends upon two arbitrary parameters, as expected. In order to be acceptable as eigenfunction, the solution must remain bounded everywhere, i.e. bounded at both of the limits $x = +\infty$ and $x = -\infty$. Note that if the energy remains lower than the potential over the entire interval $(-\infty, +\infty)$, the general solution has an exponential behavior everywhere; its second derivative is always of the same sign as the function itself. From this, one readily deduces that the latter increases exponentially at one of the limits $-\infty$ or $+\infty$, and possibly at both limits. The eigenvalue

problem has, therefore, no solution. In classical mechanics too, motion is possible only if the energy exceeds the potential in at least part of the interval.

If ε exceeds at least one of the quantities U_i, the existence and the number of eigenfunctions essentially depend upon the exponential or oscillatory character of the general solution at the two extremities of the x axis $-\infty$ and $+\infty$.

3. Potential Step. Reflection and Transmission of Waves

The simplest example of a square potential is that of a sudden jump of the potential ($n=2$) such as the one shown in Fig. III.1:

$$U(x)=\begin{cases} U_1 & \text{if} \quad x>0 \qquad \text{(region I)} \\ U_2 & \text{if} \quad x<0. \qquad \text{(region II)} \end{cases}$$

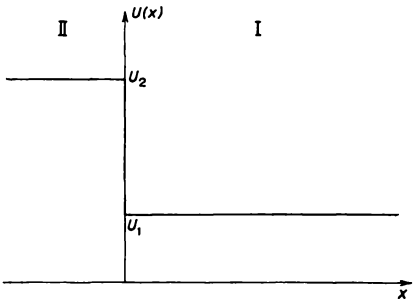


Fig. III.1. Potential step.

For definiteness, we shall assume that $U_2>U_1$.

Two cases may occur:

a) $U_2>\varepsilon>U_1$. The general solution has an oscillatory behavior in region I ($x>0$), and an exponential behavior in region II ($x<0$). To be acceptable as an eigenfunction, it must be exponentially decreasing in region II. There is always one and only one solution fulfilling this condition. Each value of ε contained in this interval is thus a non-degenerate eigenvalue. The energy spectrum is *continuous and non-degenerate*. In each of the two regions this function takes on the general form

$$y=\begin{cases} A_1 \sin(k_1 x+\varphi) & x>0 \\ A_2 \, e^{\varkappa_2 x} & x<0. \end{cases} \qquad \text{(III.6)}$$

The conditions of continuity define y to within a constant. Instead of writing down the continuity condition of the function and its derivative, it is more convenient to write down the continuity of the function and of its *logarithmic derivative* y'/y. The continuity of the logarithmic derivative fixes the phase φ:

$$k_1 \cot \varphi = \varkappa_2.$$

φ is determined to within a multiple of π since a change from φ to $\varphi+\pi$ is equivalent to changing the sign of A_1. For the value of φ we take:

$$\varphi = \tan^{-1} \frac{k_1}{\varkappa_2} \tag{III.6a}$$

where $\tan^{-1}$ denotes the value of the arc tangent located in the interval

$$\left(-\frac{\pi}{2}, +\frac{\pi}{2}\right).$$

The continuity of the function determines the ratio A_2/A_1, i.e.

$$\frac{A_2}{A_1} = \sin \varphi = \frac{k_1}{\sqrt{k_1^2+\varkappa_2^2}} = \sqrt{\frac{\varepsilon - U_1}{U_2 - U_1}}. \tag{III.6b}$$

b) $\varepsilon > U_2$. The general solution has an oscillatory behavior over all space and therefore represents an acceptable eigenfunction. To each value of ε there thus correspond two linearly independent eigenfunctions: the spectrum of eigenvalues is *continuous* and its *degeneracy is of order* 2.

We shall form the eigenfunction whose character in region II is $\exp(-ik_2x)$. It is defined to within a constant which we fix by taking the coefficient of $\exp(-ik_1x)$ equal to unity in the expression for that function in region I.

In other words

$$\chi = \begin{cases} e^{-ik_1x} + R\, e^{ik_1x} & x > 0 \\ S\, e^{-ik_2x} & x < 0. \end{cases} \tag{III.7}$$

The (*a priori* complex) constants R and S are determined by the continuity conditions at $x=0$. The continuity of the logarithmic derivative gives

$$R = \frac{k_1 - k_2}{k_1 + k_2}. \tag{III.7a}$$

Continuity of the function itself yields

$$S = 1 + R = \frac{2k_1}{k_1 + k_2}. \tag{III.7b}$$

The function χ^* is an eigenfunction linearly independent of χ. All the eigenfunctions corresponding to the eigenvalue ε can thus be put in the form of a linear combination of χ and χ^*.

We shall compare the present situation to that which would obtain if the system were classical. The motion of a classical particle in this potential is very different in case (*a*) and in case (*b*).

In case (*a*) the classical motion is one of a particle of energy $(\hbar^2/2m)\varepsilon$ which, coming from $+\infty$ runs along the positive semi-axis with constant velocity $\hbar k_1/m$ in the direction of decreasing x, rebounds elastically at the point $x=0$ and starts out again in the opposite direction with the same velocity traveling to infinity. In order to realize an analogous phenomenon in wave mechanics, one must construct a wave packet by superposing waves of the form $y \exp(-iEt/\hbar)$ with neighboring energies. Rather than the function given by eq. (III.6), it is more convenient to use the wave

$$\psi_\varepsilon(x) = \begin{cases} e^{-ik_1 x} - e^{i(k_1 x + 2\varphi)} & x > 0 \\ \dfrac{2A_2}{iA_1} e^{i\varphi} e^{\varkappa_2 x} & x < 0 \end{cases} \tag{III.8}$$

obtained by dividing y by $\frac{1}{2}iA_1 \exp(-i\varphi)$. We write it with the subscript ε in order to remind us that it is an eigenfunction corresponding to the eigenvalue ε. Consider the wave packet

$$\Psi(x, t) = \int_0^\infty f(k_1' - k_1)\psi_{\varepsilon'}(x)\, e^{-iE't/\hbar}\, dk_1'. \tag{III.9}$$

The function $f(k_1' - k_1)$ is a real, more or less regular function of k_1', with a very pronounced peak at the point $k_1' = k_1$ (the "prime" no longer denotes the derivative with respect to x for the time being; the meaning of the quantities k_1', ε', E' and their relationships are self-evident). For the sake of simplicity, we further assume that $f(k_1' - k_1)$ vanishes when $k_1^2 > U_2 - U_1$. $\Psi(x, t)$ is thus a superposition of eigenfunctions of case (*a*), multiplied by the exponential $\exp(-iE't/\hbar)$ corresponding to their respective time dependences. By its

very construction Ψ is clearly a solution of the time-dependent Schrödinger equation. Its motion is easily exhibited if one refers back to the study of free wave packets of Chapter II.

In region I, $\Psi(x, t)$ is a superposition of two quantities: an "incident wave packet",

$$\Psi_{\mathrm{i}}(x, t) = \int_0^\infty f(k_1' - k_1)\, \mathrm{e}^{-\mathrm{i}k_1' x}\, \mathrm{e}^{-\mathrm{i}E't/\hbar}\, \mathrm{d}k_1' \qquad \text{(III.10}a\text{)}$$

whose center $x = -(1/\hbar)(\mathrm{d}E/\mathrm{d}k_1)t = -v_1 t$ travels at the velocity $v_1 = \hbar k_1/m$ in the negative direction and reaches the point $x = 0$ at the time $t = 0$; and a "reflected wave packet",

$$\Psi_{\mathrm{r}}(x, t) = -\int_0^\infty f(k_1' - k_1)\, \mathrm{e}^{\mathrm{i}[k_1' x + 2\varphi' - (E't/\hbar)]}\, \mathrm{d}k_1' \qquad \text{(III.10}b\text{)}$$

whose center $x = v_1 t - 2\,\mathrm{d}\varphi/\mathrm{d}k_1$ travels with the velocity v_1 in the opposite direction and leaves the origin at the time

$$\tau = \frac{2}{v_1}\frac{\mathrm{d}\varphi}{\mathrm{d}k_1} = 2\hbar\frac{\mathrm{d}\varphi}{\mathrm{d}E} \qquad \text{(III.11)}$$

later than the time $t = 0$ at which the center of the "incident wave packet" arrived there. The motion of the center of the packet is thus almost identical to that of the classical particle. The only difference consists in the *delay* τ suffered by the center of the packet upon reflection at the point of discontinuity $x = 0$, whereas the rebound of the classical particle is instantaneous. We note in this connection that the consideration of the motion of the center of the packet makes sense only if the shape of the packet is not too violently modified in the course of its motion. This is actually so for the incident wave packet as long as the center is at a distance from the origin large compared to the width Δx of the packet. For this to hold for the reflected wave packet as well, it is further necessary that the width Δk of the peak of the function f be sufficiently small so that φ does not vary appreciably over the region which contributes most to the integral (III.9), namely $\Delta k(\mathrm{d}\varphi/\mathrm{d}k_1) \ll 1$. As the spatial extension Δx of the packet is of the order $1/\Delta k$, this condition may also be written

$$\Delta x \gg \frac{\mathrm{d}\varphi}{\mathrm{d}k_1}. \qquad \text{(III.12)}$$

Consequently, $\Delta x/v_1 \gg \tau$. The wave packet is so wide that the time

it spends passing a point of the axis is clearly longer than the delay τ caused by the reflection.

Besides the delay τ there exists another difference between the motion of the classical particle and the reflection of the quantum-mechanical wave packet. The wave Ψ does not always vanish in region II. A study analogous to the one above shows that Ψ is equal to the product of $2A_2 \exp(\varkappa_2 x)/A_1$ and a quantity which takes on appreciable values during a certain time interval about the instant $t = \frac{1}{2}\tau$, a time which one can interpret as the collision time with the potential "wall" located at the point $x = 0$. Therefore at that moment there exists a non-vanishing probability of finding the particle in region II, while the classical particle never penetrates into that region.

Let us now examine case (*b*). There are two possible classical motions corresponding to the same value of the energy [1]). One is the motion of a particle running along the entire length of the x axis from $+\infty$ to $-\infty$; its velocity which is constant and equal to $v_1 = \hbar k_1/m$ in region I, suddenly jumps from v_1 to $v_2 = \hbar k_2/m$ at the point of discontinuity of the potential, and continues its motion to $-\infty$ at the velocity v_2. The other is the exactly reversed motion, namely that of a particle travelling along the x axis in the positive direction with velocity v_2 in region II, and velocity v_1 in region I.

Let us compare these classical motions to those of wave packets with similar initial conditions. We shall carry out that comparison for the first of these motions (displacement in the negative direction). Proceeding as in case (*a*), we form a wave packet analogous to the one of eq. (III.9) by superposition of eigenfunctions corresponding to neighboring values of the energy ε. Let us attach the subscript ε to the eigenfunction χ of the type (III.7) to remind ourselves that it depends upon the energy. *A priori*, the packet has to be formed by the superposition of functions χ_ε and χ_ε^*. But in order to realize the desired initial conditions, it must only contain the functions χ_ε as this investigation will show. Let us write therefore

$$\Psi(x, t) = \int_0^\infty f(k_1' - k_1)\chi_{\varepsilon'}(x)\, \mathrm{e}^{-\mathrm{i}E't/\hbar}\, \mathrm{d}k_1'.$$

The only difference with expression (III.9) is that the peak $k_1 = \sqrt{\varepsilon - U_1}$

[1]) This fact is to be compared with the existence of a degeneracy of order 2 in the corresponding quantum-mechanical problem.

of the function f lies in the energy region (b) instead of the energy region (a). The motion of the above wave packet is found in similar fashion to that of the wave packet (III.9). One readily verifies that the desired initial conditions are actually fulfilled, namely that when $t \ll 0$, $\Psi(x, t)$ remains practically zero in region II, and that in region I the only appreciable contribution comes from the term $\exp(-ik_1x)$. This yields a wave packet whose center $x = -v_1t$ moves like the classical particle at velocity v_1 in the direction of decreasing x and reaches the origin at time $t = 0$; later, $\Psi(x, t)$ splits into two packets, a "transmitted wave packet"

$$\Psi_t(x, t) = \int_0^\infty f(k_1' - k_1)\, S' e^{ik_1'x}\, e^{-iE't/\hbar}\, dk_1'$$

whose center $x = -v_2t$ rigorously continues the motion of the classical particle, and a "reflected wave packet",

$$\Psi_r(x, t) = \int_0^\infty f(k_1' - k_1)\, R' e^{ik_1'x}\, e^{-iE't/\hbar}\, dk_1'$$

whose center $x = v_1t$ moves as the classical particle would have if it had suffered an elastic collision at $x = 0$. Thus there is a very important difference with the classical motion: *the quantum "particle" has a non-vanishing probability of being "reflected" at the potential discontinuity.*

In order to pursue this analysis, one must define the probabilities in a precise manner; this will be done in Ch. IV. We merely indicate here without proof that the probability of finding the particle in the reflected wave is equal to $|R|^2$, that of finding it in the transmitted wave is equal to $(k_2/k_1)|S|^2$ (cf. Problem IV.2). These results are quite consistent, since the sum of these two quantities is equal to unity:

$$|R|^2 + \frac{k_2}{k_1}|S|^2 = 1 \tag{III.13}$$

as one may easily verify by substituting in that equation the values given by eqs. (III.7a) and (III.7b).

The quantity

$$T = \frac{k_2}{k_1}|S|^2 = \frac{4k_1k_2}{(k_1 + k_2)^2} \tag{III.14}$$

is called the *transmission coefficient* and measures the relative importance of the phenomenon of transmission. It increases with energy

and approaches unity as $\varepsilon \to \infty$. One may say that in that limit the result of classical mechanics holds.

Note that T is a symmetric function of k_1 and k_2. A wave of the same energy but propagating in the opposite direction (from II toward I) therefore has the same transmission coefficient: the transmission coefficient is independent of the direction of travel.

All these results are not surprising when we realize the great analogy which they bear to problems of lightwave propagation. The problem treated here is exactly the same as that of the propagation of a light signal in a non-absorbing medium of variable index of refraction. In the case (a), the index of refraction changes suddenly from a real value (medium I) to an imaginary value (medium II) at the point $x=0$: we have total reflection. In the case (b), the index remains real, and the media I and II have different indices of refraction: this sudden change of index causes a partial reflection of the signal.

4. Infinitely High Potential Barrier

A limiting case of the preceding one is that of a particle encountering an infinitely high potential barrier. For definiteness, let us suppose that $U(x)=+\infty$ when $x<0$. We are in a situation which is analogous to case (a) of the preceding problem as $U_2 \to +\infty$. Inspection of the wave function y as given by eqs. (III.6), (III.6a) and (III.6b), shows that it vanishes at the point $x=0$ in this limiting case ($\varkappa_2 \to \infty$).

We are dealing here with a general result which is valid no matter what the form of the function $V(x)$ in the region $x>0$. Indeed, the wave function necessarily takes the form $A \exp(\varkappa_2 x)$ in the region $x<0$; its logarithmic derivative is therefore $\varkappa_2$. In the limit where the potential V_2 tends toward infinity, $\varkappa_2$ also becomes infinite. Consequently, the function must have an infinite logarithmic derivative at the point $x=0$: thus, the function itself must vanish there.

In conclusion, *the wave function must vanish at the edge of an infinitely high potential barrier.*

5. Infinitely Deep Square Potential Well. Discrete Spectrum

As a second simple example, consider an infinitely deep square potential well. The value of the potential at the bottom of the well is taken as the origin for the energy scale. This region of zero potential

occupies a certain interval $(-L/2, +L/2)$ of the x axis; it is bounded on both sides by infinite potential barriers (Fig. III.2).

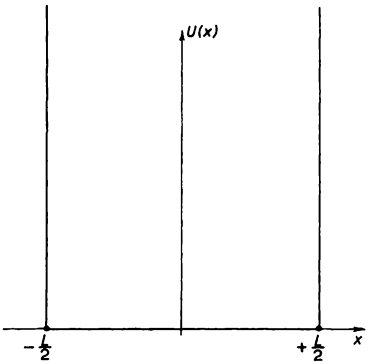


Fig. III.2. Infinitely deep square well.

The eigenvalue problem reduces to the search for a function ψ which vanishes at the points $+L/2$ and $-L/2$ and which satisfies the Schrödinger equation in the interval $(-L/2, +L/2)$

$$\psi'' + \varepsilon\psi = 0.$$

The general solution is a linear combination of $\sin kx$ and $\cos kx$ $(k = \sqrt{\varepsilon})$. Solutions satisfying simultaneously both boundary conditions exist for certain discrete values of ε, namely

$$\varepsilon_n = \frac{n^2\pi^2}{L^2} \qquad (n = 1, 2, \ldots, \infty) \tag{III.15}$$

(solutions for which $kL = n\pi$). To each of these eigenvalues there corresponds one and only one eigenfunction (no degeneracy), namely

if n is odd $$\psi_n = \cos\frac{n\pi}{L}x, \tag{III.16a}$$

if n is even $$\psi_n = \sin\frac{n\pi}{L}x. \tag{III.16b}$$

This very simple result calls for a certain number of remarks of a general nature.

In the first place, the present result differs profoundly from that obtained when dealing with a classical system. In the same potential,

a classical particle can actually move with any energy as long as it is positive; it is a (periodic) back-and-forth motion between the two potential walls located at the ends of the interval $(-L/2, +L/2)$. In wave mechanics, the motion takes place only for certain discrete values of the energy [1]: *the energy is quantized.*

The second remark concerns the *parity* of the eigenfunctions [2]. They are even if n is odd [eq. (III.16a)], odd if n is even [eq. (III.16b)]. The fact that the eigenfunctions have a definite parity is due to the potential $U(x)$ being invariant under reflection through the origin:

$$U(x) = U(-x).$$

The question of parity will be examined in all generality in § 14.

The last remark concerns the *number of nodes* of the eigenfunctions. By definition, the nodes are the zeros of the eigenfunction (except for those which lie at the ends $x = L/2$ and $x = -L/2$). The number of nodes increases regularly with the eigenvalue of the energy, increasing by unity when going from one eigenvalue to the one lying immediately above: the eigenfunction of the ground state ψ_1 has no nodes, ... that of the $(n-1)$th excited state ψ_n has $(n-1)$ nodes, etc. It is instructive to stress the analogy of this result with the one concerning the number of nodes of the stationary states of vibrating strings. Indeed, the analogy may be carried all the way since the two problems are mathematically identical.

6. Study of a Finite Square Well. Resonances

The results we have obtained with the potential step and the infinitely deep square potential well reappear in more complicated cases. As a new example, we consider the potential of Fig. III.3. Here the function $U(x)$ takes the form

$$U(x) = \begin{cases} U_1 & x > a \quad \text{(region I)} \\ U_2 & a > x > b \quad \text{(region II)} \\ U_3 & b > x \quad \text{(region III)} \end{cases}$$
$$(U_2 < U_1 < U_3).$$

---

[1] The period of the classical motion is in fact equal to $h/\Delta E$, where ΔE is the distance between neighboring energy levels, in accord with the correspondence principle.

[2] A function $f(x)$ is even if $f(x) = f(-x)$, and odd if $f(x) = -f(-x)$.

The eigenvalue problem appears in a different manner depending on the value of ε compared to the constants U_1, U_2, U_3.

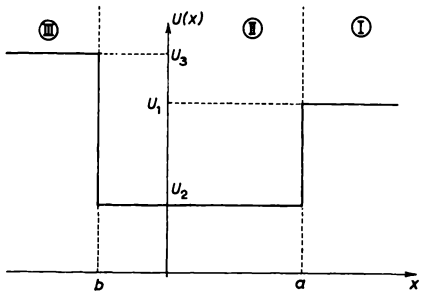


Fig. III.3. Finite square well.

a) $U_2 < \varepsilon < U_1$. DISCRETE SPECTRUM AND BOUND STATES

The general solution has an exponential behavior in the two external regions I and III and an oscillatory behavior in the interior region. In order to be acceptable as an eigenfunction, it must be exponentially decreasing in the two external regions. There exists one and only one exponentially decreasing solution in region I, and one and only one exponentially decreasing solution in region III; these two solutions join smoothly only for certain discrete values of ε. From this we conclude that the energy spectrum is certainly discrete and non-degenerate.

The function ψ is assumed real (cf. Ch. III, § 1) and assumes the following forms in each of the three regions

$$\psi = \begin{cases} A_1 \, e^{-\varkappa_1 x} & x > a \\ A_2 \sin(k_2 x + \varphi) & a > x > b \\ A_3 \, e^{\varkappa_3 x} & b > x. \end{cases} \qquad \text{(III.17)}$$

If the phase φ is known, the two continuity conditions of the function determine the constants A_1, A_2, A_3 (to within a factor). As for φ, it must simultaneously satisfy both continuity conditions of the logarithmic derivative

$$k_2 \cot(k_2 a + \varphi) = -\varkappa_1 \qquad k_2 \cot(k_2 b + \varphi) = \varkappa_3 \qquad \text{(III.18)}$$

which can also be written

$$\begin{aligned} \varphi &= -k_2 a - \tan^{-1}\frac{k_2}{\varkappa_1} + n\pi \quad (n \text{ positive integer}) \\ \varphi &= -k_2 b + \tan^{-1}\frac{k_2}{\varkappa_3} \end{aligned} \qquad \text{(III.19)}$$

(φ must be determined to within a multiple of π) we have fixed this arbitrariness by forcing $k_2 b + \varphi$ to lie within the interval $(-\pi/2, +\pi/2)$. This is possible if and only if the right-hand sides of these last two equations are equal. This equality can be achieved only for certain discrete values ε_n of ε, namely those which satisfy the equation

$$n\pi - k_2(a-b) = \tan^{-1}\frac{k_2}{\varkappa_1} + \tan^{-1}\frac{k_2}{\varkappa_3}. \qquad \text{(III.20)}$$

Let us introduce the following notations:

$$K = \sqrt{U_1 - U_2}, \quad L = b - a, \quad \cos\gamma = \sqrt{\frac{U_1 - U_2}{U_3 - U_2}} \qquad \left(0 < \gamma < \frac{\pi}{2}\right)$$

and the new variable

$$\xi = \frac{k_2}{K} = \sqrt{\frac{\varepsilon - U_2}{U_1 - U_2}}.$$

Equation (III.20) may also be written in the form of a condition on ξ:

$$n\pi - \xi KL = \sin^{-1}\xi + \sin^{-1}(\xi\cos\gamma).$$

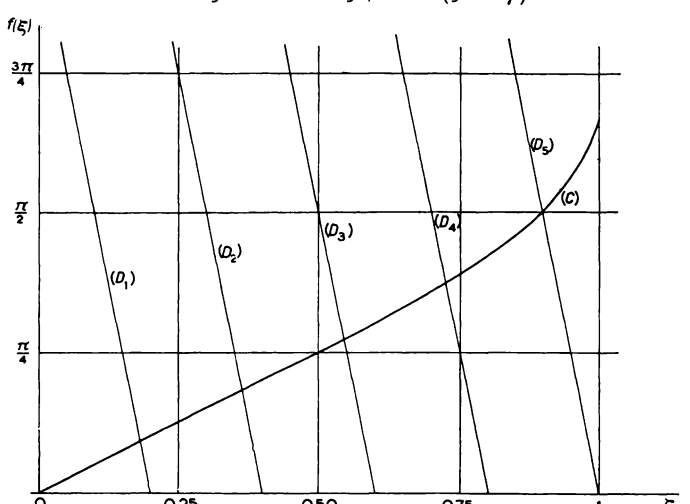


Fig. III.4. Graphical determination of the discrete eigenvalues: $\xi = [(\varepsilon - U_2)/(U_1 - U_2)]^{\frac{1}{2}}$. The eigenvalues are the points of intersection of the curve (C) whose equation is:

$$f(\xi) = \sin^{-1}\xi + \sin^{-1}(\xi\cos\gamma)$$

with each of the straight lines (D_n) with equations: $y(\xi) = n\pi - \xi KL$ (we have chosen: $\gamma = \pi/3$, $KL/\pi = 5$).

This last equation has been solved graphically in Fig. III.4. When ε increases from U_2 to U_1, ξ increases from 0 to 1, the right-hand side increases from 0 to $\pi-\gamma$ following curve (C) (which depends only upon parameter γ), and the left-hand side decreases from $n\pi$ to $n\pi-KL$ along the straight-line segment (D_n). In order that (C) and (D_n) intersect, it is necessary and sufficient that the integer n be sufficiently small:

$$KL \geqslant (n-1)\pi+\gamma.$$

If $KL<\gamma$, there are no eigenvalues; if $\gamma \leqslant KL \leqslant \pi+\gamma$ there is just one eigenvalue ε_1; if $\pi+\gamma \leqslant KL<2\pi+\gamma$, there are two eigenvalues ε_1 and ε_2 $(\varepsilon_1<\varepsilon_2)$; and so on. It is easy to see that the eigenvalues are always arranged in increasing order of n. They form a *discrete* and *finite* sequence, from the ground state eigenvalue ε_1 to a maximum eigenvalue corresponding to the largest integer contained in the number $1+(KL-\gamma)/\pi$.

The quantum number n has a quite precise mathematical meaning. Inspection of eqs. (III.19) shows that the function $\sin(k_2x+\varphi)$ vanishes $(n-1)$ times as x crosses the interval (a, b). But, according to eq. (III.17), the zeros of that function are those of the function ψ. Consequently, the *number of nodes* of the eigenfunction corresponding to the nth eigenvalue ε_n is $n-1$.

To conclude, let us make a comparison with the classical situation as we did for the infinitely deep well. In the present case, there exists a further point of difference, besides the quantization of the energy: since the wave function has values different from zero in regions I and III, there is a non-vanishing probability of finding the particle in regions whose access is forbidden for the corresponding classical particle.

b) $U_1<\varepsilon<U_3$. CONTINUOUS, NON-DEGENERATE SPECTRUM. WAVE REFLECTION

The situation here is analogous to that of case (a) in the problem of the potential step. To each value of ε corresponds one and only one solution bounded everywhere: the one which is exponentially decreasing in region III. In the interval (U_1, U_3) the spectrum of eigenvalues is *continuous* and *non-degenerate*.

We seek a solution of the form

$$\psi=\begin{cases} e^{-ik_1x}+e^{i(k_1x+2\varphi_1)} & x>a \\ 2A\,e^{i\varphi_1}\sin(k_2x+\varphi_2) & a>x>b \\ 2B\,e^{i\varphi_1}e^{\varkappa_1 x} & b>x. \end{cases} \qquad \text{(III.21)}$$

As in the preceding problems, the continuity conditions of the

logarithmic derivative determine the phases φ_1 and φ_2. One finds

$$\varphi_2 = -k_2 b + \tan^{-1}\frac{k_2}{\varkappa_3}, \quad \varphi_1 = -k_1 a - \frac{\pi}{2} + \tan^{-1}\left[\frac{k_1}{k_2}\tan\left(k_2 L + \tan^{-1}\frac{k_2}{\varkappa_3}\right)\right].$$

Then the value of A and B follow from the continuity of the function.

We assume below that $U_3 - \varepsilon \gg \varepsilon - U_2$; hence $k_2 \ll \varkappa_3$ and *a fortiori* $k_1 \ll \varkappa_3$. Everything behaves as if region III were occupied by an infinitely repulsive potential so that $B = 0$. The interesting quantities to be considered here are φ_1 and A^2.

Let us take $a = 0$, $b = -L$, and set

$$\eta = \frac{k_1}{K} = \sqrt{\xi^2 - 1}.$$

After a straightforward calculation, one obtains

$$\varphi_1 = \tan^{-1}\left(\frac{\eta}{\xi}\tan \xi KL\right) - \frac{\pi}{2}$$

$$A^2 = \frac{\eta^2}{\eta^2 + \cos^2 \xi KL}.$$

As the energy increases, the phase φ_1 grows in more or less regular fashion, while the quantity A^2 which measures the relative intensity of the wave in region II oscillates between the extreme values $\eta^2/(1+\eta^2)$ and 1, These oscillations are more pronounced the larger KL and the smaller η. Thus let us assume that

$$KL \gg \pi, \qquad \eta \ll 1.$$

Then A^2, considered as a function of η^2 (i.e. of the energy) exhibits a series of sharp peaks of width $4\eta/KL$ separated from each other by $2\pi/KL$. Figure III.5 illustrates this remarkable behavior of A^2 as well as that of φ_1.

We are faced with a typical wave phenomenon, the *resonance* phenomenon. Over certain restricted energy regions (of width $4\eta/KL$) the intensity of the wave in the interior region is of order unity: the resonance energies are those for which $\varphi_2 = (n + \frac{1}{2})\pi$, and hence where region II contains $(n + \frac{1}{2})$ "half wavelengths". Outside of these resonance regions, the intensity is comparatively very weak.

As in the problem of the potential step, we may compare the motion of a wave packet of the type given by eq. (III.9) with the motion

of a classical particle in the same potential. Coming from $+\infty$ at constant velocity $v_1 = (\hbar K/m)\eta$, the classical particle suffers a sudden acceleration at $x = 0$, crosses the region II at velocity $v_2 = (\hbar K/m)\sqrt{1+\eta^2}$,

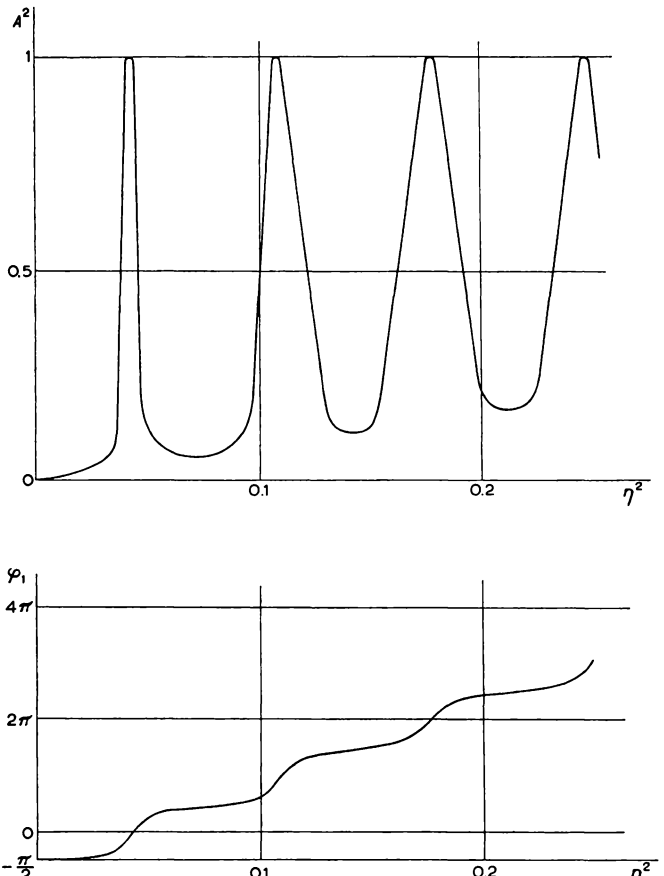


Fig. III.5. Resonances of reflection. Variation of A^2 and of φ_1 [cf. eq. (III.21)] as function of the energy. The curves correspond to $KL = (b-a)\sqrt{U_1 - U_2} = 100$. The energy scale is expressed in units of the variable

$$\eta^2 = (\varepsilon - U_1)/(U_1 - U_2).$$

bounces back at $x = -L$; it starts back in the opposite direction at velocity v_2 in region II, then at velocity v_1 in region I. The time spent in region II is $\tau_{\text{cl.}} = 2L/v_2$. The center of the wave packet carries out

an analogous motion, at least in the region of very large x, where the shape of the wave packet remains practically unchanged so that the motion of a center remains meaningful. Everything behaves as if it had carried out the classical motion except for the following slight difference: the "time spent in region II" is no longer $\tau_{cl.}$ but $\tau = (2/v_1)\, d\varphi_1/dk_1 = (2/v_1 K)\, d\varphi_1/d\eta$. We shall not dwell upon the details of this study which is in all respects similar to the one of § 3. The behavior of the various quantities entering the discussion is summarized in the following table:

Incident energy	φ_1	$\frac{d\varphi_1}{dk_1}$	$\tau/\tau_{cl.}$	A^2
At resonance	$n\pi$	L	$1/\eta$	1
Half-way between two resonances .	$(n + \frac{1}{2})\pi$	$L\eta^2$	η	η^2

Between the resonances, A^2 remains very small, and the time of passage τ in region II is very short compared to $\tau_{cl.}$: the wave packet practically does not penetrate into region II. The wave is almost entirely reflected upon arriving at the point $x = 0$. One meets an identical situation in optics where a sudden and appreciable change of index of refraction almost always entails total reflection. At resonance on the other hand, $A^2 = 1$, the wave penetrates entirely into region II, and remains concentrated there during a relatively long time, in fact much longer than $\tau_{cl.}$. In conformity with condition (III.12), the picture obtained here is applicable only for wave trains whose extent is considerably greater than that of region II itself ($d\varphi_1/dk_1 = L$ at resonance); consequently, the front of the wave packet reaches the point of reflection $x = -L$ long before the wave has finished crossing the potential step at $x = 0$. The effect is typical of waves, namely one of interference between incident and reflected wave in region II.

c) $\varepsilon > U_3$. Continuous and Degenerate Spectrum. Reflection and Transmission of Waves

The situation here is analogous to that of case (*b*) in the problem of the potential step. To each value of ε correspond two linearly independent eigenfunctions: in the interval (U_3, ∞) the eigenvalue spectrum is continuous and degenerate of order 2.

As in the problem of the potential step we build the eigenfunction of the form

$$\chi = \begin{cases} e^{-ik_1 x} + R\, e^{ik_1 x} & x > a \\ P\, e^{-ik_2 x} + Q\, e^{ik_2 x} & a > x > b \\ S\, e^{-ik_3 x} & b > x. \end{cases} \qquad \text{(III.22)}$$

The continuity conditions at points a and b give the values of R, Q, P, and S. Without entering into calculation details, we simply list the results concerning the quantities R and S. We use the following notation and conventions:

$$a = 0, \qquad b = -L, \qquad K = \sqrt{U_1 - U_2},$$
$$\xi = \frac{k_2}{K}, \qquad \eta = \frac{k_1}{K}, \qquad \zeta = \frac{k_3}{K}.$$

One obtains

$$R = \frac{\xi(\eta - \zeta)\cos \xi KL + i(\xi^2 - \eta\zeta)\sin \xi KL}{\xi(\eta + \zeta)\cos \xi KL - i(\xi^2 + \eta\zeta)\sin \xi KL}$$

$$S = e^{-i\zeta KL} \frac{2\eta\xi}{\xi(\eta + \zeta)\cos \xi KL - i(\xi^2 + \eta\xi)\sin \xi KL}$$

With the help of these expressions, one is able to compare the motion of a wave packet formed with waves of the type (III.22) and neighboring energies, with the motion of a classical particle of the same energy in the same potential.

The incident wave packet [made up of waves $\exp(-ik_1 x)$ of region I] moves in region I with constant velocity $v_1 = \hbar k_1/m$ and enters region II; once the collision has taken place, it splits into a reflected wave packet [built of waves $R \exp(ik_1 x)$ of region I] moving with velocity v_1 toward $+\infty$, and a transmitted wave packet [built of waves $S \exp(-ik_3 x)$ of region III] travelling with velocity v_3 toward $-\infty$. In contrast to the classical particle, *the wave is in general only partially transmitted*, and we can define a transmission coefficient

$$T = \frac{k_3}{k_1}|S|^2 = \frac{4\eta\zeta\xi^2}{\xi^2(\eta + \zeta)^2 \cos^2 \xi KL + (\xi^2 + \eta\zeta)^2 \sin^2 \xi KL} \qquad \text{(III.23)}$$

as we did in the case of the potential step.

Here too, we note that the transmission coefficient at a given energy is independent of the direction of travel (η and ζ occur symmetrically in T).

One further verifies the equality

$$|R|^2 + \frac{k_3}{k_1}|S|^2 = 1. \qquad \text{(III.24)}$$

The ratio of reflected to transmitted wave amplitudes varies with the energy, and we can observe *resonance phenomena* of the same kind as those of case (*b*). The latter are particularly sharp when

$$KL \gg \pi \qquad \zeta < \eta \ll 1. \qquad \text{(thus } \xi = 1\text{)}.$$

In that case, inspection of equation (III.23) shows that the transmission coefficient considered as a function of η^2 (i.e. of the energy) remains very small, of the order of $4\eta\zeta$, almost everywhere, but exhibits a series of sharp maxima equal to $4\eta\zeta/(\eta+\zeta)^2$; the width of these peaks is about $4(\eta+\zeta)/KL$. Their positions correspond to the energies for which region II contains an integral number n of "half wavelengths", namely $\xi KL = n\pi$ (their separation is approximately $2\pi/KL$).

One may pursue this study by considering the phases of the amplitudes R and S, and defining a "transit time" of the transmitted wave, or a "reflection time" of the reflected wave; one then compares these times with the time necessary for the classical particle to cross region II. One is led to the following qualitative picture: at resonance, the wave remains concentrated in region II during a time which is very much greater [$(\eta+\zeta)^{-1}$ times greater] than the classical time before splitting up into a transmitted and a reflected wave; off resonance, the wave practically does not penetrate into region II, and is almost totally reflected at the boundary of regions I and II, in near-instantaneous fashion (cf. Problem III.1).

7. Penetration of a Square Potential Barrier. The "Tunnel" Effect

As a last example, we examine the problem of the penetration of a square potential barrier (Fig. III.6).

$$U(x) = \begin{cases} 0 & x > L \qquad \text{(region I)} \\ U_0 (> 0) & 0 < x < L \qquad \text{(region II)} \\ 0 & x < 0. \qquad \text{(region III)} \end{cases}$$

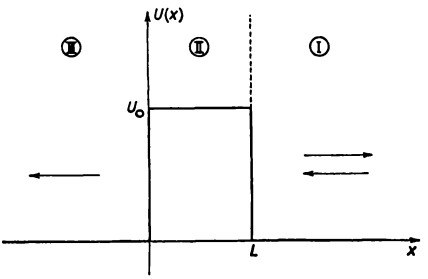


Fig. III.6. Square potential barrier.

In this case, all positive values of ε are doubly degenerate eigenvalues. Two cases occur according to whether ε is larger or smaller than U_0. In both cases we construct the solution representing a wave traveling in the negative direction in region III, i.e. the solution of the form

$$\begin{aligned} &e^{-i\sqrt{\varepsilon}x} + R\, e^{+i\sqrt{\varepsilon}x} \quad \text{for} \quad x > L \\ &S\, e^{-i\sqrt{\varepsilon}x} \qquad\qquad\quad \text{for} \quad x < 0. \end{aligned}$$

Its behavior in region II is

exponential: $A\, e^{\varkappa x} + B\, e^{-\varkappa x}$, if $\varepsilon < U_0$ $\quad (\varkappa = \sqrt{U_0 - \varepsilon})$

sinusoidal: $C\, e^{ikx} + D\, e^{-ikx}$, if $\varepsilon > U_0$ $\quad (k = \sqrt{\varepsilon - U_0})$.

Let us merely give the result of the calculation of the transmission coefficient (Fig. III.7):

$$T = |S|^2 = \begin{cases} \dfrac{4\varepsilon(\varepsilon - U_0)}{4\varepsilon(\varepsilon - U_0) + U_0^2 \sin^2 kL}, & \text{if} \quad \varepsilon > U_0 \\ \dfrac{4\varepsilon(U_0 - \varepsilon)}{4\varepsilon(U_0 - \varepsilon) + U_0^2 \sinh^2 \varkappa L}, & \text{if} \quad \varepsilon < U_0. \end{cases}$$

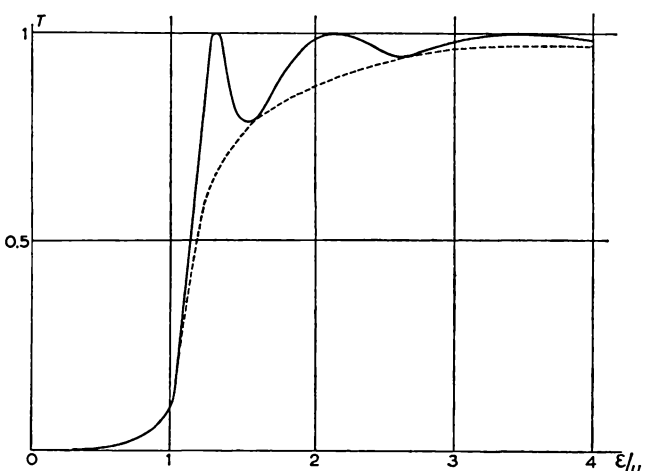


Fig. III.7. Variation of the transmission coefficient through a square potential barrier as a function of the energy (barrier of Fig. III.6). We have used:

$$U_0 L^2 = 40.$$

As in the preceding sections, one can compare the motion of a wave packet of the type (III.9) built of wave functions of that form,

with the motion of a classical particle of the same energy coming from $+\infty$.

The most spectacular difference occurs when $\varepsilon < U_0$. The classical particle rebounds from the barrier without being able to penetrate it. The wave packet splits into a reflected packet, and a transmitted packet whose intensity never vanishes: as ε increases from 0 to U_0, the transmission coefficient increases regularly from 0 to the value $[1 + U_0L^2/4]^{-1}$. This effect is known as the *tunnel effect* and plays an important role in the theory of alpha radioactivity. It is more important the lower and thinner the barrier.

When $\varepsilon > U_0$, the classical particle is slowed down in region II but traverses it nevertheless and continues its journey in region III toward $-\infty$. The wave packet, on the other hand, is almost always partially reflected. Complete transmission ($T = 1$) occurs only for certain values of the energy, namely those for which kL is a multiple of π. As the energy increases, the transmission coefficient oscillates between this maximum value and a minimum value of the order of $4\varepsilon(\varepsilon - U_0)/(2\varepsilon - U_0)^2$. The effect is particularly marked when the barrier is very high or very thick, and when the kinetic energy $\varepsilon - U_0$ in region II is small. Note the similarity with the resonance phenomena discussed in the preceding sections (cf. Problem III.2).

II. GENERAL PROPERTIES OF THE ONE-DIMENSIONAL SCHRÖDINGER EQUATION

8. Property of the Wronskian

Let us return to the equation

$$y'' + [\varepsilon - U(x)]y = 0. \tag{III.25}$$

We propose to show some very general properties of this eigenvalue equation. In all that follows the only restrictions imposed upon the real function $U(x)$ are: to be bounded from below, and to be piecewise continuous over the entire interval $(-\infty, +\infty)$.

A large number of these properties derive directly from an important theorem concerning the Wronskian of two solutions, which we shall henceforth call the *Wronskian Theorem*.

By definition, the Wronskian of two functions y_1, y_2 is

$$W(y_1, y_2) \equiv y_1y_2' - y_2y_1'.$$

It is a bilinear expression in y_1 and y_2, antisymmetric in the exchange of these two functions. If it vanishes at a point of the x axis, the functions y_1 and y_2 have the same logarithmic derivative at that point; if it vanishes over the entire interval $(-\infty, +\infty)$, the two functions are multiples of each other.

WRONSKIAN THEOREM. — *If z_1 and z_2 are solutions of the equations*

$$z_1'' + F_1(x)z_1 = 0 \qquad \text{(III.25')}$$

$$z_2'' + F_2(x)z_2 = 0 \qquad \text{(III.25'')}$$

respectively, in an interval (a, b) where the functions $F_1(x)$ and $F_2(x)$ are piecewise continuous, the overall variation of their Wronskian in this interval is given by the expression

$$W(z_1, z_2)\Big|_a^b = \int_a^b [F_1(x) - F_2(x)]z_1 z_2 \, dx. \qquad \text{(III.26)}$$

To prove this theorem, we multiply eq. (III.25') by z_2, eq. (III.25'') by z_1, and subtract term by term. We obtain

$$[z_2 z_1'' - z_1 z_2''] + (F_1 - F_2)z_1 z_2 = 0.$$

The first term is, except for a sign, the derivative with respect to x of the Wronskian $W(z_1, z_2)$. Upon integrating term by term over the interval (a, b), one obtains relation (III.26), Q.E.D.

This theorem if of particular interest to us when the equations (III.25') and (III.25'') are equations of the type (III.25) with the same potential $U(x)$. One then obtains the three following important corollaries:

Corollary I. — If y_1 and y_2 are solutions of the equation (III.25) corresponding to the values ε_1, ε_2 of the constant ε, respectively, one has for any pair of values a, b of the variable x, located in the interval where these solutions are defined:

$$W(y_1, y_2)\Big|_a^b = (\varepsilon_1 - \varepsilon_2)\int_a^b y_1 y_2 \, dx. \qquad \text{(III.27)}$$

Corollary II. — If y and z are two solutions of eq. (III.25) corresponding to the same value of ε, their Wronskian is independent of x:

$$W(y, z) = \text{constant}.$$

Corollary III. — Let $Y(x;\varepsilon)$ be the solution of eq. (III.25) whose logarithmic derivative (with respect to x) has a fixed value f_a at point a of the x axis, and let $f(x;\varepsilon)$ be its logarithmic derivative at a point x of the axis. When considered as a function of ε, $f(x;\varepsilon)$ is a *monotonic* function of that variable, increasing if $x<a$, decreasing if $x>a$, whose derivative is

$$\frac{\partial f}{\partial \varepsilon} = -\frac{1}{Y^2(x;\varepsilon)} \int_a^x Y^2(\xi;\varepsilon)\, d\xi. \tag{III.28}$$

[Considered as a function of ε, $f(x;\varepsilon)$ has a behavior analogous to the tangent or the cotangent, with a vertical asymptote at each point where $Y(x)$ vanishes.]

Corollaries I and II are direct applications of the Wronskian Theorem. The proof of Corollary III is as follows. ε being fixed, a solution of eq. (III.25) is completely determined if one gives its value and that of its derivative at a given point $x=a$ of the x axis. Let $Y(x;\varepsilon)$ be that particular solution

$$Y(a;\varepsilon)=y_a, \qquad Y'(a;\varepsilon)=y_a'.$$

If one varies ε while maintaining these boundary conditions constant, $Y(x;\varepsilon)$ is a certain continuous function of ε (and of x). To two infinitely close values ε, $\varepsilon+\delta\varepsilon$ correspond two expressions Y, $Y+\delta Y$ lying infinitely close together. Apply Corollary I to the latter in the interval (a, b):

$$W(Y,\ Y+\delta Y)\Big|_a^b = -\delta\varepsilon \int_a^b Y^2\, dx.$$

For $x=a$, $W(Y, Y+\delta Y)=0$ by hypothesis. For all other values of x,

$$W(Y,\ Y+\delta Y)=W(Y,\ \delta Y)=Y\,\delta Y'-Y'\,\delta Y=Y^2\,\delta\left(\frac{Y'}{Y}\right)=Y^2\,\delta f.$$

We have introduced $f=Y'/Y$. The logarithmic derivative f is, like Y, a continuous function of ε and of x. We have

$$-Y^2\,\delta f\Big|_{x=b} = \delta\varepsilon \int_a^b Y^2\, dx.$$

In other words

$$\frac{\partial f}{\partial \varepsilon}\bigg|_{x=b} = -\frac{1}{Y^2(b)} \int_a^b Y^2(x)\, dx. \qquad \text{Q.E.D.}$$

The properties of the solutions of the Schrödinger equation contained in these three corollaries derive their interest from the fact that they are independent of the particular shape of the potential $U(x)$.

9. Asymptotic Behavior of the Solutions

The asymptotic form of the general solution of eq. (III.25) at the boundaries of the interval $(-\infty, +\infty)$ is very different depending on the sign of $\varepsilon - U$ as x approaches these limits. We look for the asymptotic form in the limit $x = +\infty$. Similar conclusions are applicable at the limit $x = -\infty$. Suppose thus that $\varepsilon - U(x)$ keeps the same sign when x exceeds a certain value x_0. Two cases can then occur:

First Case: $\varepsilon > U(x)$ *when* $x > x_0$.

We shall assume — as is always the case in practice — that $U(x)$ tends monotonically toward a finite limit U_+ as $x \to \infty$. Let us put

$$k = \sqrt{\varepsilon - U_+}.$$

We shall show that when $x \to \infty$:

(a) the real solutions of equation (III.25) remain bounded and oscillate indefinitely between two opposite values;

(b) if, furthermore, $U(x)$ tends toward U_+ faster than $1/x$,

$$y \underset{x \to \infty}{\sim} A_+ \sin(kx + \varphi_+) \tag{III.29}$$

where A_+ and φ_+ are two suitable real constants.

For this purpose, one notes that equation (III.25) "tends asymptotically" toward the equation $z'' + k^2 z = 0$ whose general solution is $A \sin(kx + \varphi)$ and depends on two arbitrary constants A and φ. In order to determine the asymptotic form of y, let us introduce (method of variation of constants) the functions $A(x)$ and $\varphi(x)$ defined by

$$y = A \sin(kx + \varphi), \qquad y' = Ak \cos(kx + \varphi). \tag{III.30}$$

Equation (III.25) is equivalent to the two first-order differential equations

$$\frac{A'}{A} = \frac{U - U_+}{2k} \sin 2(kx + \varphi), \qquad \varphi' = -\frac{U - U_+}{k} \sin^2(kx + \varphi).$$

Upon integration they yield

$$A(x) = A(x_0) \exp\left\{ \int_{x_0}^{x} \frac{U(\xi) - U_+}{2k} \sin 2[k\xi + \varphi(\xi)] \, d\xi \right\} \tag{III.31}$$

$$\varphi(x) = \varphi(x_0) - \int_{x_0}^{x} \frac{U(\xi) - U_+}{k} \sin^2 [k\xi + \varphi(\xi)] \, d\xi. \tag{III.32}$$

The integral of the right-hand side of (III.31) certainly converges: $A(x)$ approaches a finite limit A_+ as $x \to \infty$; furthermore, since $\varphi' \to 0$, the function $\sin(kx + \varphi)$ in expression (III.30) for y oscillates with a period which tends asymptotically to $2\pi/k$. This proves the first stated result. If in addition $U(x)$ approaches U_+ more rapidly than $1/x$, the integral of the right-hand side of eq. (III.32) also converges: A and φ both approach finite limits A_+ and φ_+, respectively, whence the asymptotic form (III.29). Q.E.D.

Second Case: $\varepsilon < U(x)$ when $x > x_0$.

The results we shall obtain are independent of the behavior of $U(x)$ at infinity. We merely assume that

$$U(x) - \varepsilon \geqslant M^2 > 0 \quad \text{when} \quad x > x_0.$$

This case corresponds to the exponential solutions in square well potential problems.

We shall show that when $x \to \infty$:

(a) there exists *one* particular solution (defined to within a constant) of equation (III.25) which approaches 0 at least as rapidly as $\exp(-Mx)$;

(b) all other solutions tend toward ∞ at least as rapidly as $\exp(Mx)$.

Since the solutions are defined to within a constant, let us fix this constant by the condition $y(x_0) = 1$, and let us look at the behavior of the solutions $y(x)$ satisfying this normalization condition. Some of them are represented in the graph of Fig. III.8.

If we designate by $Y(x)$ and $Z(x)$ the particular solutions defined by the boundary conditions

$$Y(x_0) = 1, \qquad Y'(x_0) = 0 \qquad \text{and} \qquad Z(x_0) = 0, \qquad Z'(x_0) = 1$$

respectively, the solutions considered above are of the form

$$y(x) = Y(x) + fZ(x). \tag{III.33}$$

The parameter $f = y'(x_0)$ may take on all values between $-\infty$ and $+\infty$.

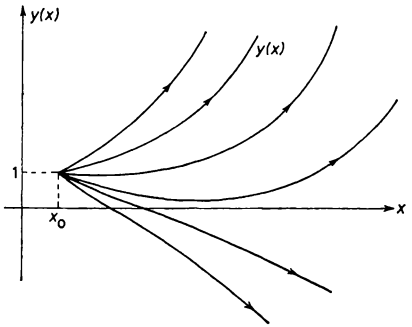


Fig. III.8. Diagram representing several solutions of eq. (III.25) satisfying the normalization condition $y(x_0) = 1$ for the case where $U(x) - \varepsilon \geqslant M^2 > 0$ for $x > x_0$.

$Y(x)$ and $Z(x)$ remain positive over the entire interval (x_0, ∞) and tend toward infinity at least as fast as $\exp(Mx)$. Indeed, like all solutions of eq. (III.25), these functions keep the same sign as their second derivatives. From this one derives by simple inspection of their boundary conditions that these functions necessarily increase indefinitely since they are concave upward. In order to evaluate the rapidity of their growth, note that $Y'' \geqslant M^2 Y$ and $Z'' \geqslant M^2 Z$ and compare these functions with the solutions of the differential equation $u'' - M^2 u = 0$ having the same initial conditions at x_0, namely $\cosh M(x-x_0)$ and $\sinh M(x-x_0)$, respectively. Y and Z remain everywhere superior (or equal) to these comparison functions.

By applying the Wronskian Theorem one has in fact

$$W[Y, \cosh M(x-x_0)] \leqslant 0,$$

therefore

$$Y'/Y \geqslant M \tanh M(x-x_0)$$

and, by integration

$$Y \geqslant \cosh M(x-x_0);$$

one shows similarly that $Z \geqslant \sinh M(x-x_0)$. Note that

$$Y' \geqslant MY \tanh M(x-x_0),$$

thus at infinity $Y' \geqslant MY$; similarly, at infinity $Z' \geqslant MZ$.

On the other hand (Corollary II)

$$Z'Y - Y'Z = 1 \quad \text{for any } x. \tag{III.34}$$

Let us introduce the functions

$$u(x) \equiv \frac{Y}{Z}, \qquad v(x) \equiv \frac{Y'}{Z'}.$$

From the property (III.34) and from the fact that Y and Z are solutions of equation (III.25), one obtains:

$$\begin{aligned} u - v &\equiv \frac{Y}{Z} - \frac{Y'}{Z'} = \frac{1}{ZZ'} \\ u' &= \frac{Y'Z - YZ'}{Z^2} = -\frac{1}{Z^2} \\ v' &= \frac{Y''Z' - Y'Z''}{Z'^2} = \frac{U-\varepsilon}{Z'^2}. \end{aligned} \tag{III.35}$$

In the interval (x_0, ∞), u is a decreasing function, v is an increasing function and their difference vanishes at infinity. They therefore have a common (positive) limit C as $x \to \infty$, and we have

$$v(x) < C < u(x),$$

an inequality which can be rewritten with the help of eq. (III.35)

$$-\frac{1}{ZZ'} < v - C < 0 < u - C < \frac{1}{ZZ'}. \tag{III.36}$$

The particular solution

$$\hat{y}(x) \equiv Y - CZ = [u(x) - C]\, Z(x)$$

and its derivative

$$\hat{y}'(x) \equiv Y' - CZ' = [v(x) - C]\, Z'(x)$$

always satisfy the inequalities

$$-\frac{1}{Z} < \hat{y}' < 0 < \hat{y} < \frac{1}{Z'}.$$

Thus, $\hat{y}$ is positive everywhere, and tends toward 0 at least as fast as $1/Z'$, and *a fortiori* at least as fast as $\exp(-Mx)$. Similarly, $\hat{y}'$, a negative function everywhere, tends toward 0 at least as rapidly as $\exp(-Mx)$. The solution $\hat{y}$ is the solution vanishing at infinity which we were seeking.

Obviously, there exist no others, because if $f \neq -C$, the solution $y(x)$ can also be written

$$y = \hat{y} + (f + C)Z,$$

and its asymptotic behavior is the same as that of Z, to within the non-vanishing factor $f+C$. Q.E.D.

10. Nature of the Eigenvalue Spectrum

Let U_+ and U_- be the respective limits of $U(x)$ when x tends toward $+\infty$ and $-\infty$. (The following conclusions hold when either one of these limits is replaced by $+\infty$.) U_+ and U_- divide the domain of variation of ε into three regions in which the eigenvalue spectra have different properties. To be specific, let us assume that $U_+ < U_-$.

When $\varepsilon > U_-$, $\varepsilon - U(x)$ remains positive at the two ends of the interval $(-\infty, +\infty)$; since any solution of eq. (III.25) which remains bounded as $x \to \pm\infty$, is acceptable as eigenfunction, ε is a degenerate eigenvalue of order 2. The eigenvalue spectrum is *continuous* and *degenerate*. On the other hand, in the two asymptotic regions the eigenfunctions oscillate indefinitely between two finite and opposite limits: they represent *unbound states*.

When $U_- > \varepsilon > U_+$, since $\varepsilon - U(x)$ is negative in the limit $x = -\infty$, only one solution remains bounded (exponentially decreasing) in this asymptotic region. This solution remains bounded and oscillates indefinitely in the other asymptotic region since $\varepsilon - U(x)$ is positive there; it is therefore an acceptable eigenfunction and represents an *unbound state*. The eigenvalue spectrum is *continuous* and *non-degenerate*.

When $U_+ > \varepsilon$, $\varepsilon - U(x)$ is negative in both asymptotic regions. The bounded solution, if it exists, vanishes (exponentially) at the two extremities of the interval and represents a *bound state*. However, it exists only for discrete values of ε. Indeed, let $\hat{y}_-$ be the solution vanishing at the limit $-\infty$, $\hat{y}_+$ the solution vanishing at the limit $+\infty$, and f_- and f_+ their respective logarithmic derivatives at a definite point of the x axis. ε is an eigenvalue if, and only if, $\hat{y}_-$ and $\hat{y}_+$ are equal (to within a constant factor), i.e. if $f_- = f_+$. Considered as functions of ε, f_- is a monotonically decreasing function, f_+ a monotonically increasing function (Corollary III [1])). Therefore, the values of ε for which these two functions are equal, are necessarily

---

[1]) In fact, we are dealing with an extension of Corollary III to the case where $a = +\infty$ or $-\infty$. It is easy to see that this corollary remains valid with a slight change in the definition of the function $Y(x;\varepsilon)$. The latter is the solution of eq. (III.5) which vanishes (and whose derivative vanishes as well) at the point a $(= \pm\infty)$. Note that f_- and f_+ can exhibit vertical asymptotes for some values of ε.

isolated from each other. The spectrum is *discrete* and *non-degenerate*.

The *number of eigenvalues of the discrete spectrum* depends upon the particular function $U(x)$ considered. It may go from 0 to ∞. It is certainly zero if $U(x)$ everywhere exceeds the smaller of its two asymptotic values, U_+. If not, one can show – and we state it here without proof – that it is of the order of magnitude of

$$\frac{1}{\pi}\int \sqrt{U_+ - U(x)}\,\mathrm{d}x = \frac{1}{\pi}\int \frac{\sqrt{2m[V_+ - V(x)]}}{\hbar}\,\mathrm{d}x,$$

this integral being extended over the entire region of the x axis where $U(x) < U_+$. In particular, if this integral diverges, the eigenvalues are infinite in number.

11. Unbound States: Reflection and Transmission of Waves

With the eigenfunctions of the continuous, doubly-degenerate part of the spectrum, one may construct wave packets which are partly transmitted and partly reflected by the potential $U(x)$. To simplify matters we assume that $U(x)$ reaches its two asymptotic limits, U_+ and U_- more rapidly than $1/|x|$; we can then use the asymptotic expressions given by eq. (III.29).

To form the wave packets under consideration, two types of eigenfunctions may be used. The functions of the first type are the functions $u(x)$ whose behavior in the two asymptotic zones is

$$\begin{aligned} u &\underset{x\to-\infty}{\sim} \mathrm{e}^{\mathrm{i}k_- x} + R_u\,\mathrm{e}^{-\mathrm{i}k_- x} \\ &\underset{x\to+\infty}{\sim} S_u\,\mathrm{e}^{\mathrm{i}k_+ x} \end{aligned} \qquad (k_\pm = \sqrt{\varepsilon - U_\pm}).$$

The wave packet one obtains by superposing functions of neighboring energies in a way analogous to that given by eq. (III.9) represents an incident wave $\exp(\mathrm{i}k_- x)$ traveling from $-\infty$ in the positive direction; it enters the potential $U(x)$ and splits up into a reflected wave $R_u \exp(-\mathrm{i}k_- x)$ moving in the opposite sense, and a transmitted wave $S_u \exp(\mathrm{i}k_+ x)$ propagating toward $+\infty$ (Fig. III.9*a*).

The functions of the second type are the functions $v(x)$ whose asymptotic behavior is

$$\begin{aligned} v &\underset{x\to-\infty}{\sim} S_v\,\mathrm{e}^{-\mathrm{i}k_- x} \\ &\underset{x\to+\infty}{\sim} \mathrm{e}^{-\mathrm{i}k_+ x} + R_v\,\mathrm{e}^{\mathrm{i}k_+ x} \end{aligned}$$

and which can represent an analogous wave packet traveling in the opposite sense (Fig. III. 9*b*).

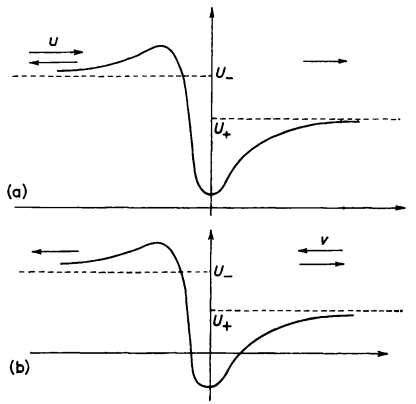


Fig. III.9. Reflection and transmission of waves by a potential:
(*a*) solution of type u: wave coming from negative x;
(*b*) solution of type v: wave coming from positive x.

The functions u and v and their complex conjugates u^* and v^* are solutions of the same Schrödinger equation. The Wronskian of any two such functions is independent of x (Corollary II); in particular it takes on the same value in the two asymptotic regions; equating these two values we obtain a relation between the coefficients R_u, S_u, R_v, and S_v, or their complex conjugates. Six such relations can be formed with the four functions u, v, u^* and v^*. They are very general relations, which hold whatever the shape of the potential function $U(x)$.

We obtain [1])

$$\frac{i}{2}W(u, u^*) = k_-[1-|R_u|^2] = k_+|S_u|^2 \qquad \text{(III.37)}$$

$$\frac{i}{2}W(v, v^*) = -k_-|S_v|^2 = -k_+[1-|R_v|^2] \qquad \text{(III.38)}$$

$$\frac{i}{2}W(u, v) = k_-S_v = k_+S_u \qquad \text{(III.39)}$$

$$\frac{i}{2}W(u, v^*) = -k_-R_uS_v^* = k_+S_uR_v^* \qquad \text{(III.40)}$$

---

[1]) This calculation is quite simple if one makes use of the fact that the Wronskian of two functions is bilinear and antisymmetric, and if one notes that $W(e^{ikx}, e^{ikx}) = W(e^{-ikx}, e^{-ikx}) = 0$ and that $W(e^{-ikx}, e^{+ikx}) = 2ik$.

and two relations deduced from the two last ones by complex conjugation.

The equations (III.37) and (III.38) are called the *relations of conservation of flux*; we have already verified them in special cases [eqs. (III.13) and (III.24)]. The origin of this name comes from the following interpretation of the wave function ψ of an unbound state in the asymptotic region. We shall examine the justification of that interpretation when we shall discuss collision problems (Ch. X). Let $A \exp(ikx) + B \exp(-ikx)$ be the expression of the wave function ψ in one of the asymptotic regions, $-\infty$ for instance. If one builds with this wave function a wave packet analogous to the one given by expression (III.9), it is made up of two terms; the first one formed with $A \exp(ikx)$ is a term of relative intensity $|A|^2$ which is traveling in the direction of increasing x with velocity $\hbar k/m$; the other, formed with $B \exp(-ikx)$, has an intensity $|B|^2$ and travels at the same velocity but in the opposite direction. Sufficiently far in the asymptotic region these two terms do not interfere, the first leaving that region long before the other reaches it. However, instead of forming with this wave function a packet representing the motion of a particle, one can equally well consider — and this is precisely the interpretation mentioned above — that the wave itself represents a statistical ensemble of a very large number of particles. This wave is composed of a beam of density $|A|^2$ in the asymptotic region under consideration, traveling in the direction of increasing x with velocity $\hbar k/m$, and of a beam of density $|B|^2$ moving at the same speed but in the opposite direction. The total flux (counted positively in the direction of increasing x) of particles passing a given point is the difference between the flux $(\hbar k/m)|A|^2$ of particles traveling in the positive sense, and the flux $(\hbar k/m)|B|^2$ of particles traveling in the negative sense. This flux is equal, to within a constant, to the Wronskian $W(\psi, \psi^*)$:

$$\frac{\hbar k}{m}[|A|^2 - |B|^2] = \frac{i}{2}\frac{\hbar k}{m} W(\psi, \psi^*).$$

The equality of the Wronskian $W(\psi, \psi^*)$ at both ends of the interval $(-\infty, +\infty)$ means that the number of particles entering the interaction region per unit time is equal to the number which leave it.

According to this interpretation, one or the other of eqs. (III.37) and (III.38) may just as well be written:

$$\text{incident flux} - \text{reflected flux} = \text{transmitted flux}.$$

Following this same interpretation, the *transmission coefficient* T may be defined as follows:

$$T \equiv \frac{\text{transmitted flux}}{\text{incident flux}}.$$

We have in particular

$$T_u = \frac{k_+}{k_-}|S_u|^2, \qquad T_v = \frac{k_-}{k_+}|S_v|^2.$$

Writing that the absolute values of the two sides of equation (III.39) are equal, one obtains the equality

$$T_u = T_v. \tag{III.41}$$

At a given energy, the transmission coefficient of a wave is independent of the direction of travel. This is the *reciprocity property of the transmission coefficient*, which was already pointed out in special cases in § 3 and § 6. A potential barrier is as difficult to traverse in one direction as in the other.

The equality of the absolute values of the two sides of eq. (III.40), coupled with the conservation relations (III.37) and (III.38), again yields the reciprocity relation (III.41).

From eqs. (III.39) and (III.40) we also obtain relations between the phases of the reflection and transmission amplitudes:

$$\text{phase}\,(S_u) = \text{phase}\,(S_v)$$

$$\text{phase}\left(\frac{R_u}{S_u}\right) = \pi - \text{phase}\left(\frac{R_v}{S_v}\right).$$

These relations are of interest due to the fact that the phases are related to "retardation" effects in the propagation of the wave packets. We have seen several times in Section I (§§ 3 and 9) that the quantity $\hbar\partial(\text{phase})/\partial E$, the product of $\hbar$ and the derivative with respect to the energy of the phase of the reflection or transmission amplitude, a quantity with the dimension of a time, may be interpreted as the "retardation" of the wave during reflection or transmission. This interpretation proves to be quite general.

12. Number of Nodes of Bound States

Let us now consider the eigenfunctions (if they exist) of the discrete spectrum. There is no degeneracy in that case; consequently these functions are certainly real to within a phase.

Let us return to the notation of § 8. We assume the functions y_1 and y_2 to be real, $\varepsilon_2 > \varepsilon_1$, and we apply relation (III.27) by taking as limits of the integration interval two consecutive zeros of y_1. We obtain

$$y_2 y_1' \Big|_a^b = (\varepsilon_2 - \varepsilon_1) \int_a^b y_1 y_2 \, dx.$$

In the interval (a, b), y_1 keeps the same sign. Let us suppose for instance, that $y_1 > 0$. In that case $y_1'(a) > 0$, $y_1'(b) < 0$. Hence y_2 certainly changes sign in the interval (a, b); if not, the right-hand side of the equation would have the same sign as y_2, the left-hand side would have the opposite sign. Hence y_2 certainly has at least one zero within the interval (a, b). Between two nodes of y_1, there is always at least one node of y_2.

Assume now that y_1 and y_2 are eigenfunctions of the discrete spectrum. They both vanish ("exponentially") at the two boundaries of the interval $(-\infty, +\infty)$. The n_1 nodes of y_1 subdivide that interval into $(n_1 + 1)$ partial intervals, in each of which y_2 has at least one zero: thus, the function y_2 has at least $(n_1 + 1)$ nodes. Consequently an eigenfunction has the more nodes the higher its eigenvalue.

By repeating the argument used in § 10 to build up the successive eigenfunctions, and following the increase of the number of nodes of the functions $\hat{y}_-$ and $\hat{y}_+$ as the energy ε increases continuously, one obtains the more precise statement which follows (Problems III.4 and III.5):

If one arranges the eigenstates in the order of increasing energies $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n, \ldots$, *the eigenfunctions likewise fall in the order of increasing number of nodes; the nth eigenfunction has* $(n-1)$ *nodes between each of which the following eigenfunctions all have at least one node.*

13. Orthogonality Relations

One obtains another very important consequence of the Wronskian Theorem (Corollary I) by letting the limits of integration a and b in eq. (III.27) go to $-\infty$ and $+\infty$, respectively.

Assume that y_1 and y_2 are two eigenfunctions belonging to two distinct eigenvalues of the *discrete spectrum*. They both vanish at infinity, and so does their Wronskian; and since $\varepsilon_2 - \varepsilon_1 \neq 0$,

$$\int_{-\infty}^{+\infty} y_1 y_2 \, dx = 0. \tag{III.42}$$

When the integral of the product $y_1 y_2$ of two (real) functions extended over all space is zero, these two functions are said to be orthogonal. More generally, two complex functions y_1 and y_2 are said to be *orthogonal* if

$$\int_{-\infty}^{+\infty} y_1^* y_2 \, dx = 0.$$

Thus, the eigenfunctions of the discrete spectrum are orthogonal.

Clearly, this result remains valid if only one of the two eigenfunctions belongs to the discrete spectrum.

The limiting procedure is more delicate when both eigenfunctions belong to the continuous spectrum. The Wronskian $W(y_1, y_2)$ exhibits infinite oscillations at least at one of the limits of integration; the integral $\int y_1 y_2 \, dx$ therefore possesses the same property. However, if one replaces at least one of the eigenfunctions in the integral, say y_2, by a wave packet made up of the eigenfunctions of a small domain of energy $\delta\varepsilon$ surrounding ε_2, the orthogonality relation is verified in the limit where $\delta\varepsilon \ll |\varepsilon_1 - \varepsilon_2|$. Indeed, let us write y_2 in the form $y(x; \varepsilon)$ in order to recall that it is a function of the energy ε [1]). Consider also the wave packet

$$Y_2(x; \delta\varepsilon) = \frac{1}{\sqrt{\delta\varepsilon}} \int_{\varepsilon_2}^{\varepsilon_2 + \delta\varepsilon} y(x; \varepsilon) \, d\varepsilon. \tag{III.43}$$

As the Wronskian $W(y_1, y_2)$ depends linearly upon the function y_2, one obtains upon integrating equation (III.27) term by term

$$W(Y_2, y_1)\Big|_a^b = (\varepsilon_2 - \varepsilon_1) \int_a^b y_1 Y_2 \, dx + \int_a^b y_1 \left[\frac{1}{\sqrt{\delta\varepsilon}} \int_{\varepsilon_2}^{\varepsilon_2 + \delta\varepsilon} (\varepsilon - \varepsilon_2) \, y(x; \varepsilon) \, d\varepsilon \right] dx.$$

The interest in this procedure lies in the fact that Y_2 goes to zero (as $1/x$) in the asymptotic regions where y_2 has an oscillatory behavior. When one let a and b go to $-\infty$ and $+\infty$, respectively, the left-hand side vanishes; thus the two converging integrals on the right-hand

---

[1]) Knowing ε is not sufficient to define the solution $y(x; \varepsilon)$ since the latter further depends upon one or two arbitrary constants, according to whether the eigenvalue is simple or degenerate. One resolves this arbitrariness by a suitable condition on the asymptotic form of $y(x; \varepsilon)$ at one of the limits, $-\infty$ or $+\infty$, of the range of integration.

side add up to zero. In the limit where $\delta\varepsilon \ll |\varepsilon_2 - \varepsilon_1|$, the second integral becomes negligible. We can therefore write

$$\lim_{\delta\varepsilon \to 0} \int_{-\infty}^{+\infty} y_1(x)\, Y_2(x;\, \delta\varepsilon)\, \mathrm{d}x = 0. \qquad \text{(III.42')}$$

One calls "eigendifferential" of the function $y_2(x)$ the wave packet $Y_2(x;\, \delta\varepsilon)$ defined by eq. (III.43), in which $\delta\varepsilon$ is a very small quantity which will eventually be made to vanish at the end of the calculations.

In conclusion, *two eigenfunctions belonging to two distinct eigenvalues are orthogonal*, with the convention that at least one of the eigenfunctions must be replaced by its eigendifferential in the orthogonality relation [eq. (III.42')] when the two eigenfunctions belong to the continuous spectrum.

We have mentioned the concept of eigendifferential in a rather summary fashion. It is never used in practice. We shall describe later a rather elegant mathematical device which will enable us to formulate the orthogonality relations in a quite general way, without having recourse to that concept.

14. Remark on Parity

Let us return to the concept of parity, encountered for the first time in connection with the infinitely deep square well. The property is entirely general.

If the potential is even, that is to say if

$$U(x) = U(-x),$$

the Schrödinger Hamiltonian does not change when we replace x by $-x$: *it is invariant under reflection through the origin*. Consequently, if $\psi(x)$ is an eigenfunction of the eigenvalue E,

$$H\psi(x) = E\psi(x),$$

this equation continues to hold true if one changes x to $-x$, hence

$$H\psi(-x) = E\psi(-x).$$

Moreover, the even function $\psi(x) + \psi(-x)$ and the odd function $\psi(x) - \psi(-x)$ are also eigenfunctions for the same eigenvalue E; at least one of these two functions certainly does not vanish identically. Two cases may then arise:

1. The eigenvalue E is not degenerate. The four functions above are multiples of each other: $\psi(x)$ is a multiple of that one of the two functions $\psi(x)+\psi(-x)$ or $\psi(x)-\psi(-x)$ which is not identically equal to zero (the other one is necessarily identically zero). In other words, the eigenfunctions of the non-degenerate portion of the spectrum have a well-defined parity: some are even, others are odd. In fact, an even function necessarily has an even number of nodes, an odd function has an odd number of nodes. Thus, if the eigenfunctions are arranged in the order of increasing eigenvalues of the energy, these functions are alternately even and odd, the eigenfunction of the ground state is always even. The results of § 5 agree with these predictions quite well.

2. The eigenvalue E is degenerate. In that case all functions can be put into the form $\lambda\psi+\mu\varphi$ where ψ and φ are two linearly independent eigenfunctions. Let us assume that at least one of these two functions, ψ say, has no well-defined parity: under these conditions, neither of the functions $\psi_+=\psi(x)+\psi(-x)$ and $\psi_-=\psi(x)-\psi(-x)$ is identically zero. These two functions, being of opposite parity, are of necessity linearly independent; as noted above, they are eigenfunctions of the same eigenvalue E. One can thus express ψ, φ, and consequently $\lambda\psi+\mu\varphi$ in the form of linear combinations of these two functions. Therefore, the eigenfunctions of a degenerate eigenvalue may be written as linear combination of two functions, each having a well-defined parity.

One readily sees by inspection that the eigenvalues of the continuous spectrum are all doubly degenerate, and that to each of them corresponds an even eigenfunction (whose derivative vanishes at the origin and an odd function (which vanishes at the origin).

It often happens in Quantum Mechanics that the Hamiltonian of the system under study is invariant under some specific transformations. From that invariance property there result certain symmetry properties characteristic of the eigensolutions of the Schrödinger equation. Parity represents a very simple example of this situation.

EXERCISES AND PROBLEMS

1. In the square well problem of § 6, calculate the constants P, Q, R, S occurring in the expression (III.22) of the solution χ as a function of the parameters of the well; verify expression (III.23) giving the transmission coefficient,

and the conservation relation (III.24). Assuming that $KL \gg \pi$ and $\zeta < \eta \ll 1$, define and calculate the "transit time" of the transmitted wave, and the "reflection time" of the reflected wave. Show the existence of resonances and compare the motion of the transmitted wave to that of the corresponding classical particle.

2. Calculate the transmission coefficient of the square barrier defined in § 7. Calculate the "transit time" of the transmitted wave and compare the motion of this wave to that of the classical particle.

3. Investigate the motion of a particle in a square potential consisting of an infinite barrier for $x < 0$, and in the region of positive x

$$\begin{cases} V(x) = V_{\mathrm{I}}, & 0 < x < a \\ V(x) = V_{\mathrm{II}}, & a < x < b \\ V(x) = 0, & x > b. \end{cases}$$

One assumes that $V_{\mathrm{I}} < 0 < V_{\mathrm{II}}$. Compare the motion of a wave packet undergoing reflection at $x = 0$ with that of the coresponding classical particle. Investigate the "reflection delay" when the incident energy E is less than V_{II}.

Show the existence of resonances and discuss the connection between the width of these resonances and the "reflection delay" when $E \ll V_{\mathrm{II}}$ and $(b-a)\sqrt{2mV_{\mathrm{II}}} \gg \hbar$.

4. The energy spectrum of a particle in one dimension in an arbitrary potential may occasionally include a discrete portion. Show that if one arranges the eigenstates of the discrete spectrum in the order of increasing energy, the eigenfunctions are found to lie also in the order of increasing number of nodes, the nth eigenfunction having $(n-1)$ nodes, between any two of which the following eigenfunctions all have at least one node.

5. A particle in one dimension in the interval $(0, \infty)$ lies in a potential $V(x)$ tending asymptotically to 0, and bounded at $x = 0$ by an infinitely repulsive barrier. Show that the number of bound states is equal to the number of nodes of the solution of the Schrödinger equation which vanishes at the origin, corresponding to an infinitely small negative energy.

CHAPTER IV

STATISTICAL INTERPRETATION OF THE WAVE-CORPUSCLE DUALITY AND THE UNCERTAINTY RELATIONS

1. Introduction

If one attempts to locate the underlying causes of discrepancy between the experimental facts concerning microscopic phenomena, and the predictions of classical theory, one arrives at two conclusions of entirely general applicability.

The first is what we have called the *atomism of action*, responsible for the appearance of discontinuity in microscopic physics: the fact that the change of action of a physical system, and the exchanges of action between physical systems can occur only by *discrete* and *indivisible quanta*. We have discussed this new concept at length in the first chapter, and we have seen how the description of phenomena within the framework of Classical Theory can be made only in the limit where the quantum of action $\hbar$ can be treated as negligibly small.

The second conclusion is the wave-corpuscle duality, that very general property of microscopic objects of appearing under one of two contradictory aspects, as the case may be: that of waves or that of corpuscles. This character of duality is in fact very intimately connected with the atomism of action, as is illustrated by the appearance of the constant $\hbar$ in the general correspondence relations (II.5) between waves and corpuscles. We have seen how the very general character of this duality property was recognized only rather late, and what a decisive role the recognition of this fact had played in the development of the Quantum Theory of natural systems.

A systematic examination of the actual course of diffraction experiments (cf. Ch. I, §§ 5 and 6, Ch. II, §§ 7 and 8) lends support to a very simple interpretation of the wave-corpuscle duality of microscopic objects, namely that there exists a statistical bond between their wave aspect and their corpuscular aspect. The intensity of the wave in a region of space gives the probability of finding the particle in this region. It is upon this statistical interpretation that we shall elaborate in this chapter, as well as check its internal consistency and compatibility with the experimental facts.

In Section I, we develop the statistical interpretation of the wave function of Wave Mechanics for systems of matter. The discussions and results of that section will lead us in Section II to formulate a very general consequence of the statistical interpretation of the wave-corpuscle duality: the uncertainty relations of Heisenberg. In Section III, we show how these uncertainty relations can in no way be invalidated by experiment, no matter how they might appear; we must only take into account the fact that the measuring instruments are themselves quantum objects obeying the same relations. Consequently, when a system is subjected to a given measurement, the perturbation caused in its evolution by the intervention of the measuring device can never be made arbitrarily small or perfectly controllable.

In fact, at the quantum level of accuracy, it is impossible to separate the object to be measured from the measuring instrument. Now, when one describes a measuring operation in *ordinary language*, one implicitly assumes a clear-cut separation between the object to be measured and everything that might serve to carry out that measurement. The intervention of the measuring instrument then appears as an uncontrollable perturbation, which cannot be made to vanish because of the atomism of action. The existence of such an uncontrollable disturbance in fact sets a limit to that necessary distinction between subject and object, and entails a revision of the classical ideas concerning the description of phenomena. This question is treated in Section IV of this chapter.

I. STATISTICAL INTERPRETATION OF THE WAVE FUNCTIONS OF WAVE MECHANICS

2. Probabilities of the Results of Measurement of the Position and the Momentum of a Particle

Let us first treat the case of a quantized system consisting of a single particle. Let us designate its wave function by $\Psi(\boldsymbol{r}, t)$. It obeys the Schrödinger equation and is fully determined at each instant if one knows its value $\Psi(\boldsymbol{r}, t_0)$ at the initial time t_0. For the time being we analyze the situation at a given time t and simply denote by $\Psi(\boldsymbol{r})$ the wave function of the particle at time t.

The dynamical state of a classical particle is defined at every instant

upon specifying precisely its position $\boldsymbol{r}(x, y, z)$ and its momentum $\boldsymbol{p}(p_x, p_y, p_z)$. On the other hand, since the wave function $\Psi(\boldsymbol{r})$ has a certain spatial extension one cannot attribute to a quantum particle a precise position; one can only define the probability of finding the particle in a given region of space *when one carries out a measurement of position*. Denote by $P(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r}$ the probability of finding the particle in the volume element $(\boldsymbol{r}, \boldsymbol{r}+\mathrm{d}\boldsymbol{r})$. The probability $P(V)$ of finding it in a finite volume V is obtained by integrating the "probability density" $P(\boldsymbol{r})$ over that volume: $P(V)=\int_V P(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r}$. Similarly, one cannot in general attribute a precise momentum to a quantum particle. If the associated wave is a plane wave $\exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r})$, it actually represents a particle with momentum $\boldsymbol{p}=\hbar\boldsymbol{k}$, according to de Broglie's law of correspondence; however, Ψ is in general a *superposition* of plane waves of variable wave vector $\boldsymbol{k}$. One can only define the probability of finding the momentum in a given region of momentum space *when one carries out a measurement of momentum*; we shall denote by $\Pi(\boldsymbol{p})\,\mathrm{d}\boldsymbol{p}$ the probability of finding the momentum of the particle in the interval $(\boldsymbol{p}, \boldsymbol{p}+\mathrm{d}\boldsymbol{p})$. The probability $\Pi(D)$ of finding the momentum in the finite region D of momentum space derives from it by integration $\Pi(D)=\int_D \Pi(\boldsymbol{p})\,\mathrm{d}\boldsymbol{p}$. The probability densities $P(\boldsymbol{r})$ and $\Pi(\boldsymbol{p})$ are necessarily positive quantities, satisfying the conditions

$$\int P(\boldsymbol{r})\,\mathrm{d}\boldsymbol{r}=1, \qquad \int \Pi(\boldsymbol{p})\,\mathrm{d}\boldsymbol{p}=1, \tag{IV.1}$$

with the integrals of the left-hand side to be taken over the whole configuration space and the whole momentum space, respectively.

The distributions $P(\boldsymbol{r})$ and $\Pi(\boldsymbol{p})$ are completely determined once the wave function $\Psi(\boldsymbol{r})$ is known. We define $P(\boldsymbol{r})$ by the relation

$$P(\boldsymbol{r})=\Psi^*(\boldsymbol{r})\,\Psi(\boldsymbol{r})=|\Psi(\boldsymbol{r})|^2. \tag{IV.2}$$

This definition agrees well with the idea that the larger the probability of presence of a particle the more intense the wave at that point. Relation (IV.1) leads to the *normalization condition*

$$N \equiv \int |\Psi(\boldsymbol{r})|^2\,\mathrm{d}\boldsymbol{r}=1. \tag{IV.3}$$

Thus, the wave function $\Psi(\boldsymbol{r})$ must be square-integrable; furthermore, its normalization must remain constant in the course of time. We shall prove in § 3 that this second condition of consistency of the statistical interpretation is actually fulfilled.

To define $\Pi(\boldsymbol{p})$, let us consider the process of momentum measurement of the particle associated with the wave Ψ. It looks analogous to that of the spectral analysis of a composite light wave. The analogy is particularly striking if the momentum measurement is carried out with the aid of a diffraction arrangement, but the present discussion is independent of the particular nature of the measuring device used. We introduce the Fourier transform $\Phi(\boldsymbol{p})$ of the wave function defined as follows [1])

$$\Phi(\boldsymbol{p}) = (2\pi\hbar)^{-3/2} \int \Psi(\boldsymbol{r})\, \mathrm{e}^{-\mathrm{i}(\boldsymbol{p}\cdot\boldsymbol{r})/\hbar}\, \mathrm{d}\boldsymbol{r} \tag{IV.4}$$

$$\Psi(\boldsymbol{r}) = (2\pi\hbar)^{-3/2} \int \Phi(\boldsymbol{p})\, \mathrm{e}^{+\mathrm{i}(\boldsymbol{p}\cdot\boldsymbol{r})/\hbar}\, \mathrm{d}\boldsymbol{p}. \tag{IV.5}$$

According to eq. (IV.5), $\Psi(\boldsymbol{r})$ may be viewed as a linear combination of elementary waves $\exp(\mathrm{i}\boldsymbol{p}\cdot\boldsymbol{r}/\hbar)$ of well-defined momentum $\boldsymbol{p}$, each elementary wave having a coefficient $(2\pi\hbar)^{-3/2}\Phi(\boldsymbol{p})$. If there was just one single term $\exp(\mathrm{i}\boldsymbol{p}_0\cdot\boldsymbol{r}/\hbar)$, the result of the measurement would with certainty be $\boldsymbol{p}_0$. If $\Phi(\boldsymbol{p})$ has appreciable values only in a small region surrounding $\boldsymbol{p}_0$ as is the case for the wave packets studied in Ch. II, experiment shows that the value found is almost certainly a value lying in the neighborhood of $\boldsymbol{p}_0$. More generally speaking, the probability $\Pi(\boldsymbol{p})\,\mathrm{d}\boldsymbol{p}$ of finding a value of the momentum in the volume element $(\boldsymbol{p}, \boldsymbol{p}+\mathrm{d}\boldsymbol{p})$ is large when $|\Phi(\boldsymbol{p})|$ is large. We are thus led to the definition

$$\Pi(\boldsymbol{p}) = \Phi^*(\boldsymbol{p})\Phi(\boldsymbol{p}) = |\Phi(\boldsymbol{p})|^2. \tag{IV.6}$$

Since the scalar product is conserved by Fourier transformation (Theorem IV, § A.16),

$$\int |\Phi(\boldsymbol{p})|^2\, \mathrm{d}\boldsymbol{p} = \int |\Psi(\boldsymbol{r})|^2\, \mathrm{d}\boldsymbol{r}$$

and the normalization condition (IV.1) is automatically satisfied if the function $\Psi(\boldsymbol{r})$ is normalized to unity.

The Fourier transformation establishes a one-to-one correspondence between the square-integrable functions $\Psi(\boldsymbol{r})$ and $\Phi(\boldsymbol{p})$. The knowledge of $\Phi(\boldsymbol{p})$ is sufficient, as is a knowledge of $\Psi(\boldsymbol{r})$ to define the dynamical state of the particle; hence one calls $\Phi(\boldsymbol{p})$ the *wave function in mo-*

[1]) If the integrals of the right-hand sides of equations (IV.4) and (IV.5) do not converge, the latter must be modified in conformity with the prescriptions of Theorem I of Appendix A (§ 16) (recall that the function $\Psi(\boldsymbol{r})$ is square-integrable). The results below remain valid with this slight modification. In the following, these points of mathematical rigor will be omitted.

mentum space, a terminology which is justified inasmuch as the functions Ψ and Φ play entirely symmetrical roles in the definitions (IV.2) and (IV.6). The functions Ψ and Φ are also said to be equivalent *representations* of the same dynamical state.

The physical meaning of the quantities $P(\boldsymbol{r})$ and $\Pi(\boldsymbol{p})$ should be clearly understood. The particle associated with the wave generally possesses neither a precise position, nor a precise momentum. When carrying out a measurement on either one of these dynamical variables on an individual system represented by Ψ, no definite prediction can be made about the result. The predictions defined here apply to a very large number $\mathcal{N}$ of equivalent systems independent of each other, each system being represented by the same wave function Ψ. If one carries out a position measurement on each of them, $P(\boldsymbol{r})$ gives the distribution of the $\mathcal{N}$ results of measurement in the limit where the number $\mathcal{N}$ of members of this statistical ensemble approaches infinity. Similarly, if one measures the momentum, $\Pi(\boldsymbol{p})$ gives the distribution of the results of measurement of the momentum.

The above definitions of $P(\boldsymbol{r})$ and $\Pi(\boldsymbol{p})$ have been based on plausibility and self-consistency arguments. It is not obvious that expressions (IV.2) and (IV.6) are the only ones which follow from basing oneself on arguments of this type. In fact, the quantities $P(\boldsymbol{r})$ and $\Pi(\boldsymbol{p})$ may in principle be directly compared with the experimental results; the justification of definitions (IV.2) and (IV.6) definitely rests on the success of this comparison.

3. Conservation in Time of the Norm

In order that the definition of the probabilities of the preceding section be consistent, the norm N of the wave function must remain constant in time. Now the functions Ψ and Ψ^* satisfy respectively the Schrödinger equation (II.33) and its complex conjugate equation, namely

$$\mathrm{i}\hbar \frac{\partial}{\partial t} \Psi = H\Psi,$$

$$\mathrm{i}\hbar \frac{\partial}{\partial t} \Psi^* = -(H\Psi)^*.$$

Hence

$$\frac{\partial}{\partial t} |\Psi|^2 = \Psi^* \left(\frac{\partial}{\partial t} \Psi\right) + \left(\frac{\partial}{\partial t} \Psi^*\right) \Psi = \frac{1}{\mathrm{i}\hbar} [\Psi^*(H\Psi) - (H\Psi)^* \Psi]. \quad \text{(IV.7)}$$

Upon integrating the two sides of eq. (IV.7) over all of configuration space, we obtain

$$\frac{dN}{dt} = \frac{1}{i\hbar} \int [\Psi^*(H\Psi) - (H\Psi)^* \Psi] \, d\mathbf{r}.$$

In order that the norm remain constant in time, it is necessary and sufficient that

$$\int \Psi^*(H\Psi) \, d\mathbf{r} = \int (H\Psi)^* \Psi \, d\mathbf{r}. \tag{IV.8}$$

This property must hold no matter what the dynamical state of the particle, hence for any square-integrable function Ψ in configuration space.

In the language of operators, one says that an operator such as H is *Hermitean* if it satisfies the property (IV.8) for all functions Ψ of function space in which this operator was defined. The main properties of Hermitean operators will be studied in Chapter V.

The Schrödinger Hamiltonian actually possesses this hermiticity property. Let us verify it in the simple case of a particle in a scalar potential (the case of a charged particle in an electromagnetic field is the subject of Problem IV.1):

$$H \equiv -\frac{\hbar^2}{2m} \triangle + V(\mathbf{r}).$$

Since $V(\mathbf{r})$ is a real quantity, eq. (IV.8) can in this case be written

$$\int [\Psi^*(\triangle \Psi) - (\triangle \Psi)^* \Psi] \, d\mathbf{r} = 0.$$

If the integral of the left-hand side were extended over a certain volume bounded by a surface S, it would be, according to Green's theorem, equal to the surface integral

$$\int_S \left(\Psi^* \frac{d\Psi}{dn} - \frac{d\Psi^*}{dn} \Psi \right) dS,$$

where d/dn is the external normal derivative. One obtains the integral of the left-hand side by letting the volume of integration extend over all of space. In that limit all elements of the surface S are removed to infinity. However, since the wave Ψ represents the dynamical state of a physical system, it is necessarily square-integrable; consequently the surface integral goes to zero. Q.E.D.

Thus, it is sufficient that the normalization condition (IV.3) be fulfilled at the initial instant for it to remain satisfied at all later times. Since the Schrödinger equation is homogeneous, its solutions are defined to within a constant. The normalization constant at the initial time fixes the absolute value of that constant; its phase remains arbitrary.

4. Concept of Current

The property of conservation of the norm has a simple interpretation if one introduces the notion of current. The right-hand side of eq. (IV.7) can always be put in the form of a divergence of a suitably defined vector, the vector probability current density or *vector current*. Let us restrict ourselves here to the case of a particle in a scalar potential (cf. Problem IV.1). We define the current $\boldsymbol{J}(\boldsymbol{r}, t)$ at point $\boldsymbol{r}$ and at time t by

$$\boldsymbol{J}(\boldsymbol{r}, t) = \mathrm{Re}\left[\Psi^* \frac{\hbar}{\mathrm{i}m} \nabla\Psi\right]. \tag{IV.9}$$

It is easily verified that

$$\mathrm{div}\,\boldsymbol{J} = \frac{\mathrm{i}}{\hbar}\left[\Psi^*(H\Psi) - (H\Psi)^*\,\Psi\right]; \tag{IV.10}$$

thus we can write eq. (IV.7) in the form

$$\frac{\partial}{\partial t} P + \mathrm{div}\,\boldsymbol{J} = 0. \tag{IV.11}$$

Relation (IV.11) commonly occurs in hydrodynamics. It is the conservation law for a fluid of density P and current $\boldsymbol{J}$ in a medium without source or sink. We are thus led to liken the motion of the quantum particle to that of a classical fluid [1]). The mass of fluid contained in a given volume $\mathscr{V}$ is equal to the integral of the density P extended over that volume. From eq. (IV.11) comes the well-known result that the fluid contained in $\mathscr{V}$ is equal to

$$-\int_{\mathscr{V}} \mathrm{div}\,\boldsymbol{J}\,\mathrm{d}\boldsymbol{r} = -\int_{S} \boldsymbol{J}\cdot\mathrm{d}\boldsymbol{S},$$

---

[1]) Of course the analogy between this probability fluid and a classical fluid should not be pushed too far. All pictures based on this analogy contain no more than the property (IV.11) (cf. Ch. VI).

i.e. to the flux of the vector current crossing the surface S bounding that volume. The total mass remains constant (conservation of the norm) if the flux across S goes to zero in the limit where $\mathscr{V}$ extends over all space.

The definition of $\boldsymbol{J}$ inherently has some arbitrariness: the property (IV.11) remains valid when adding to vector $\boldsymbol{J}$ a vector with vanishing divergence. However, the definition (IV.9) has the advantage of simplicity. Besides, it can be deduced by correspondence from the classical definition of current. Indeed, according to the correspondence principle, the operator $(\hbar/im)\boldsymbol{V}$ represents the quantity $\boldsymbol{p}/m$, i.e. the velocity of the particle; the quantity $\boldsymbol{J}$ corresponds to the product of the velocity and the density, i.e. to the current. In particular if Ψ is a plane wave $A \exp[(i/\hbar)(\boldsymbol{p}\cdot\boldsymbol{r} - Et)]$, $\boldsymbol{J}(\boldsymbol{r}, t) = A^2(\boldsymbol{p}/m)$ is actually equal to the product of the (probability) density and the velocity.

Property (IV.11) is much more general than the conservation of the norm. If the function Ψ is a stationary solution of the Schrödinger equation

$$\Psi(\boldsymbol{r}, t) = \psi(\boldsymbol{r})\, e^{-iEt/\hbar},$$

the conservation of the norm is either obvious, or meaningless. It is obvious if the state is bound; it is meaningless if the state is unbound since the function ψ is not square-integrable in this case. In both cases, however, eq. (IV.11) remains valid, and as the density $|\Psi|^2$ is time-independent, it may be written

$$\operatorname{div} \boldsymbol{J} = 0. \tag{IV.12}$$

This property of the eigenfunction ψ is of interest because it does not depend upon the specific form of the potential term in the Hamiltonian [1]).

5. Mean Values of Functions of $\boldsymbol{r}$ or of $\boldsymbol{p}$

Having checked the consistency of our definitions for the probability densities P and Π, we may use them to calculate the mean values of functions of $\boldsymbol{r}$ or of $\boldsymbol{p}$.

Knowing the distribution $P(\boldsymbol{r})$ of the results of position measurements at a given time, one can define the mean value assumed at

[1]) In three-dimensional problems its role is analogous to the property of conservation of the Wronskian $W(y^*, y)$ in one-dimensional problems (cf. Ch. III, § 11).

that time by some arbitrary function $F(\boldsymbol{r}) \equiv F(x, y, z)$ of the coordinates of the particle. The physical meaning to be attributed to this mean value is the same as we have stated in the definition of $P(\boldsymbol{r})$: it is the average value of the measurements of $F(\boldsymbol{r})$ carried out on a very large number $\mathcal{N}$ of equivalent systems, independent of each other and represented by the same wave function Ψ.

We shall designate this quantity by the notation $\langle F(\boldsymbol{r}) \rangle$. It is evident that

$$\langle F(\boldsymbol{r}) \rangle = \int P(\boldsymbol{r})\, F(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r}.$$

Similarly, we obtain for the average value of a function of the momentum $G(\boldsymbol{p}) \equiv G(p_x, p_y, p_z)$:

$$\langle G(\boldsymbol{p}) \rangle = \int \Pi(\boldsymbol{p}) G(\boldsymbol{p})\, \mathrm{d}\boldsymbol{p}.$$

Using the definitions of probability densities of § 2, we obtain the expressions (provided that the integrals converge)

$$\langle F(\boldsymbol{r}) \rangle = \int \Psi^*(\boldsymbol{r})\, F(\boldsymbol{r})\, \Psi(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r} \tag{IV.13}$$

$$\langle G(\boldsymbol{p}) \rangle = \int \Phi^*(\boldsymbol{p}) G(\boldsymbol{p})\, \Phi(\boldsymbol{p})\, \mathrm{d}\boldsymbol{p}. \tag{IV.14}$$

Thus the average value of the coordinate x of the particle is

$$\langle x \rangle = \int \Psi^*(\boldsymbol{r})\, x\, \Psi(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r} \tag{IV.15}$$

and that of the component p_x of the momentum

$$\langle p_x \rangle = \int \Phi^*(\boldsymbol{p})\, p_x\, \Phi(\boldsymbol{p})\, \mathrm{d}\boldsymbol{p}. \tag{IV.16}$$

We shall obtain another form of expression (IV.16) by using the properties of Fourier transforms stated in appendix A. If the function $p_x \Phi(\boldsymbol{p})$ is square-integrable – and we shall always assume this to be so – its Fourier transform is $(\hbar/\mathrm{i})(\partial/\partial x)\Psi(\boldsymbol{r})$ (Theorem A.III). Applying the conservation property of the scalar product (Theorem A.IV) to the functions Φ and $p_x \Phi$, we have

$$\langle p_x \rangle = \int \Psi^*(\boldsymbol{r}) \left(\frac{\hbar}{\mathrm{i}} \frac{\partial}{\partial x} \right) \Psi(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r}. \tag{IV.17}$$

Note the formal analogy between the right-hand sides of eqs. (IV.16) and (IV.17): one goes over from the one to the other by replacing the integration over $\boldsymbol{p}$ by an integration over $\boldsymbol{r}$, $\Phi(\boldsymbol{p})$ by its Fourier

transform, $\Phi^*(\boldsymbol{p})$ by its complex conjugate function, and p_x by $(\hbar/\mathrm{i})(\partial/\partial x)$, where $(\partial/\partial x)$ stands for partial derivation with respect to x applied to the function on its right.

In analogous fashion, one goes over from equation (IV.15) to the expression

$$\langle x \rangle = \int \Phi^*(\boldsymbol{p}) \left(\mathrm{i}\hbar \frac{\partial}{\partial p_x} \right) \Phi(\boldsymbol{p}) \, \mathrm{d}\boldsymbol{p}. \tag{IV.18}$$

These results can be generalized to more complex types of functions. Thus, from the fact that $p_x^2 \Phi(\boldsymbol{p})$ (assumed to be square-integrable) is the Fourier transform of

$$\left(\frac{\hbar}{\mathrm{i}} \frac{\partial}{\partial x} \right)^2 \Psi(\boldsymbol{r}) = -\hbar^2 \frac{\partial^2 \Psi}{\partial x^2}$$

(repeated application of Theorem A.III), one deduces

$$\langle p_x^2 \rangle = \int \Phi^* \, p_x^2 \, \Phi \, \mathrm{d}\boldsymbol{p} = -\hbar^2 \int \Psi^* \frac{\partial^2 \Psi}{\partial x^2} \, \mathrm{d}\boldsymbol{r}. \tag{IV.19}$$

More generally, if $G(\boldsymbol{p})$ is a polynomial or an absolutely converging series in p_x, p_y, p_z, one has

$$\langle G(\boldsymbol{p}) \rangle = \int \Psi^*(\boldsymbol{r}) \, G\left(\frac{\hbar}{\mathrm{i}} \nabla \right) \Psi(\boldsymbol{r}) \, \mathrm{d}\boldsymbol{r} \tag{IV.20}$$

provided that some straightforward conditions of convergence hold. Under the same conditions one finds for the average value of $F(\boldsymbol{r})$:

$$\langle F(\boldsymbol{r}) \rangle = \int \Phi^*(\boldsymbol{p}) \, F(\mathrm{i}\hbar \nabla_p) \, \Phi(\boldsymbol{p}) \, \mathrm{d}\boldsymbol{p}. \tag{IV.21}$$

Since the results obtained above are sufficient to undertake the main discussion of this chapter, we shall not pursue further the development of the statistical interpretation of the function Ψ. In addition to the statistics of the measurements of position and momentum, and the results derived therefrom concerning the mean values of quantities such as $F(\boldsymbol{r})$ and $G(\boldsymbol{p})$, the specification of Ψ must yield the statistics of the measurement of any measurable physical quantity. This question will be treated in Chapter V. We shall limit ourselves here to a few remarks which will guide us in its study.

The quantities $\langle x \rangle$ and $\langle p_x \rangle$ are real as a consequence of their definition. The right-hand sides of eqs. (IV.15) and (IV.17) are real as

well. In other words – and this is precisely the definition of hermiticity [cf. eq. (IV.8)] – the operators x and $(\hbar/\mathrm{i})(\partial/\partial x)$ are *Hermitean*. Similarly, the two other components of $\boldsymbol{r}$, and the two other components of $-\mathrm{i}\hbar\boldsymbol{\nabla}$ are Hermitean operators, and so are the operators $F(\boldsymbol{r})$, $G(-\mathrm{i}\hbar\boldsymbol{\nabla})$ if the functions F and G are real.

Consider the expressions of the mean values formed with the function Ψ [eqs. (IV.13), (IV.20), (IV.15), and (IV.17)]. They are all of the same form. With the quantity whose average value is to be taken is associated some linear (Hermitean) operator A and the value sought is given by the expression

$$\int \Psi^* A \Psi \, \mathrm{d}\boldsymbol{r} \tag{IV.22}$$

in which, according to usual convention, the operator acts on the function of $\boldsymbol{r}$ on its right. Operator A is obtained by a very simple correspondence rule: if we are dealing with a function $F(\boldsymbol{r})$ of the coordinates of its position, the corresponding operator is the function itself. If it is a function $G(\boldsymbol{p})$, the corresponding operator is obtained by substituting in G the respective components of $-\mathrm{i}\hbar\boldsymbol{\nabla}$ for those of $\boldsymbol{p}$. This is just the correspondence rule (II.17) between the momentum $\boldsymbol{p}$ and the operator $-\mathrm{i}\hbar\boldsymbol{\nabla}$ which led to the Schrödinger equation.

Each of these mean values may be calculated indifferently with either one of the wave functions Ψ and Φ: expressions (IV.21), (IV.14), (IV.18), and (IV.16) constructed with function Φ are respectively equivalent to expressions (IV.13), (IV.20), (IV.15), and (IV.17) constructed with function Ψ. They are even formally alike. To the quantity whose average value is to be taken there corresponds some linear (Hermitean) operator B – acting this time on functions of $\boldsymbol{p}$ – and the value sought is given by the expression

$$\int \Phi^* B \Phi \, \mathrm{d}\boldsymbol{p}. \tag{IV.23}$$

The operator B is found by a correspondence rule similar to that which led to A: if one is dealing with a function $G(\boldsymbol{p})$, the operator is the function itself; if it is a function $F(\boldsymbol{r})$, it is obtained by substituting in F the components of $\mathrm{i}\hbar\boldsymbol{\nabla}_{\boldsymbol{p}}$ [$\boldsymbol{\nabla}_{\boldsymbol{p}} \equiv (\partial/\partial p_x, \partial/\partial p_y, \partial/\partial p_z)$] for those of $\boldsymbol{r}$, respectively.

Just as the wave functions Φ and Ψ are equivalent *representations* of the same dynamical state, the operators B and A are equivalent representations of one and the same physical entity. The calculation

of physically measurable quantities such as the mean values considered here may be carried out in a formally identical manner in either of these representations. This suggests that Quantum Theory might be formulated in a general way, independent of any representation. This general formulation will be given in Chapters VII and VIII.

6. Generalization to Systems of Several Particles

The preceding definitions and results may be easily generalized to quantum systems of several particles.

Most generally, let $\Psi(q_1, \ldots, q_R; t)$ be the wave function of an R-dimensional quantum system, whose dynamical variables are its R coordinates $q_1, \ldots, q_R$ and the R conjugate momenta $p_1, \ldots, p_R$. We assume the coordinates to be cartesian, and denote by $\mathrm{d}\tau \equiv \mathrm{d}q_1 \ldots \mathrm{d}q_R$ and by $\mathrm{d}\omega \equiv \mathrm{d}p_1 \ldots \mathrm{d}p_R$ the volume elements in q space and p space, respectively. The wave function in p space is

$$\Phi(p_1, \ldots, p_R; t) = (2\pi\hbar)^{-\frac{1}{2}R} \int \Psi(q_1, \ldots, q_R; t)\, \mathrm{e}^{-(\mathrm{i}/\hbar)\sum_i p_i q_i}\, \mathrm{d}\tau.$$

$|\Psi|^2\, \mathrm{d}\tau$ is the probability of finding the coordinates q in region $(\tau, \tau + \mathrm{d}\tau)$, $|\Phi|^2\, \mathrm{d}\omega$ that of finding the momenta p in region $(\omega, \omega + \mathrm{d}\omega)$. Taking this as a starting point, one can repeat the entire procedure of the preceding sections.

Consider, for example, a two-particle system, and denote by $\boldsymbol{r}_1(x_1, y_1, z_1)$, $\boldsymbol{r}_2(x_2, y_2, z_2)$ their respective positions, and by $\boldsymbol{p}_1(p_{x_1}, p_{y_1}, p_{z_1})$, $\boldsymbol{p}_2(p_{x_2}, p_{y_2}, p_{z_2})$ their respective momenta. $P(\boldsymbol{r}_1, \boldsymbol{r}_2)\, \mathrm{d}\boldsymbol{r}_1 \mathrm{d}\boldsymbol{r}_2$ is the probability of finding particle 1 in volume element $(\boldsymbol{r}_1, \boldsymbol{r}_1 + \mathrm{d}\boldsymbol{r}_1)$ and particle 2 in volume element $(\boldsymbol{r}_2, \boldsymbol{r}_2 + \mathrm{d}\boldsymbol{r}_2)$; $\Pi(\boldsymbol{p}_1, \boldsymbol{p}_2)\mathrm{d}\boldsymbol{p}_1 \mathrm{d}\boldsymbol{p}_2$ is the probability of finding the momentum of particle 1 in the interval $(\boldsymbol{p}_1, \boldsymbol{p}_1 + \mathrm{d}\boldsymbol{p}_1)$ and that of particle 2 in interval $(\boldsymbol{p}_2, \boldsymbol{p}_2 + \mathrm{d}\boldsymbol{p}_2)$. One may also introduce the probability density of presence $P_1(\boldsymbol{r}_1)$ of particle 1 at the point $\boldsymbol{r}_1$, the position of particle 2 being unspecified. This quantity is the statistical distribution obtained when measuring the position of particle 1 without caring about the position of particle 2. Obviously:

$$P_1(\boldsymbol{r}_1) = \int P(\boldsymbol{r}_1, \boldsymbol{r}_2)\, \mathrm{d}\boldsymbol{r}_2.$$

Similarly one may introduce the probability densities $P_2(\boldsymbol{r}_2)$, $\Pi_1(\boldsymbol{p}_1)$, and $\Pi_2(\boldsymbol{p}_2)$. All these statistical distributions are essentially positive

quantities satisfying the normalization conditions

$$\iint P(\boldsymbol{r}_1, \boldsymbol{r}_2)\, \mathrm{d}\boldsymbol{r}_1\, \mathrm{d}\boldsymbol{r}_2 = 1, \ldots, \int \Pi_2(\boldsymbol{p}_2)\, \mathrm{d}\boldsymbol{p}_2 = 1$$

[the symbol $\iint \mathrm{d}\boldsymbol{r}_1\, \mathrm{d}\boldsymbol{r}_2$ designates the six-fold integral extended over all configuration space, the symbol $\int \mathrm{d}\boldsymbol{p}_2$ denotes a triple integral extended over all momentum space of particle 2, etc.]

The dynamical state of the system of two particles at a given time is defined by the wave function $\Psi(\boldsymbol{r}_1, \boldsymbol{r}_2)$ at that time. Its Fourier transform is the wave function in momentum space:

$$\Phi(\boldsymbol{p}_1, \boldsymbol{p}_2) = (2\pi\hbar)^{-3} \iint \mathrm{e}^{-\mathrm{i}(\boldsymbol{p}_1\cdot\boldsymbol{r}_1 + \boldsymbol{p}_2\cdot\boldsymbol{r}_2)/\hbar}\, \Psi(\boldsymbol{r}_1, \boldsymbol{r}_2)\, \mathrm{d}\boldsymbol{r}_1\, \mathrm{d}\boldsymbol{r}_2$$

$$\Psi(\boldsymbol{r}_1, \boldsymbol{r}_2) = (2\pi\hbar)^{-3} \iint \mathrm{e}^{\mathrm{i}(\boldsymbol{p}_1\cdot\boldsymbol{r}_1 + \boldsymbol{p}_2\cdot\boldsymbol{r}_2)/\hbar}\, \Phi(\boldsymbol{p}_1, \boldsymbol{p}_2)\, \mathrm{d}\boldsymbol{p}_1\, \mathrm{d}\boldsymbol{p}_2.$$

The generalization of definitions (IV.2) and (IV.6) is clearly

$$P(\boldsymbol{r}_1, \boldsymbol{r}_2) = |\Psi(\boldsymbol{r}_1, \boldsymbol{r}_2)|^2, \qquad \Pi(\boldsymbol{p}_1, \boldsymbol{p}_2) = |\Phi(\boldsymbol{p}_1, \boldsymbol{p}_2)|^2 \qquad \text{(IV.24)}$$

and the normalization conditions of the probabilities imply the normalization condition of the wave functions:

$$\iint |\Psi(\boldsymbol{r}_1, \boldsymbol{r}_2)|^2\, \mathrm{d}\boldsymbol{r}_1\, \mathrm{d}\boldsymbol{r}_2 = \iint |\Phi(\boldsymbol{p}_1, \boldsymbol{p}_2)|^2\, \mathrm{d}\boldsymbol{p}_1\, \mathrm{d}\boldsymbol{p}_2 = 1.$$

This normalization condition can actually be fulfilled at *any instant* if the Hamiltonian in the Schrödinger equation of the system is a Hermitean operator. This hermiticity property is easily verified. The wave functions Ψ and Φ are then determined to within an arbitrary phase constant.

From the preceding definitions one derives the definitions of various other distributions introduced above. Thus

$$P_1(\boldsymbol{r}_1) = \int |\Psi(\boldsymbol{r}_1, \boldsymbol{r}_2)|^2\, \mathrm{d}\boldsymbol{r}_2.$$

Similarly, one also obtains the average values of functions $F(\boldsymbol{r}_1, \boldsymbol{r}_2)$ and $G(\boldsymbol{p}_1, \boldsymbol{p}_2)$ of the positions and the momenta. Thus for instance

$$\langle x_1 \rangle = \iint \Psi^*\, x_1 \Psi\, \mathrm{d}\boldsymbol{r}_1\, \mathrm{d}\boldsymbol{r}_2 = -\frac{\hbar}{\mathrm{i}} \iint \Phi^* \frac{\partial \Phi}{\partial p_{x_1}}\, \mathrm{d}\boldsymbol{p}_1\, \mathrm{d}\boldsymbol{p}_2,$$

$$\langle p_{x_2} \rangle = \iint \Phi^*\, p_{x_2}\, \Phi\, \mathrm{d}\boldsymbol{p}_1\, \mathrm{d}\boldsymbol{p}_2 = \frac{\hbar}{\mathrm{i}} \iint \Psi^* \frac{\partial \Psi}{\partial x_2}\, \mathrm{d}\boldsymbol{r}_1\, \mathrm{d}\boldsymbol{r}_2.$$

All the remarks made at the end of § 5 remain entirely valid. If the wave function $\Psi(\boldsymbol{r}_1, \boldsymbol{r}_2)$ can be factored:

$$\Psi(\boldsymbol{r}_1, \boldsymbol{r}_2) = \Psi_1(\boldsymbol{r}_1)\, \Psi_2(\boldsymbol{r}_2)$$

[$\Psi_1(\boldsymbol{r}_1)$ and $\Psi_2(\boldsymbol{r}_2)$ are assumed normalized to unity], the same is true for the wave function in momentum space, and for the distributions P and Π:

$$P(\boldsymbol{r}_1, \boldsymbol{r}_2) = P_1(\boldsymbol{r}_1)\, P_2(\boldsymbol{r}_2), \qquad \Pi(\boldsymbol{p}_1, \boldsymbol{p}_2) = \Pi_1(\boldsymbol{p}_1)\, \Pi_2(\boldsymbol{p}_2),$$

as can easily be seen from the definitions of these quantities. In other words, there is no correlation in the statistics of measurements carried out on each of these particles. The statistical predictions concerning the results of measurement on one of them, particle 1, say, are the same as if it were in the dynamical state represented by the wave function $\Psi_1(\boldsymbol{r}_1)$. Indeed, one easily verifies that $P_1(\boldsymbol{r}_1) = |\Psi_1(\boldsymbol{r}_1)|^2$ and that $\Pi_1(\boldsymbol{p}_1) = |\Phi_1(\boldsymbol{p}_1)|^2$, $\Phi_1(\boldsymbol{p}_1)$ denoting the wave function of momentum space associated with Ψ_1. In all calculations bearing on the measurements carried out on this particle (average values, fluctuations, etc.) one may simply ignore the other particle and treat the problem as if one were dealing with a single particle whose wave function is $\Psi_1(\boldsymbol{r}_1)$.

If the two particles do not interact or if for some reason their interaction may be considered negligible, this property of factorization of the wave function persists in the course of time. Indeed the Hamiltonian of the system may then be put in the form of a sum of two terms:

$$H = H_1 + H_2,$$

of which one, H_1, acts only on the functions of the variable $\boldsymbol{r}_1$, and the other, H_2, on those of the variable $\boldsymbol{r}_2$. Assume that at the initial instant

$$\Psi(\boldsymbol{r}_1, \boldsymbol{r}_2; t_0) = \Psi_1(\boldsymbol{r}_1, t_0)\Psi_2(\boldsymbol{r}_2, t_0)$$

and let $\Psi_1(\boldsymbol{r}_1, t)$ and $\Psi_2(\boldsymbol{r}_2, t)$ be the solutions of the Schrödinger equation:

$$\left[i\hbar \frac{\partial}{\partial t} - H_1\right] \Psi_1 = 0, \qquad \left[i\hbar \frac{\partial}{\partial t} - H_2\right] \Psi_2 = 0,$$

with the respective initial conditions $\Psi_1(\boldsymbol{r}_1, t_0)$ and $\Psi_2(\boldsymbol{r}_2, t_0)$. The factored wave function

$$\Psi(\boldsymbol{r}_1, \boldsymbol{r}_2; t) = \Psi_1(\boldsymbol{r}_1, t)\Psi_2(\boldsymbol{r}_2, t)$$

satisfies the Schrödinger equation of the system, since

$$\begin{aligned} i\hbar \frac{\partial}{\partial t} \Psi &= i\hbar \left(\frac{\partial \Psi_1}{\partial t} \Psi_2 + \Psi_1 \frac{\partial \Psi_2}{\partial t} \right) = (H_1 \Psi_1) \Psi_2 + \Psi_1 (H_2 \Psi_2) \\ &= H_1 \Psi_1 \Psi_2 + H_2 \Psi_1 \Psi_2 = (H_1 + H_2) \Psi_1 \Psi_2 \\ &= H\Psi. \end{aligned}$$

The motions of each particle remain completely independent, as required by logic, and there is no correlation at any moment between the statistics of measurements carried out on either of them.

II. HEISENBERG'S UNCERTAINTY RELATIONS

7. Position-Momentum Uncertainty Relations of a Quantized Particle

Let us return to the definitions of probabilities of § 2. The distributions $P(\boldsymbol{r})$ and $\Pi(\boldsymbol{p})$ being defined in terms of the same wave function $\Psi(\boldsymbol{r})$, are not independent of each other, although the function $\Psi(\boldsymbol{r})$ could *a priori* be any square-integrable function. One of them can always be fixed arbitrarily by means of a suitable choice of Ψ; if, for instance, one picks $P(\boldsymbol{r})$ it suffices to take a function with absolute value $\sqrt{P}$: $\Psi(\boldsymbol{r}) = \sqrt{P(\boldsymbol{r})} \exp [i\alpha(\boldsymbol{r})]$, the phase $\alpha(\boldsymbol{r})$ remaining completely indeterminate. However, it is not possible to obtain by a suitable choice of $\alpha(\boldsymbol{r})$ any distribution $\Pi(\boldsymbol{p})$ specified in advance, although the distribution $\Pi(\boldsymbol{p})$ when considered as a function of $\alpha(\boldsymbol{r})$ can vary over a rather wide range. The fact that there always exists some correlation between the distributions $P(\boldsymbol{r})$ and $\Pi(\boldsymbol{p})$ is characteristic of Quantum Theory [1]). It is expressed quantitatively by *Heisenberg's uncertainty relations.*

---

[1]) Herein lies an essential difference between the statistical distributions $P(\boldsymbol{r})$ and $\Pi(\boldsymbol{p})$ on the one hand, and the corresponding distributions $P_{\text{cl.}}(\boldsymbol{r})$ and $\Pi_{\text{cl.}}(\boldsymbol{p})$ of classical *statistical* mechanics with which one would be tempted to make a comparison. The latter are obtained by means of the density in phase space $\varrho(\boldsymbol{r}, \boldsymbol{p})$:

$$P_{\text{cl.}}(\boldsymbol{r}) = \int \varrho(\boldsymbol{r}, \boldsymbol{p})\, d\boldsymbol{p}, \qquad \Pi_{\text{cl.}}(\boldsymbol{p}) = \int \varrho(\boldsymbol{r}, \boldsymbol{p})\, d\boldsymbol{r};$$

$\varrho(\boldsymbol{r}, \boldsymbol{p})$ is a positive function, subject to the condition $\iint \varrho(\boldsymbol{r}, \boldsymbol{p})\, d\boldsymbol{r}\, d\boldsymbol{p} = 1$, but otherwise arbitrary. Now one may fix arbitrarily and simultaneously the distributions $P_{\text{cl.}}(\boldsymbol{r})$ and $\Pi_{\text{cl.}}(\boldsymbol{p})$; indeed there exists at least one density in phase space, $\varrho(\boldsymbol{r}, \boldsymbol{p}) = P_{\text{cl.}}(\boldsymbol{r}) \Pi_{\text{cl.}}(\boldsymbol{p})$ which leads to these distributions.

Consider at first a particle in one dimension. Denote its position by x and its momentum by p; let $\psi(x)$ and

$$\varphi(p) = (2\pi\hbar)^{-\frac{1}{2}} \int_{-\infty}^{+\infty} \psi(x)\, e^{-(ipx)/\hbar}\, dx,$$

be the wave functions representing its dynamical state in x-space and p-space, respectively.

Heisenberg's result essentially expresses the mathematical fact that the extension of the wave ψ and that of its Fourier transform φ cannot simultaneously be made arbitrarily small. *If the wave ψ occupies a region of order Δx in x-space, and the wave φ occupies a region of order Δp in p-space, the product $\Delta x \cdot \Delta p$ always remains larger than a quantity of order $\hbar$*:

$$\Delta x \cdot \Delta p \gtrsim \hbar. \tag{IV.25}$$

We already met this result in connection with the construction of wave packets in the theory of matter waves. It also entered, although less directly, in the discussion of Chapter III on the motion of wave packets formed by superposition of unbound states.

One may convince oneself of its validity by the following semi-quantitative argument (which merely reproduces the analysis of Ch. II, § 3 in different terms). Any wave $\psi(x)$ is a superposition of plane waves $\exp(ikx)$ of wavelength $2\pi/k$. Let Δk be the extent of the domain of variation of parameter k. In order that the wave $\psi(x)$ be restricted to a certain region Δx, it is necessary that constructive interference among these different waves occurs only in that region, and that they interfere destructively everywhere else. The number of wavelengths $2\pi/k$ contained in Δx is $k\Delta x/2\pi$. In order that the various plane waves forming $\psi(x)$ may interfere destructively at the limits of the interval Δx, this number must change at least by one unit when k runs over its domain of variation: $\Delta x \cdot \Delta k \gtrsim 2\pi$. We are dealing here with an order of magnitude argument; we shall therefore not bother about the factor 2π and simply write

$$\Delta x \cdot \Delta k \gtrsim 1.$$

Since the momentum p is connected to k by the relation $p = \hbar k$, we arrive immediately at relation (IV.25).

One usually calls the quantities Δx and Δp the uncertainties in position and in momentum, respectively, and often states the result of Heisenberg in the following manner:

The product of the uncertainty in position and the uncertainty in momentum is necessarily greater than a quantity of order $\hbar$.

Let us illustrate this result by some examples.
The Gaussian wave packet (not normalized to unity)

$$\psi(x) = \exp\left[\frac{\mathrm{i}}{\hbar} p_0 x - \frac{(x-x_0)^2}{2\xi^2}\right]$$

occupies a region of dimension ξ about the point x_0. The wave

$$\varphi(p) = \frac{\xi}{\sqrt{\hbar}} \exp\left[\frac{\mathrm{i}}{\hbar} x_0(p_0-p) - \tfrac{1}{2}\frac{\xi^2}{\hbar^2}(p-p_0)^2\right]$$

of momentum space corresponding to it (cf. table at the end of Appendix A) occupies in that space a region of extension $\hbar/\xi$

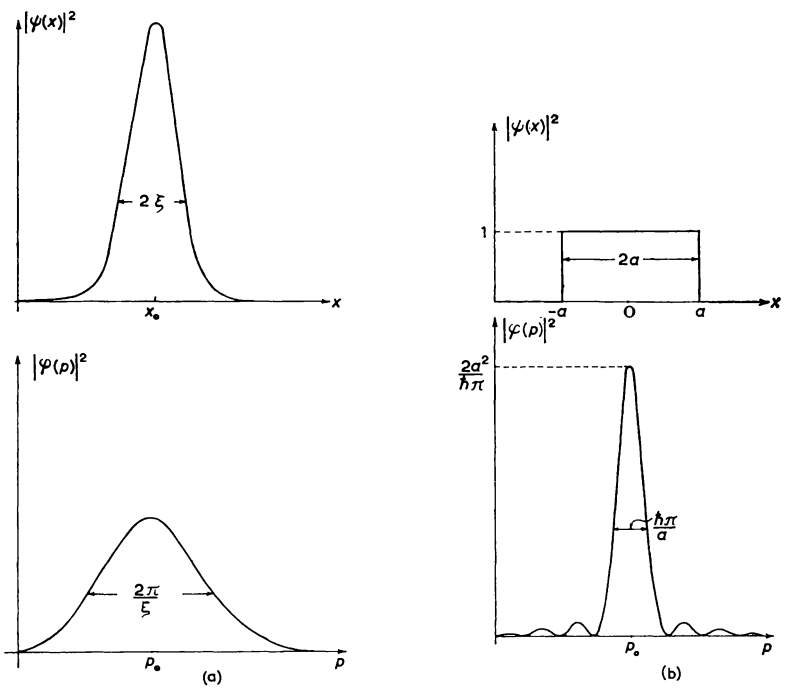


Fig. IV.1. Square modulus of the wave packets $\psi(x)$ and $\varphi(p)$ in the case: (*a*) of a Gaussian wave packet; (*b*) of a square pulse.

about the point p_0. Upon diminishing ξ, one decreases Δx, but one increases Δp by a corresponding amount so that their product remains of order $\hbar$.

As another example, let us consider the "square pulse"

$$\psi(x) = \begin{cases} e^{ip_0 x/\hbar} & \text{for} \quad |x| < a \\ 0 & \text{for} \quad |x| > a \end{cases}$$

which extends over a region $2a$ surrounding the point $x=0$. We have in this case (cf. Appendix A)

$$\varphi(p) = \frac{\sqrt{(2\hbar/\pi)}}{p-p_0} \sin \frac{(p-p_0)a}{\hbar}.$$

The function $|\varphi(p)|^2$ exhibits a very sharp peak at the point p_0 surrounded on either side by a succession of vanishing minima [reached when $p = p_0 + (n\pi\hbar/a)$] separated by maxima whose height decreases as $[1/(p-p_0)]^2$. One can say that the wave $\varphi(p)$ is mainly concentrated between the first zeros of $|\varphi(p)|^2$ located on either side of the main peak, i.e. an extension $\Delta p \approx 2\pi\hbar/a$. The smaller the extension $\Delta x \approx 2a$, the larger Δp:

$$\Delta x \cdot \Delta p \approx 4\pi\hbar.$$

Figure IV.1 shows the curves $|\psi(x)|^2$ as a function of x and $|\varphi(p)|^2$ as a function of p for these two wave packets.

All the considerations on the extension of the wave ψ compared to that of its Fourier transform may be easily extended to the case of a particle in three-dimensional space. Denote by $\Delta x, \Delta y, \Delta z$ the respective uncertainties in the three position coordinates, and by $\Delta p_x, \Delta p_y, \Delta p_z$ the corresponding quantities for the momentum. The correlation between the statistical distributions $P(\mathbf{r}) = |\Psi(\mathbf{r})|^2$ and $\Pi(\mathbf{p}) = |\Phi(\mathbf{p})|^2$ is expressed by the uncertainty relations

$$\begin{aligned} \Delta x \cdot \Delta p_x &\gtrsim \hbar, \\ \Delta y \cdot \Delta p_y &\gtrsim \hbar, \\ \Delta z \cdot \Delta p_z &\gtrsim \hbar. \end{aligned} \tag{IV.26}$$

Up to now, the uncertainty relations have been presented as order of magnitude relations. Of course this is inevitable as long as one has not adopted a precise definition of the quantities $\Delta x, \Delta p_x$, etc., measuring the various uncertainties. By taking suitable definitions

for these quantities, we shall arrive at a precise statement. However, without underestimating the virtues of a precise statement might have, one must insist on the fact that the essential interest of the uncertainty relations is already contained in the order of magnitude result: one may not, under any circumstances, attribute simultaneously to the quantum particle a rigorously precise position and a rigorously precise momentum. To think of the particle as endowed with precise position and momentum is justified only to the extent that the quantum of action $\hbar$ may be considered negligible, that is, in the domain of validity of Classical Theory.

8. Precise Statement of the Position-Momentum Uncertainty Relations

Let us first treat in detail the case of a particle in one dimension. We take for the uncertainties in x and p the following precise definitions [1]: Δx and Δp are the root-mean-square deviations of the distributions $|\psi(x)|^2$ and $|\varphi(p)|^2$. Adopting the notation of Chapter IV, § 5, we therefore have

$$\begin{aligned} \Delta x &\equiv \sqrt{\langle x^2 \rangle - \langle x \rangle^2} \\ \Delta p &\equiv \sqrt{\langle p^2 \rangle - \langle p \rangle^2}. \end{aligned} \tag{IV.27}$$

The quantity Δx is thus directly related to the position measurement: it is the statistical fluctuation of the result of measurement around the average value $\langle x \rangle$; the same remark applies to Δp relative to the momentum measurement. We proceed to show the following very general result:

$$\Delta x \cdot \Delta p \geqslant \tfrac{1}{2}\hbar. \tag{IV.28}$$

Let us consider the positive-definite expression

$$I(\lambda) = \int_{-\infty}^{+\infty} \left| x\psi + \lambda\hbar \frac{\partial \psi}{\partial x} \right|^2 \mathrm{d}x \qquad [I(\lambda) \geqslant 0 \text{ for any } \lambda].$$

---

[1] This is probably the most convenient definition one might imagine. In most of the cases, the quantities Δx and Δp thus defined give a good idea of the uncertainty in x and p. Nevertheless, it sometimes happens that this mathematical definition of the uncertainty differs rather substantially from the rough estimates one might make. The "square pulse" is an example of this type of accident: the quantity Δp defined by equation (IV.27) is infinite whereas a rough estimate of this quantity gave $2\pi\hbar/a$.

Developing and integrating by parts, we obtain successively

$$I(\lambda) = \int_{-\infty}^{+\infty} |x\psi|^2 \, dx + \lambda\hbar \int_{-\infty}^{+\infty} \left(\frac{\partial \psi^*}{\partial x} x\psi + x\psi^* \frac{\partial \psi}{\partial x} \right) dx + \lambda^2\hbar^2 \int_{-\infty}^{+\infty} \left| \frac{\partial \psi}{\partial x} \right|^2 dx$$

$$= \int_{-\infty}^{+\infty} \psi^* x^2 \psi \, dx - \lambda\hbar \int_{-\infty}^{+\infty} \psi^* \psi \, dx - \lambda^2\hbar^2 \int_{-\infty}^{+\infty} \psi^* \frac{\partial^2 \psi}{\partial x^2} \, dx,$$

which yields, assuming ψ to be normalized to unity and using the results of § 5:

$$I(\lambda) = \langle x^2 \rangle - \lambda\hbar + \lambda^2 \langle p^2 \rangle. \tag{IV.29}$$

Since the polynomial $I(\lambda)$ of the second degree in λ is positive-definite (or zero) its discriminant $\hbar^2 - 4\langle p^2 \rangle \langle x^2 \rangle$ is necessarily negative (or zero), thus:

$$\langle x^2 \rangle \langle p^2 \rangle \geqslant \tfrac{1}{4}\hbar^2. \tag{IV.30}$$

Condition (IV.30) is less restrictive than condition (IV.28). But one can make an identical calculation starting from a slightly different expression $I(\lambda)$, i.e. by replacing in the definition of $I(\lambda)$, x by $x - \langle x \rangle$ and $\hbar(\partial/\partial x)$ by $\hbar(\partial/\partial x) - \mathrm{i}\langle p \rangle$ [or, what amounts to the same thing, replacing $\psi(x)$ by $\exp(-\mathrm{i}\langle p \rangle x/\hbar)\ \psi(x + \langle x \rangle)$]. The result is analogous to eq. (IV.29):

$$I(\lambda) = (\Delta x)^2 - \lambda\hbar + \lambda^2 (\Delta p)^2 \geqslant 0.$$

Condition (IV.28) expresses the fact that the discriminant of this second-degree expression in λ is never positive. Q.E.D.

The foregoing proof applies equally well to a particle in three-dimensional space. The wave function $\Psi(r)$ is a function of three-dimensional space and the integrals introduced in the proof extend over the three dimensions of configuration space instead of extending only along the x axis. The reader will easily verify that all the operations carried out on these integrals remain valid. Similarly, the definition (IV.27) of the root-mean-square deviations is easily generalized, and one obtains the *Heisenberg uncertainty relations*:

$$\begin{aligned} \Delta x \cdot \Delta p_x &\geqslant \tfrac{1}{2}\hbar \\ \Delta y \cdot \Delta p_y &\geqslant \tfrac{1}{2}\hbar \\ \Delta z \cdot \Delta p_z &\geqslant \tfrac{1}{2}\hbar. \end{aligned} \tag{IV.31}$$

9. Generalization: Uncertainty Relations Between Conjugate Variables

One finds the same relations in the general case of R-dimensional quantum systems. We use the notation of § 6. By extension of the definition (IV.27) we measure the uncertainty in q_i and p_i, respectively, by the root-mean-square deviations of their statistical distribution:

$$\begin{aligned} \Delta q_i &= \sqrt{\langle q_i^2 \rangle - \langle q_i \rangle^2}, \\ \Delta p_i &= \sqrt{\langle p_i^2 \rangle - \langle p_i \rangle^2}. \end{aligned} \tag{IV.32}$$

The proof of the foregoing section can be repeated without modification leading to the uncertainty relations between conjugate (cartesian) variables

$$\Delta q_i \cdot \Delta p_i \geqslant \tfrac{1}{2}\hbar \qquad (i = 1, 2, \ldots, R).$$

10. Time-Energy Uncertainty Relation

Position-momentum uncertainty relations originate from the fact that the momentum is defined, to within a constant, as the characteristic wave number of a plane wave, and that, rigorously speaking, a plane wave extends over all space; to localize the momentum at an exact point of space has no more meaning than to localize a plane wave.

Just as momentum is a wave number and cannot be localized in space, so energy is a frequency and cannot be localized in time. Thus there exists, as suggested by the principle of relativity, a time-energy uncertainty relation, analogous to the position-momentum uncertainty relations discussed previously:

$$\Delta t \cdot \Delta E \gtrsim \hbar. \tag{IV.33}$$

However, its physical interpretation is quite different. In the position-momentum uncertainty relations, the position and momentum variables play exactly symmetrical roles; they can both be measured at *a given time* t. The statistical distributions of the results of measurement and consequently the uncertainties Δq_i, Δp_i are all derivable from the value of the wave function at that time. In relation (IV.33), on the other hand, energy and time play fundamentally different roles: the energy is a dynamical variable of the system, whereas the time t is a parameter. Relation (IV.33) *connects the uncertainty* ΔE in the value taken by this dynamical variable *to a time interval* Δt *characteristic of the rate of change of the system.*

At first, let us consider the free particle case. The monochromatic plane wave $\exp[i(\mathbf{k}\cdot\mathbf{r}-\omega t)]$ represents a particle of well-defined momentum $\hbar\mathbf{k}$ and energy $\hbar\omega$. By superposition of plane waves, one can form a wave packet such as the one of eq. (II.11). For simplicity we consider a wave packet in one dimension and take a wave train such as the square pulse studied above. Denote its length by Δx and its group velocity by v. It travels with velocity v along the x axis; however, the instant at which it passes a given point of the axis is not determined precisely, but carries an uncertainty

$$\Delta t \approx \frac{\Delta x}{v}.$$

Furthermore, this wave train has a certain spread in momentum space, hence a certain uncertainty ΔE in the value of the energy of the particle

$$\Delta E \approx \frac{\partial E}{\partial p}\Delta p = v\Delta p.$$

From the two foregoing equations we deduce

$$\Delta t\cdot\Delta E \approx \Delta x\cdot\Delta p$$

and, applying the momentum-position uncertainty relation, we obtain relation (IV.33) which sets a lower limit to the product of the spread ΔE of the energy spectrum of the particle, and the precision Δt with which the instant of passage of the particle at a given point can be predicted.

The preceding proof can be extended without difficulty to the case of a wave packet in a slowly varying field, but ceases to be valid in more general cases. In order to obtain a relation such as (IV.33) one must consider the time dependence of the wave function in these more general cases.

The simplest situation obtains when the quantum system has a well-defined energy. We know (cf. Ch. II, § 16) that the wave function of a quantum system of given energy $E=\hbar\omega$ varies in time proportionally to the exponential factor $\exp(-i\omega t)$. Consider, for example, the case of a particle in a force field. If its quantum state corresponds to a well-defined value of the energy $E=h\omega$, its wave function is written $\Psi(\mathbf{r},t)=\psi(\mathbf{r})\exp(-i\omega t)$. Consequently, the distribution of position of this particle $P(\mathbf{r})=|\psi(\mathbf{r})|^2$ is independent of time. The

momentum distribution is easily seen to possess this same property. Consequently, the result of a measurement (of position or momentum) carried out on this system is independent of the moment at which the measurement is made. This is expressed in short by saying that the physical properties of such a system are time-independent, or else that the system is in a *stationary state*.

Assume now that the quantum state of the particle is the superposition of two stationary states of energy E_1 and E_2, respectively. Its wave function is of the form

$$\psi_1(\boldsymbol{r})\, e^{-iE_1t/\hbar} + \psi_2(\boldsymbol{r})\, e^{-iE_2t/\hbar}$$

and the distribution

$$P(\boldsymbol{r}, t) = |\psi_1(\boldsymbol{r})|^2 + |\psi_2(\boldsymbol{r})|^2 + 2\mathrm{Re}\ \psi_1^* \psi_2\, e^{i(E_1-E_2)t/\hbar}$$

oscillates in time [between two extreme values $(|\psi_1| - |\psi_2|)^2$ and $(|\psi_1| + |\psi_2|)^2$] with the period

$$\tau = \frac{\hbar}{|E_1 - E_2|}.$$

The momentum distribution has the same property.

τ thus appears as a characteristic time for the rate of change of the physical properties of the system. The result of measurement — more precisely, the statistical distribution of the results of measurement — made at two different times t_1 and t_2 will be practically identical if the difference $\Delta t = |t_1 - t_2|$ is small compared to τ. In other words, in order that the physical properties of the system be notably modified over a time interval Δt, the product of Δt and the energy uncertainty $\Delta E = |E_1 - E_2|$ must at least be equal to a quantity of the order of $\hbar$: $\Delta t \cdot \Delta E \gtrsim \hbar$. Expressed in this fashion, this result remains valid when the state of the system is any superposition of any number of stationary states. It is thus entirely general. A precise proof thereof will be given later on (Ch. VIII, § 13).

An important application of relation (IV.33) is the *lifetime-width relation* for radioactive systems (radioactive nucleus, excited state of an atom, unstable elementary particle, etc.). A radioactive system is not stationary; it does not correspond to a well-defined value of the energy, but to an energy spectrum with a certain spread ΔE, usually called the level width. The mean lifetime τ here plays the role of the

characteristic time considered above. One must wait for a time of order τ to observe an appreciable change in the properties of the system. Consequently

$$\tau \cdot \Delta E \approx \hbar.$$

Another consequence of (IV.33) has to do with energy measurements in general. *The precision ΔE of the measurement is connected with the time Δt required for the measurement by relation* (IV.33). Thus, one may consider the measurement of the excitation energy of the first-excited state of the hydrogen atom by bombarding it with monoërgic electrons and measuring the energy lost by the electrons in the corresponding inelastic collision. The duration of the measurement is at least equal to the collision time, that is to say to the time of passage Δt of the wave packet representing the electron at the location of the hydrogen atom. The measurement error is at least equal to the uncertainty ΔE in the energy of the incident electron, and we have in fact $\Delta t \cdot \Delta E \gtrsim \hbar$.

11. Uncertainty Relations for Photons

The uncertainty relations for systems of material particles follow from the wave-corpuscle duality of matter. For the same reason, there exist analogous relations for photons. But when formulating them one must be aware that the number of photons contained in a physical system is not in general a well-defined quantity, and that one cannot, strictly speaking, consider the motion of a particular photon unless it is free from any interaction.

With these restrictions in mind we may nevertheless represent a free photon by a wave packet formed by superposition of plane, monochromatic waves propagating at velocity c [1]).

Thus, by illuminating a screen equipped with a diaphragm, which one keeps open during a certain time Δt, one obtains a light signal which eventually may contain but a single light quantum. This photon is represented by a wave packet whose extension along the three directions of the axes, Δx, Δy, Δz, depends upon the dimensions of the diaphragm, and upon its open time. This wave packet is a super-

[1]) To simplify matters we ignore the existence of the polarization of light. To take it into account, one must attribute to the photon an internal degree of freedom.

position of monochromatic waves. It has all the properties of the wave packet we studied previously. It does not represent a photon of well-defined momentum and energy, but the components of its momentum and its energy cover a certain finite region, namely Δp_x, Δp_y, Δp_z, and ΔE, respectively.

All these quantities satisfy the relations

$$\Delta x \cdot \Delta p_x \gtrsim \hbar, \qquad \Delta y \cdot \Delta p_y \gtrsim \hbar, \qquad \Delta z \cdot \Delta p_z \gtrsim \hbar, \qquad \Delta t \cdot \Delta E \gtrsim \hbar.$$

From these relations, some interesting conclusions may be drawn concerning the mechanism of interaction between such a photon and matter. Thus, in the absorption of a photon by an atom (photoelectric effect), the product of the uncertainty ΔE in the energy transfer to the atom, and the uncertainty Δt concerning the instant at which this energy is transferred, is at least of order $\hbar$. Conversely, if an atom initially in an excited state, decays to the ground state by emitting a photon, the moment at which this transition occurs cannot be defined with a precision greater than the mean life τ of this excited state. The emitted photon is represented by a wave packet of spatial extension $c\tau$ and, consequently, its spread in energy ΔE is such that $\tau \cdot \Delta E \approx \hbar$. This result is in good agreement with experiment. One could also have obtained it by starting from the relation between mean life and width discussed earlier, and noting that, owing to the conservation of energy, the spread in energy of the emitted photon (final state) must be equal to the spread in energy of the excited atom (initial state).

III. UNCERTAINTY RELATIONS AND THE MEASUREMENT PROCESS

12. Uncontrollable Disturbance During the Operation of Measurement

In the following, we shall concentrate particularly on the position-momentum uncertainty relations. The spreads Δx, Δp which enter into these relations refer to measurements carried out under the conditions stated earlier when we defined the probabilities. They must no be confused with the ordinary errors of measurement caused by the fact that the measuring instrument is never perfect and does not permit the determination of the measured quantities with infinite precision. In all previous reasoning this type of error is assumed negligible.

We proceed to discuss more closely the mechanism of a measurement and its consistency with the statistical interpretation outlined above. Consider, for instance, the measurement of the position coordinate of a particle in a quantized system whose dynamical state is represented by the wave function Ψ (we shall henceforth say, more briefly, that the system is in the state Ψ). We assume that the measuring device at our disposal is *infinitely precise*. Thus, the impossibility to predict with certainty the result of observation is not due to any imperfection of the measuring device; the state Ψ does not in general correspond to a precise value of x; rather, it is a superposition of dynamical states, each corresponding to a given value of x. Immediately after the operation of measurement, we can assert that the system is in a dynamical state where the x coordinate has the precise value x' indicated by the measuring apparatus. Such a state can definitely not be represented by the wave function Ψ: *the intervention of the measuring device has modified the dynamical state of the measured system. Moreover, this perturbation of the system during the measuring operation is to a certain extent uncontrollable* in the sense that one cannot predict exactly what the state of the system will be after the measurement, but only the probability that it is in a dynamical state corresponding to one of several values x' of the x coordinate.

It is not too surprising that there should be a disturbance of the system in the course of measurement. In order to obtain a measurement, it is necessary to let the system interact in some way with an appropriate measuring apparatus; the latter undergoes some change during the process. Such a change is necessary, since it determines the response of the apparatus, that is to say, the value taken by the physical quantity one wishes to measure. Conversely, one cannot avoid some perturbation of the system itself during the process of measurement.

On the scale of macroscopic phenomena, one can always proceed in such a way that this disturbing action of the measuring device on the system is negligible, or at least known to a good precision. Let us illustrate this point with an example. One can determine the position of a macroscopic object by forming its image with the aid of a set of lenses on a photographic plate. The dynamical state of the object is unavoidably modified during the measurement owing to the very fact that it is exposed to light (radiation pressure). In the classical approximation where the incident light can be treated as a continuous

wave (numerous photons present), this modification may in principle be calculated exactly; provided that the photographic plate is sensitive enough, one can in fact decrease the illumination of the object indefinitely and thus make the modification arbitrarily small.

The above argument is evidently valid only in the limit of the classical approximation. In fact, the action of the measuring instrument on the object cannot be decreased indefinitely since this mutual interaction takes place via discrete quanta. The deflection angle of a light quantum by the object is not defined in a precise manner, in other words, the momentum transfer during the deflection may be considered uncontrollable. The position measurement is thus accompanied by an uncontrollable change in the momentum of the object (the semi-quantitative analysis of this effect will be given in the following section). Since modifications of this type are entirely negligible on the macroscopic scale, classical theory postulates that *all* the dynamical variables of a physical system can be measured simultaneously with an arbitrarily small error; the dynamical state of the system is then defined at any given time by the precise specification of the values assumed by all these variables at that instant. On the level of precision of microscopic phenomena, this postulate lacks experimental support and must be abandoned. *The quantum theory, however, assumes that the unpredictable and uncontrollable disturbance suffered by the physical system during a measurement is always sufficiently strong so that the uncertainty relations always holds true.* Thus, in the measurement of x contemplated above, the system has gone over from state Ψ to state Ψ'. As was already pointed out, state Ψ in general corresponds neither to a precise value of x, nor to a precise value of p, but to a probability distribution $P(x)$ of finding some value of x when a precise measurement of this quantity is undertaken; or to a probability distribution $\Pi(p)$ of finding some value of p in the case of a precise measurement of p. To the new state Ψ' correspond new distributions $P'(x)$ and $\Pi'(p)$; in particular the root-mean-square deviations $\Delta' x$ and $\Delta' p$ of these distributions necessarily satisfy the relation $\Delta' x \cdot \Delta' p \geqslant \frac{1}{2}\hbar$. In particular, if the measurement of x is infinitely precise ($\Delta' x = 0$), the quantity $\Delta' p$ must be infinite. One often expresses this result by stating that one cannot diminish the uncertainty in the position variable x without increasing the uncertainty in the momentum variable p by at least a corresponding amount, and *vice versa*.

We shall examine a few specific examples of measurements and show that the perturbation suffered by the measured system is actually always sufficiently large so that the uncertainty relations of Heisenberg are never contradicted [1]).

13. Position Measurements

a) Use of a Diaphragm. – Consider a beam of monoërgic electrons moving parallel to Oz. Let us measure the component of the electron position along some direction Ox perpendicular to Oz. To this effect, one inserts into the path of the beam a screen with a slit (Fig. IV.2). If d is the width of the slit, the x-component of the position of any electron crossing this diaphragm is defined with a precision $\Delta x = d$.

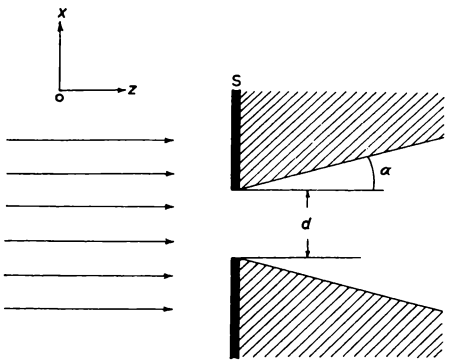


Fig. IV.2. Measurement of position by means of a diaphragm.

However, this electron is represented by a de Broglie wave of wavelength $\lambda = \hbar/p$. The crossing of the diaphragm is accompanied by a

[1]) This especially happens in all experimental arrangements which permit the "trajectory" of a particle to be displayed, as is the case for photographic plates or Wilson cloud chambers for the observation of charged particles. Thus in the cloud chamber, the particle ionizes a certain number of atoms along its path, and these ions become centers of condensation on which visible droplets form. The position of the particle is thus obtained with an uncertainty Δq at least equal to the radius of the ionized atom (in practice it is much larger). However, in the interaction with the measuring device, that is to say in the process of ionization of the atom by the particle, the latter suffers an unpredictable and uncontrollable momentum transfer Δp of the order of $\hbar/\Delta q$. The "trajectory" of the particle can thus be observed only within the limits of precision $\Delta q \cdot \Delta p \approx \hbar$.

diffraction effect; the beam therefore diverges by a certain angle α of the order

$$\sin \alpha \approx \frac{\lambda}{d} = \frac{h}{p \cdot \Delta x}.$$

This amounts to a spread $\Delta p_x = p \sin \alpha$ of the electron momentum in the direction Ox, and we have

$$\Delta x \cdot \Delta p_x \approx h.$$

The x-component of the electron momentum, assumed perfectly well-known before the measuring operation ($p_x = 0$) has been shifted by an uncontrollable amount of order $\hbar/\Delta x$ during the measuring process (crossing of the diaphragm).

It is important to convince oneself thoroughly here that the momentum transferred from the diaphragm to the electron during the measurement cannot actually be determined to better than within $\hbar/\Delta x$; otherwise the preceding argument would be vitiated. Indeed, to carry out the position measurement, the diaphragm must be kept stationary and its position (along Ox) must be known to within δx, δx being very small compared to the slit width: $\delta x \ll \Delta x$. But the diaphragm is a quantum object, just like the electron. Its momentum is therefore not defined to better than δp, and we have

$$\delta p \approx \frac{\hbar}{\delta x} \gg \frac{h}{\Delta x}.$$

It can nevertheless remain practically motionless in the course of the measurement provided it is sufficiently heavy; this limitation therefore does not interfere with the measuring operation itself. But it is obviously not possible to determine the momentum variation of the diaphragm with a greater precision than δp and *a fortiori* with a greater precision than $h/\Delta x$.

This discussion stresses an important point. One must postulate that *the measuring apparatus is a quantized object which also obeys the uncertainty relations. This supposes that the uncertainty relations are of an entirely universal nature.* Otherwise, the physical interpretation of the Quantum Theory would have to be profoundly revised.

b) Use of the Microscope. — The use of a diaphragm is undoubtedly the most direct way to measure the position of an object. A less

direct but equally valid method consists in illuminating that object and observing the image through a microscope. That is, we are considering the determination of the position x of an electron by observing it through a microscope (Fig. IV.3). The precision of the measurement is limited by the fact that the image of each point is a

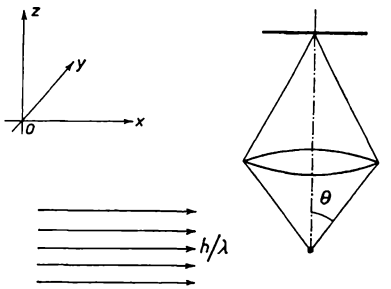


Fig. IV.3. Measurement of position by means of a microscope.

diffraction spot of finite extension. From the extension of that spot one can deduce the limit of precision Δx of the measurement:

$$\Delta x \approx \frac{\lambda}{\sin \theta}$$

where λ is the wavelength of the light, and θ is the half-angle of divergence of the beam scattered by the electron and focused in the microscope. However, the scattering of light proceeds by discrete quanta and is accompanied by a partly uncontrollable momentum transfer (Compton effect). This effect is at a minimum when the scattered light contains but one single photon; the photon momentum has a well-defined magnitude, $p = h/\lambda$; its direction of propagation, however, is defined only to within the angle θ. The momentum transferred to the electron is therefore not exactly known; it has an inherent uncertainty

$$\Delta p \approx \frac{h}{\lambda} \sin \theta.$$

The more precise the position measurement, the more important is this effect, and one always has: $\Delta x \cdot \Delta p \approx \hbar$.

One might object that the precision in x is not given by the size of the diffraction spot, but by the precision to which the center of that spot can be determined. This precision is greater the larger

the number N of photons participating in the formation of the spot. According to the laws of statistics, the error in x calculated in this way is $\sqrt{N}$ times smaller than the error computed above:

$$\Delta x \approx \frac{1}{\sqrt{N}} \frac{\lambda}{\sin \theta}.$$

However, if the momentum transferred by each photon has an inherent uncertainty $(h \sin \theta)/\lambda$, the uncertainty in the momentum transferred by the N photons is $\sqrt{N}$ times greater (addition of the square of the errors), hence

$$\Delta p \approx \sqrt{N} \frac{h}{\lambda} \sin \theta.$$

Consequently: $\Delta x \cdot \Delta p \approx h$.

14. Momentum Measurements

The same conclusions hold for momentum measurements. The momentum of a particle may *a priori* be measured with arbitrary precision; however, the measuring process is necessarily accompanied by a perturbation which increases the uncertainty in the corresponding position coordinate, in such a way that the uncertainty relations continue to be satisfied. This will be shown in the following two examples.

a) Deflection in a Magnetic Field. – The momentum of a charged particle is usually measured by deflection in a constant magnetic field. The momentum p is related to the radius of curvature R of the particle trajectory by the well-known relation

$$p = \frac{e}{c} \mathscr{H} R,$$

where $\mathscr{H}$ is the magnitude of the magnetic field and e is the charge of the particle.

We shall examine the measurement of the momentum of an electron by this method. Figure IV.4 shows the schematic diagram of the measuring set-up. The electron enters the magnet after passing through the diaphragm A and leaves it through diaphragm B after having suffered a 180-degree deflection (this particular angle has been chosen for simplicity). At the instant immediately *preceding* the

beginning of the measurement (i.e. just before the crossing of diaphragm A), we assume the direction of propagation perfectly defined (direction Oy) and the y coordinate of the electron in this direction

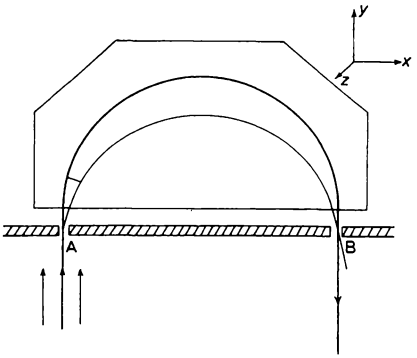


Fig. IV.4. Measurement of momentum by deflection in a magnetic field.

perfectly known ($\Delta y = 0$). These initial conditions ($p_x = p_z = 0$, $y = y_A$) can always be realized in principle by means of a collimator equipped with a shutter whose "exposure time" is sufficiently short. The radius of curvature R is equal to half the distance between the two diaphragms; if $2d$ and $2d'$ are the respective widths of these diaphragms, R is measured to within $d + d'$. The momentum of the electron is therefore known to a precision

$$\Delta p = \frac{e}{c} \mathscr{H}(d + d') = \frac{p}{R}(d + d').$$

The quantity measured here is the momentum component along the axis Oy. Let us show that once the measurement is carried out, the y coordinate of the electron has an uncertainty Δy such that $\Delta y \cdot \Delta p \gtrsim h$. The quantum effect which enters here is the diffraction of the electron wave upon crossing the diaphragm A (the reader can convince himself that diffraction by diaphragm B plays no role whatsoever in this problem). If this effect did not exist, the momentum would be exactly parallel to Oy when entering the magnet; the electron would then very accurately describe a semi-circle from A to B and the time of flight from A to B would be independent of the value of p and equal to $\pi mc/e\mathscr{H}$ (m = mass of the electron). Because of this effect, the angle between the momentum and the Oy axis at the entrance

of the magnet has an inherent indeterminacy $\alpha \approx \lambda/d = h/pd$; the trajectory of the electron (i.e. the ray in the geometrical optics approximation) is a circular segment defined to within 2α. The moment when the electron reaches B has an indeterminacy $\Delta t = 2\alpha mc/e\mathscr{H}$, rather than being defined in a precise manner; the uncertainty Δy is p/m times larger, namely

$$\Delta y \approx 2\alpha \frac{pc}{e\mathscr{H}} \approx 2h \frac{c}{e\mathscr{H}d}.$$

Therefore

$$\Delta y \cdot \Delta p \approx 2h \left(1 + \frac{d'}{d}\right).$$

b) Collision with a Photon. – Another method of measuring momentum consists in letting the particle under consideration collide with another particle, a photon, say, whose initial momentum is perfectly well-known and then to measure the momentum transferred to that second particle in the collision. Once again let us take the electron of the foregoing problem characterized by the initial conditions $p_x = p_z = 0$, $y = y_A$. To measure its momentum p_y, we illuminate it with perfectly monochromatic light of frequency ν and direction of propagation parallel to the y axis. One of the photons of this radiation might suffer a Compton collision and one measures its final momentum. For simplicity, we shall assume the latter to be also directed along Oy (in the opposite sense) (Fig. IV.5). Let ν' be the

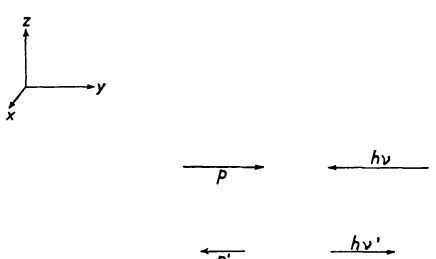


Fig. IV.5. Measurement of momentum of an electron by Compton collision with a photon. Scheme of the collision.

frequency of the photon after collision; the theory of the Compton effect yields the initial momentum p and final momentum p' of the electron as a function of the frequencies ν and ν'. We shall assume

that the conditions of non-relativistic approximation hold ($p, p' \ll mc$ and $\nu, \nu' \ll mc^2/h$). One obtains after some calculations

$$p = mc\,\frac{\nu'-\nu}{\nu'+\nu} + \frac{h}{2c}(\nu'+\nu), \qquad p' = mc\,\frac{\nu'-\nu}{\nu'+\nu} - \frac{h}{2c}(\nu'+\nu),$$

and the precision in the determination of these quantities is connected with the precision in the determination of ν' by the relation

$$\Delta p \approx \Delta p' \approx mc\,\frac{\Delta\nu'}{\nu'+\nu}.$$

The position y of the electron after measurement can be calculated knowing its initial value y_A and that the electron travels with velocity p/m before the Compton collision, and with velocity p'/m afterwards. If the momentum and the position of the scattered photon could be measured simultaneously with extreme precision, the moment when the collision takes place could be determined in a precise manner, and the indeterminacies in p and y could be simultaneously made very small. The reader can easily verify this fact. Indeed, the measurement of ν' is a frequency measurement; the more precise it is, the larger the indeterminacy in the time when the photon passes a given point, and in particular the larger the uncertainty Δt in the time of collision: $\Delta\nu' \cdot \Delta t \gtrsim 1$. Now, the uncertainty in the position of the electron after collision Δy is at least equal to the product of Δt and the change in velocity, hence

$$\Delta y \gtrsim \frac{|p-p'|}{m\Delta\nu'} = \frac{h}{mc}\,\frac{\nu+\nu'}{\Delta\nu'},$$

which yields

$$\Delta y \cdot \Delta p \quad \text{and} \quad \Delta y \cdot \Delta p' \gtrsim h.$$

In this measuring operation as in the foregoing one, the measured quantity p_y is itself modified in the course of the measuring process. *This modification must not be confused with the unpredictable and uncontrollable perturbation suffered by the system during the measurement.* Indeed, this modification is known exactly or, at any rate, the values p and p' of p_y before and after collision are known with equal precision; this precision may be made as great as desired. On the other hand, short of being able to predict and control the perturbation suffered by the particle during the measuring operation (uncertainty in the

time at which the momentum and energy are being transferred), one cannot know y once the measurement is carried out, without some uncertainty Δy; the latter is the larger the more precise the measurement of p.

Let us stress again that the uncertainty relations must be of a universal character. Let us assume, for the sake of argument, that the quantum of action for photons is a quantity $\hbar'$ very much smaller than $\hbar$. Clearly, all our arguments concerning measurement (b) could be taken over with $\hbar'$ instead of $\hbar$ in all formulae; this would lead us to a value of order $\hbar'$ for the product of the uncertainties $\Delta y \cdot \Delta p'$ and consequently $\Delta y \cdot \Delta p' \approx \hbar' \ll \hbar$. The uncertainty relations would be violated and the entire statistical interpretation which we have developed would be in contradiction with experiment.

The two measurements of momentum discussed here require a certain delay. This delay can actually be reduced to some extent (by increasing $\mathscr{H}$ in the first case; by increasing ν in the second case; cf. Problem IV.7) without affecting the precision of the measurement of p_y. The probabilities defined in §§ 2 and 5 are calculated from the wave function at a given time t, which ought to be specified. If the measurement is instantaneous, t is the time when it is performed. If the measurement requires some delay, t is the time when it begins, that is, when the system starts interacting with the measuring device. The probability law defined in § 2 [eq. (IV.6)] refers to the momentum p_y, i.e. to the momentum before measurement.

IV. DESCRIPTION OF PHENOMENA IN QUANTUM THEORY. COMPLEMENTARITY AND CAUSALITY

15. Problems Raised by the Statistical Interpretation

Clearly, the representation of a quantum system by its wave function is unfamiliar because of its abstract character, and the statistical interpretation discussed above is difficult to grasp intuitively. However, when one seeks to picture microscopic phenomena in a more concrete fashion, one always runs into contradictions.

For instance, let us consider a helium atom; for simplicity, we assume the helium nucleus to be infinitely heavy and motionless, and we treat the atom as a system of two electrons represented by the wave function $\Psi(\boldsymbol{r}_1, \boldsymbol{r}_2; t)$. The simplest picture one can draw is that of

two corpuscles, the electrons, performing a somewhat complicated motion around the nucleus. This implies that each of them has a well-defined trajectory, whereas the function Ψ does not give more than the statistics of positions in the case of position measurement, and the statistics of momenta in the case of momentum measurement. Since we assume that the dynamical state of the atom is completely defined by the wave function Ψ, this corpuscular representation is partly incorrect. At the other extreme, one can imagine these two electrons to be a continuous distribution of electricity in the space surrounding the nucleus, or preferably, as a continuous wave charged with electricity. But this picture also gives rise to difficulties. In the first place, the wave function Ψ is defined in configuration space and not in ordinary space; it can therefore not be identified with the concrete wave which we are discussing, in fact, no satisfactory definition of such a concrete wave has, as yet, been proposed. In the second place, the picture of a continuous wave cannot be reconciled with such phenomena as ionization, where the existence of individual and, to some extent, localizable corpuscles is exhibited.

Of course, there is no existing proof that a consistent and concrete representation of microscopic phenomena is impossible to formulate. The fact remains, however, that nobody has ever succeeded. In fact logical necessity does not require that the more or less abstract concepts of a physical theory be expressible in concrete language. All our intuition, all our sense of what constitutes concreteness are based upon our everyday experience, and the terms used to describe a phenomenon concretely are necessarily drawn from that experience. There is no indication that such a language could be used without contradictions for phenomena which are as far removed from it as those of microscopic physics. This is not the first example of difficulties of this kind. Some results of the theory of relativity similarly offend our common sense [1]), such as for instance the contraction of lengths and the dilatation of time when the relative velocity of the reference frames is close to c. It is therefore hardly surprising that the concrete pictures one tries to draw of microscopic phenomena cannot be pushed all the way without giving rise to some contradiction or some paradox.

In this connection, it is instructive to note the parallelism between

[1]) The occasions to come to grips between human common sense and the discoveries of physics by no means date from the time of relativity theory ("Eppur' si muove . . ." etc.).

the roles played by the constant $\hbar$ in Quantum Theory, and by the constant c in Relativity Theory. The fact that c is finite rather than infinite imposes a revision of the concept of simultaneity and henceforth sets a limit to the domain of validity of Newtonian Mechanics. In the same way, the fact that $\hbar$ is finite rather than zero imposes a revision of the concept of simultaneous measurements and limits the domain of validity of Classical Theory. The concrete pictures suggested by our everyday experience are those of a world where c seems infinite and $\hbar$ seems to be zero; they cannot bodily be brought over into a domain where one or the other of these approximations ceases to be valid.

The absence of a concrete and coherent representation of phenomena in Quantum Theory may therefore in no way be considered as a shortcoming of the theory. Nevertheless it is open to criticism from another point of view.

The question arises whether the description of phenomena in Quantum Theory actually fulfills all the requirements one should expect from a completely satisfactory theory. The first thing one requires of a theory is, of course, that its predictions be in accord with experimental observations. It is quite certain that Quantum Theory meets this condition, at least in the domain of atomic and molecular physics. But a physical theory cannot claim to be *complete* if it does not go further than to state what one observes when one does a given experiment. At the outset of any scientific endeavor one establishes as a fundamental postulate that nature possesses an objective reality independent of our sensory perceptions or of our means of investigation; the object of physical theory is to give an intelligible account of this objective reality.

Now, all the conclusions of the Quantum Theory can always be put into the following form: "One obtains this or that result if one makes this or that observation". One may therefore question whether Quantum Theory actually furnishes a complete description of objective reality.

The question is even more legitimate since the predictions of the theory are of a statistical nature. In Classical Theory, one has recourse to the language and methods of statistics when the information on the physical systems under study is incomplete. The concepts introduced in Classical Statistics cannot lead to a complete description of objective reality; they only allow us to obtain certain average

properties and some results concerning the physical systems under study, despite incomplete information. These results do not apply to one particular system, but to a very large number $\mathcal{N}$ of identical and independent systems. In the same way, Quantum Theory does not generally yield with certainty the result of a given measurement performed on an individually selected system, but the statistical distribution of the results obtained when one repeats the same measurement on a very large number $\mathcal{N}$ of independent systems represented by the same wave function.

One would thus be tempted to conclude that Quantum Theory furnishes a correct description of statistical distributions of systems of microscopic objects, but that it cannot claim to describe completely each system when taken individually. According to this view, the knowledge of the wave function would not suffice to define completely the dynamical state of an individual physical system. In order to do so, one should have a certain number of additional data which it is impossible to obtain because of the insufficiency of our means of observation. In other words, the dynamical state of the physical system should be defined at each instant by a certain number of *hidden variables* whose evolution would be governed by some specific laws. The impossibility to predict with certainty the results of a given measurement would simply come from our inability to know the precise values of these hidden parameters. The wave function would not represent the objective state of the system under study; rather it would be a mathematical object containing the totality of information which one possesses on an incompletely known system.

Although this opinion is perfectly tenable, the current view holds [1]) that Quantum Theory furnishes a complete description of natural phenomena. This is based upon the analysis, due to Bohr (1927), of the very special conditions of observation on the microscopic scale, and on a general principle which evolves from Bohr's analysis – the complementarity principle.

16. Description of Microscopic Phenomena and Complementarity

Any description of natural phenomena – whose objective reality is by no means questioned here – must inevitably involve at some stage the results of our observations bearing on these phenomena.

[1]) Cf. footnote Ch. II, p. 48.

Now – and this is the first point of Bohr's analysis – *no matter how far the phenomena transcend the scope of Classical Physics, their account must be expressed in classical terms.* Indeed, to account for an experiment means to give an unambiguous description of the circumstances of the experiment and of the observed results; it means to state, for instance, that "this pointer has stopped on this dial at that point and at that moment". The point we wish to emphasize here is the necessity of using unambiguous language, in which no element of uncertainty on the part of the observer may enter. This is absolutely indispensable, since the experiment must be reproducible, and its progress must remain completely independent of the observer who performs it.

However, *on the microscopic level, one cannot make the sharp separation required by the ordinary concept of observation, between the natural phenomenon and the instrument with which it is observed.* To describe the object and the observing instrument as separate entities is justified only to the extent where the quantum $\hbar$ may be considered negligible. This sets a limit to the analysis of phenomena, when carried out in classical language; any attempt to push the analysis beyond this limit requires a modification of the experimental arrangement which introduces new possibilities of interaction between object and measuring instruments.

As a consequence, *evidences obtained under different experimental conditions cannot be comprehended within a single picture.* However, they must be regarded as *complementary* in the sense that only the totality of the observational results exhausts the possible types of information about the objects of microscopic physics. Such is the content of the complementarity principle.

17. Complementary Variables. Compatible Variables

The description of the phenomena in microscopic physics is thus made up of complementary elements; these elements which are simultaneously needed in any attempt to build a classical picture, are respectively defined by means of mutually exclusive experimental arrangements.

A given component of the position x and the corresponding momentum component p form a pair of complementary elements in the above sense. The precise measurement of x and that of p require incompatible experimental arrangements, so that the simultaneous

measurement of these two quantities cannot be carried beyond the limit of precision $\Delta x \cdot \Delta p \approx \hbar$. This is what follows from the discussions of §§ 13 and 14. The mutually exclusive character of the experimental arrangements designed for measuring x and p is very simply exhibited when observing an electron by means of a diaphragm (§ 13*a*). The diaphragm used for the position measurement can serve equally well for the momentum measurement. The latter is obtained by measuring the momentum transferred to the screen through its interaction with the electron. But since electron + screen form an indivisible quantum system, and since it is impossible to make a sharp separation between its two parts, one cannot use ordinary classical language and speak of a separate evolution of the electron (measured system) and the screen (measuring instrument) unless one assumes that their mutual interaction is to some extent uncontrollable. If one wishes to know x with precision Δx, one cannot control the momentum transferred to the screen with a precision greater than $h/\Delta x$. If one wants to know p with precision δp, adding to the screen all the experimental devices necessary for a precise determination of the momentum transfer, one cannot control the position of the screen with a precision greater than $h/\delta p$, and this limits our knowledge of x correspondingly (cf. Problem IV.10).

The variables x and p are said to form a pair of *complementary variables*. One often states the complementarity principle in the following more restrictive form:

The description of the physical properties of microscopic objects in classical language requires pairs of complementary variables; the accuracy in one member of the pair cannot be improved without a corresponding loss in the accuracy of the other member.

This statement stresses the essential point of difference between Quantum Mechanics and Classical Mechanics, namely that the dynamical variables of a quantum system cannot all be defined simultaneously with infinite accuracy.

By definition, two dynamical variables are said to be *compatible* if they can simultaneously be defined with infinite precision. As will be shown in Chapter V, two compatible variables are represented by commuting linear operators. Thus the coordinates x and y of a particle are compatible variables. Of particular importance are the *complete*

sets of compatible variables a, b, c, ..., composed of pairwise compatible variables, and such that any other variable compatible with each of them is a function $f(a, b, c, \ldots)$ of these variables. Consider, for example, a one-particle quantum system. The three position variables x, y, z form a complete set of compatible variables. Indeed, all three can be defined simultaneously with infinite precision. On the other hand, any dynamical variable is a function of x, y, z, and p_x, p_y, p_z; the only ones compatible with x, y, and z are those which are independent of p_x, p_y, and p_z, since each of the latter quantities is incompatible with x, y, and z, respectively; hence they are the functions $F(x, y, z)$. In the same way p_x, p_y, and p_z form a complete set; similarly p_x, y, and p_z (the specification of these three variables completely defined the dynamical state of the system before measurement in the examples of § 14). A precise measurement carried out on a complete set of compatible variables of a system represents the *maximum information* one can obtain on that system. It therefore defines the dynamical state completely, and a definite wave function corresponds to it. We shall come back to this point in Chapter V (§§ 15 and 16).

18. Wave-Corpuscle Duality and Complementarity

If one adopts the principle of complementarity, the wave-corpuscle duality ceases to be paradoxical: the wave aspect and the corpuscular aspect are two complementary aspects which are exhibited only in mutually exclusive experimental arrangements. Any attempt to reveal one of the two aspects requires a modification of the experimental set-up which destroys any possibility of observing the other aspect.

Consider for instance, Young's diffraction experiment. The monochromatic radiation originating from the source S crosses the screen Y through two openings separated by d, and then forms the interference pattern on the screen E placed at a distance D from Y (Fig. IV.6). The distance between fringes is $\lambda D/d$, λ being the wavelength of the radiation. In this experiment, the source S and the screens Y and E are rigidly mounted on a common support. In fact the fringes cannot be observed unless the position of Y with respect to S and E is controlled to a tolerance δx smaller than the fringe spacing:

$$\delta x < \frac{\lambda D}{d}.$$

(For the sake of simplicity the source S is assumed to be at infinity.) The corpuscular aspect of the radiation is displayed if one can find out through which hole of the screen Y the particle has passed. This

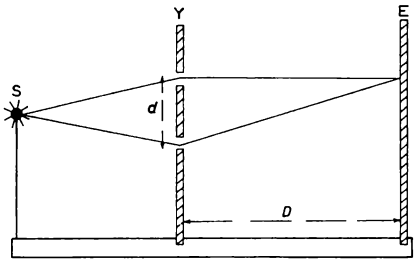


Fig. IV.6. Complementarity between the corpuscular aspect and the wave aspect in the Young double-slit experiment. S = source; Y = slotted screen; E = detection screen. In the figure, S, Y and E are rigidly attached to a common support; the experimental arrangement allows the display of the interference pattern, but all information concerning the trajectory followed by each photon escapes detection by the same token.

may be done most simply by measuring the momentum transferred to the screen Y; it differs by $\approx pd/D$ according to whether the corpuscle passed through one or the other opening, $p=h/\lambda$ denoting the momentum of the corpuscle. The measurement must therefore be made to a precision such that

$$\delta p < \frac{hd}{\lambda D}.$$

In order to reveal simultaneously the wave aspect and the corpuscular aspect, both of these inequalities must be satisfied simultaneously, i.e. one must have

$$\delta x \cdot \delta p < h,$$

which is clearly impossible since the observing instrument is a quantum object and satisfies the uncertainty relations. Any other attempt to reveal the corpuscular aspect would similarly result in a suppression of the interference fringes (cf. Problem IV.11).

19. Complementarity and Causality

The discussion of phenomena in the framework of Quantum Theory implies some definite restriction on the ordinary principle of causality.

Causality, rigorously, applies only to isolated systems. The dynamical state of such a system is represented at a given instant of time by its wave function at that instant. The causal relationship between the wave function $\psi(t_0)$ at an initial time t_0, and the wave function $\psi(t)$ at any later time, is expressed through the Schrödinger equation. However, as soon as it is subjected to observation, the system experiences some reaction from the observing instrument. Moreover, the above reaction is to some extent unpredictable and uncontrollable since there is no sharp separation between the observed system and the observing instrument. They must be treated as an indivisible quantum system whose wave function $\Psi(t)$ depends upon the coordinates of the measuring device as well as upon those of the observed system. During the process of observation, the measured system can no longer be considered *separately* and the very notion of a dynamical state defined by the simpler wave function $\psi(t)$ loses its meaning. Thus the intervention of the observing instrument destroys all causal connection between the state of the system before and after the measurement; this explains why one cannot in general predict with certainty in what state the system will be found after the measurement; one can only make predictions of a statistical nature [1]).

In order to show how the notions which enter into the description of a phenomenon are inseparable from the manner in which it is observed, and what limitations are thereby imposed on causality, we shall take the example of an atom which de-excites through emission of a photon. This example will at the same time illustrate complementarity between wave and corpuscular aspects.

The atom is assumed to be in its excited state at time $t=0$; call τ its mean life, and $\hbar\omega$ the energy of the emitted photon. We wish to

[1]) The statistical predictions concerning the results of measurement are derived very naturally from the study of the mechanism of the measuring operation itself, a study in which the measuring instrument is treated as a quantized object and the complex (system + measuring instrument) evolves in causal fashion in accordance with the Schrödinger equation. A very concise and simple presentation of the measuring process in Quantum Mechanics is given in F. London and E. Bauer, *La Théorie de l'Observation en Mécanique Quantique* (Paris, Hermann, 1939). More detailed discussions of this problem may be found in J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton, Princeton University Press, 1955), and in D. Bohm, *Quantum Theory* (New York, Prentice-Hall, 1951).

make a precise measurement of the time of photon emission. To this effect, we completely surround the atom with suitable counters; one of these counters will eventually be struck by the photon. Knowing the time at which the count took place, the distance separating the atom from the counter, and the velocity of propagation of light c, one easily calculates the moment at which the emission occurred. However, the predictions of Quantum Theory are only statistical in nature. Indeed, consider the system (atom + photon) at time $t \gg \tau$: the atom is then with near certainty in its ground state, and the photon is to a very good approximation – we shall not give the proof here – represented by a wave packet $\psi(\mathbf{r}, t)$ whose dependence upon t and r is of the form

$$f(r,t) = \begin{cases} 0 & \text{when } r > ct \\ \dfrac{1}{r} \exp\left[-\left(i\omega + \dfrac{1}{2\tau}\right)\left(t - \dfrac{r}{c}\right)\right] & \text{when } r < ct \end{cases} \tag{IV.34}$$

(r is the distance between the photon and the atom).

The probability of detecting a photon at a distance R from the atom is zero as long as $t < R/c$, and proportional to $\exp[-(t-R/c)/\tau]$ at any later time; hence, we have the well-known exponential law giving the statistical distribution of the time of decay.

This law actually fits experimental observations on a very large number of decaying atoms. However, Quantum Theory is unable to predict the time of decay of each individual atom. One is therefore tempted to conclude that the Quantum Theory describes correctly the de-excitation of statistical ensembles of excited atoms, but that it does not describe completely the phenomenon of de-excitation of a single atom.

The reply to this objection is as follows. The notion of decay time cannot be separated from the experimental arrangement which serves to reveal it, and must not be regarded as a property characterizing the evolution of the atom under consideration, independently of this arrangement. In fact, there exist other experimental arrangements, complementary to the one just considered, through which aspects of the decay phenomenon are revealed which are totally incompatible with the very existence of a precise time of decay. These are all the arrangements devised for the observation of the interference effects of the emitted light. Through those arrangements, the wave (IV.34) is split into two wave packets which are later made to recombine after

having traveled along different optical paths. Obviously, the above wave packets cannot interfere unless the difference between the two optical paths is smaller than the spatial extension of the wave (IV.34) [1]). In such phenomena, the spatial extension of the wave is clearly exhibited and the very notion of a decay time is meaningless.

As far as causality is concerned, it actually applies to the system (atom + photon) as long as the latter has not yet interacted with the observing instrument. During that entire period, the description of the dynamical state of this system by means of a wave function obeying the Schrödinger equation remains possible; in particular, the photon is represented by the packet of outgoing waves (IV.34) whose wave front ($r = ct$) recedes from the atom with velocity c. However, this causal description loses validity as soon as the system (atom + photon) starts to interact with the observing instrument. One can no longer consider the system as a separate entity since the ensemble (system + observing instrument) forms an indivisible whole. The properties commonly attributed to the system are in fact properties of the complex. Therefore, no strict causal relationship can exist between the state of the system before the measurement and its state afterwards.

This situation differs very profoundly from the strict causality of Classical Theory; the dynamical variables of a classical system are all defined at each instant in a precise manner and their evolution in time is strictly determined. Those of a quantum system are only defined in the limits of approximation required by the uncertainty relations between pairs of complementary variables, and their evolution in time is only partially determined. Surprising as it might appear, this limitation on the causality principle does not run the risk of contradicting any experimental fact, since the evolution in time of the ensemble of dynamical variables of the system can be observed experimentally only within the limits of approximation required by the uncertainty relations.

---

[1]) The lifetime of an excited state emitting visible light is usually of the order of 10^{-8} sec. Consequently, the length of the emitted wave train $c\tau$ is of the order of 3 m. Interference patterns corresponding to optical path differences of the order of a meter have actually been observed. This confirms the conclusions which we are drawing here on the spatial extension of the wave train, and on the ambiguity of the notion of decay time.

EXERCISES AND PROBLEMS

1. Show that the Hamiltonian of a particle in an electromagnetic field [eq. (II.25)] is Hermitean and that, consequently, the normalization of an arbitrary solution Ψ of the Schrödinger equation is conserved in the course of time. Show that one can write a continuity equation in that case which is identical to eq. (IV.11) if one takes as definition of the current the one suggested by correspondence with Classical Mechanics, namely

$$\boldsymbol{J} = \mathrm{Re}\left[\Psi^* \left(\frac{\hbar}{i n} \nabla - \frac{e}{mc} \boldsymbol{A}\right) \Psi\right]$$

($\boldsymbol{A}$ = vector potential of the electromagnetic field).

2. A particle in one dimension of mass m moves in a potential $V(x)$ which goes asymptotically (more rapidly than $1/|x|$) to V_+ and V_- when x approaches $+\infty$ and $-\infty$, respectively. Introduce the wave numbers $k_\pm = \sqrt{2m(E - V_\pm)}/\hbar$ and consider the solutions $u_E(x)$ whose asymptotic behavior at the two ends of the x axis is

$$u_E \underset{x \to -\infty}{\sim} e^{ik_- x} + R\, e^{-ik_- x}$$

$$u_E \underset{x \to +\infty}{\sim} S\, e^{ik_+ x}.$$

By superposition of solutions of this type, form a wave packet (normalized to unity) representing a particle moving in the direction of increasing x. Show that after a sufficiently long time this packet splits into a transmitted wave packet and a reflected wave packet, and that the probabilities of finding the particle in the one or the other of these packets are equal to $(k_+/k_-)|S|^2$ and $|R|^2$, respectively.

3. The wave function associated with a particle constrained to move along the x axis is

$$\psi(x) = (2\pi\xi^2)^{-\frac{1}{4}} \exp\left(\frac{i}{\hbar} p_0 x - \frac{x^2}{4\xi^2}\right).$$

Calculate the wave function $\varphi(p)$ of momentum space. Verify that $\psi(x)$ and $\varphi(p)$ are normalized to unity. Calculate the average values of x and x^2 by successively applying the expressions (IV.13) and (IV.21), and compare the results. Calculate the mean values of p, p^2, and $\exp(ipX/\hbar)$ (X is a given real constant) by successively applying expressions (IV.14) and (IV.20), and compare the results.

4. Show that the wave packet for which $\Delta x \cdot \Delta p = \hbar/2$ (minimum wave packet) is necessarily of the form

$$(2\pi\xi^2)^{-\frac{1}{4}} \exp\left[\frac{i}{\hbar} p_0 x - \frac{(x - x_0)^2}{4\xi^2}\right]$$

and that one has $\langle x \rangle = x_0$, $\langle p \rangle = p_0$, $\Delta x = \xi$, $\Delta p = \hbar/2\xi$.

5. Let $\langle x \rangle$ and $\langle p \rangle$ be the mean values of x and its conjugate momentum p, for a system in the dynamical state $\psi(x)$. Show that the mean values of x and p for the dynamical state

$$\exp\left(-\frac{i}{\hbar}\langle p \rangle x\right)\psi(x + \langle x \rangle)$$

vanish.

6. Show that in the position measurement of § 13*b*, the quantum nature of the measuring device (light beam + microscope) plays an essential role, and that the uncertainty relations would be violated if the light were not quantized. Show, in this example, that the constant h which enters in the uncertainty relations for the electron is necessarily no greater than the constant which occurs in the definition of light quanta.

7. In the momentum measurement by deflection in a magnetic field (Ch. IV, § 14*a*), show that one can decrease to a certain extent the duration of the measurement $\tau = \pi mc/e\mathscr{H}$ without changing the precision of the measurement, but that the precision $\Delta\tau$ with which this duration of measurement is defined, remains unchanged. Discuss the time-energy uncertainty relation in this example. Elaborate on the same questions bearing on the duration $1/\Delta\nu'$, and the measurement of momentum by collision with a photon (§ 14*b*).

8. The momentum measurement of § 14*b* may be undertaken by using any kind of particle in the place of photons. It consists in making the electron undergo a collision with that particle, and to measure the momentum transferred to the latter in the collision. Discuss the momentum-position uncertainty relations in this case in the non-relativistic approximation. Show that the uncertainty introduced by the measurement of the position of the electron depends upon the constant $\hbar$ occurring in the uncertainty relations belonging to the particle used.

9. A quantized particle is made to pass through an opening of diameter d equipped with a shutter which opens for a time τ. Show that the particle necessarily exchanges with this device (diaphragm + shutter) a momentum of the order of $\hbar/d$ and an energy of the order of $\hbar/\tau$.

10. In the position measurement of an electron by means of a microscope (§ 13*b*) one seeks to determine the momentum imparted to the electron by carrying out a precise measurement of the momentum transferred to the microscope in the course of the operation. Show that one can improve the precision of this momentum measurement only at the expense of a decrease in precision in the position measurement, in accord with the uncertainty relations.

11. One performs the experiment of Young's double slit on a monoërgic beam of charged particles and one seeks to reveal the interference fringes by using a cloud chamber as a detector. This is possible provided the chamber is sufficiently far from Young's screen. Show that the observation of the "trajectory" followed by each particle in the cloud chamber is not accurate enough to allow us to decide through which slit the particle has traveled. (Cf. footnote p. 142.)

CHAPTER V

DEVELOPMENT OF THE FORMALISM OF WAVE MECHANICS AND ITS INTERPRETATION

1. Introduction

In the preceding chapter we have given the basic elements of the statistical interpretation of Quantum Theory; taking a very general point of view, we have examined the internal consistency of that interpretation, its compatibility with the experimental facts and the very special conditions it implied for the description of natural phenomena. Restricting ourselves henceforth to the narrower framework of the Wave Mechanics of systems of particles (in the non-relativistic approximation), we shall complete and refine this interpretation in accordance with the general principles we previously established.

Thus, we postulate that the wave function completely defines the dynamical state of the system under consideration. In contrast to what occurs in Classical Theory, the dynamical variables of the system cannot in general be defined at each instant with infinite precision. However, if one performs the measurement of a given dynamical variable, the results of measurement follow a certain probability law, and that law must be completely determined upon specifying the wave function.

The probability laws for position and momentum measurements were given explicitly in Chapter IV, Section I; we have also derived a general formula for the mean values of any function of the position coordinates, and of any function of the momentum coordinates. However, the rule for deducing the probability law from the wave function has not been given in the most general case, namely in the case where the measured dynamical variable is a function of *both* position *and* momentum coordinates. We shall correct this flaw in the first three sections of this chapter. The fundamental postulates are presented in Section I. With every dynamical variable $\mathscr{A}$ one associates a certain Hermitean operator A acting on the wave functions, and the mean value of the results of measurements of $\mathscr{A}$ is an expression formed in terms of A, generalizing those of Chapter IV

(§ 5). Assuming this postulate to be valid for any dynamical variable and any function of these variables, the statistical distribution sought is exactly determined. Its explicit determination is closely related to the solution of the eigenvalue problem of A. In Section II, we study the properties of the eigenvalues and eigenfunctions of A in the special case where the spectrum is discrete, and the eigenfunctions are square-integrable. The probability law turns out to be closely related to the coefficients of the expansion of the wave function in a series of eigenfunctions of A. In Section III, the same problem is treated for the general case where the eigenvalue spectrum is partly continuous.

In Section IV, we formally treat the problem of deducing the wave function of a quantum system from the results obtained when carrying out simultaneous measurements of a complete set of compatible variables. When this "maximum observation" is not realized, the information concerning the dynamical state of the system is incomplete; the study of the system can nevertheless be pursued by the methods of statistics, the term statistics being taken here in its usual sense.

With two compatible variables there are associated commuting operators. If the commutators of all pairs of operators were to vanish all the dynamical variables of the system could be simultaneously defined with arbitrary precision. In contrast with this situation characteristic of Classical Theory, some pairs of variables are incompatible and the corresponding commutators differ from zero. The commutators therefore play an essential role in Quantum Theory. Section V is devoted to the study of commutators, to the calculation of the most common commutators, and to the establishment and discussion of some equations in which the commutator concept turns out to be particularly useful.

I. HERMITEAN OPERATORS AND PHYSICAL QUANTITIES

All examples needed for illustrative purposes in the following will be taken from the wave mechanics of one-dimensional systems (Ch. III) or three-dimensional systems (that is, one-particle systems). However, one should never lose sight of the fact that these considerations are very generally applicable to quantum systems with any number of dimensions.

2. Wave-Function Space

The wave functions capable of representing a given quantum system belong to a function space which should be specified precisely. In order that the distributions $P(\boldsymbol{r})$ and $\Pi(\boldsymbol{p})$ defined in Ch. IV, § 2 have a meaning, it is necessary and sufficient that the normalization condition (IV.3) could be applied to the wave function $\psi(\boldsymbol{r})$. We are thus led to the following definition of wave-function space:

The wave functions of wave mechanics are the square-integrable functions of configuration space, that is to say the functions $\psi(q_1, \ldots, q_R)$ such that the integral $\int|\psi(q_1, \ldots, q_R)|^2 \, \mathrm{d}\tau$ converges [1]) [$\mathrm{d}\tau$ denotes the volume element $\mathrm{d}q_1 \, \mathrm{d}q_2 \ldots \mathrm{d}q_R$].

We could restrict the function space somewhat more by requiring the wave functions to be normalized to unity [eq. (IV.3)]. However, it turns out to be more convenient to relax this normalization condition; this can be done, as we shall see below, at the price of a slight modification in the definition of the statistical distributions and probabilities.

In the language of mathematics, the function space defined above is a *Hilbert space*. It possesses indeed the properties characteristic of such a space, as shown below.

In the first place, it is a *linear space*. If ψ_1 and ψ_2 are two square-integrable functions, their sum, the product of each by a complex number and, more generally, any linear combinations

$$\lambda_1\psi_1 + \lambda_2\psi_2,$$

where λ_1 and λ_2 are arbitrarily chosen complex numbers, are also square-integrable functions.

In the second place, one can define a *scalar product* in that space. By definition, the scalar product of the function ψ by the function φ is

$$\langle\varphi, \psi\rangle \equiv \int \varphi^*(q_1, \ldots, q_R) \, \psi(q_1, \ldots, q_R) \, \mathrm{d}\tau. \qquad (\mathrm{V}.1)$$

---

[1]) The Fourier transform $\varphi(p_1, \ldots, p_R)$ of such a function always exists: it is a square-integrable function possessing the same normalization as $\psi(q_1, \ldots, q_R)$. In fact (cf. Appendix A) the correspondence between φ and ψ is one-to-one if one adopts the convention not to consider as distinct two functions which differ only by an ensemble of measure zero, which will always be done.

If it is zero, the functions φ and ψ are said to be *orthogonal*. The *norm* N_ψ of a function ψ is the scalar product of this function by itself:

$$N_\psi \equiv \langle \psi, \psi \rangle.$$

The fundamental properties of the scalar product are as follows:

a) the scalar product of φ by ψ is the complex conjugate of the scalar product of ψ by φ, namely

$$\langle \psi, \varphi \rangle = \langle \varphi, \psi \rangle^*; \tag{V.2}$$

b) the scalar product of ψ by φ is linear with respect to ψ, in other words

$$\langle \varphi, \lambda_1\psi_1 + \lambda_2\psi_2 \rangle = \lambda_1 \langle \varphi, \psi_1 \rangle + \lambda_2 \langle \varphi, \psi_2 \rangle; \tag{V.3}$$

c) the norm of a function ψ is a real, non-negative number:

$$\langle \psi, \psi \rangle \geqslant 0 \tag{V.4}$$

and if $\langle \psi, \psi \rangle = 0$, we have necessarily [1]) $\psi = 0$.

All the above properties are easily deduced from the very definition of the scalar product itself. From properties (*a*) and (*b*) one easily shows that the scalar product $\langle \varphi, \psi \rangle$ does not depend linearly, but antilinearly on φ:

$$\langle \lambda_1\varphi_1 + \lambda_2\varphi_2, \psi \rangle = \lambda_1^* \langle \varphi_1, \psi \rangle + \lambda_2^* \langle \varphi_2, \psi \rangle. \tag{V.3'}$$

From the properties (*a*), (*b*), and (*c*) follows a very important property of the scalar product, the *Schwarz inequality* (cf. Problem V.1):

$$|\langle \varphi, \psi \rangle| \leqslant \sqrt{\langle \varphi, \varphi \rangle \langle \psi, \psi \rangle}. \tag{V.5}$$

Equality obtains when the functions φ and ψ are multiples of each other, and only in that case. The Schwarz inequality insures that the integral (V.1) defining the scalar product converges when the functions φ and ψ are square-integrable functions.

In addition to the fact that it is linear and that one can define a scalar product there, the space of square-integrable functions possesses the property of being *complete*; this is what allows us to

---

[1]) Rigorously, ψ can take on values different from zero on an ensemble of measure zero. Following the conventions of the preceding footnote, such a function is not different from zero.

identify it as a Hilbert space. To be complete means that any set of square-integrable functions satisfying the Cauchy criterion, converges (in the quadratic mean) toward a square-integrable function. Conversely, any square-integrable function can be considered as the limit (in the quadratic mean) of a converging series (in the sense of Cauchy) of square-integrable functions (*separability*) [1]).

3. Definition of Mean Values

In Chapter IV, § 5 we were led to associate with any dynamical var. iable of the type $F(\boldsymbol{r})$ or of the type $G(\boldsymbol{p})$ some linear operator A equal to one of the expressions $F(\boldsymbol{r})$ or $G(-i\hbar\nabla)$; moreover — and this resulted from the definitions adopted for $P(\boldsymbol{r})$ and $\Pi(\boldsymbol{p})$ — the mean value of this dynamical variable was given by expression (IV.22) which, with our present notation, can just as well be written $\langle \Psi, A\Psi \rangle$. If the wave function is not normalized to unity, the expressions (IV.2) and (IV.6) for $P(\boldsymbol{r})$ and $\Pi(\boldsymbol{p})$ must be divided by the norm $\langle \Psi, \Psi \rangle$ of the wave function, and the expression for the mean value considered above must be replaced by the expression $\langle \Psi, A\Psi \rangle / \langle \Psi, \Psi \rangle$.

Generalizing this result to any dynamical variable, we are therefore led to *postulate* that:

a) with the dynamical variable $\mathscr{A} = A(q_1, \ldots, q_R; p_1, \ldots, p_R)$ is associated the linear operator

$$A\left(q_1, \ldots, q_R; \frac{\hbar}{i}\frac{\partial}{\partial q_1}, \ldots, \frac{\hbar}{i}\frac{\partial}{\partial q_R}\right).$$

b) the mean value of this dynamical variable when the system is in the dynamical state defined by the function $\Psi(q_1, \ldots, q_R)$, is

$$\langle A \rangle = \frac{\langle \Psi, A\Psi \rangle}{\langle \Psi, \Psi \rangle}. \tag{V.6}$$

The correspondence between the classical Hamiltonian function and the Schrödinger Hamiltonian (Ch. II, § 15) is a special case of the

[1]) For a rigorous and detailed study of Hilbert space, see M. H. Stone, *Linear Transformations in Hilbert Space* (New York, Amer. Math. Soc., 1932). The main properties are given together with their proof in J. von Neumann, *loc. cit.* For the definition of convergence in the quadratic mean, see Appendix A, footnote p. 475.

correspondence (*a*) between dynamical variables and linear operators. The restrictions made in Ch. II, § 15 concerning this correspondence equally apply to the general case. Let us recall that the q's must be cartesian coordinates. Furthermore there exists a certain ambiguity in the definition of the operator A due to the fact that one replaces in the correspondence operation, quantities of ordinary algebra by operators which might not commute. In practice, one removes this ambiguity by confirming to the empirical rules of Chapter II, § 15.

Furthermore, if the dynamical variable $\mathscr{A}$ represents a physical quantity, it is a real function of the q's and p's, and the results of measurements of $\mathscr{A}$, and *a fortiori* the average value $\langle A \rangle$, are real quantities. Hence $\langle \Psi, A\Psi \rangle$ is real

$$\langle \Psi, A\Psi \rangle = \langle A\Psi, \Psi \rangle. \tag{V.7}$$

and this is so for any dynamical state of the system to which the measurement is applied, hence for any function Ψ. To put it differently (cf. Ch. IV, § 3), *the operator A must be Hermitean.* It is easy to see – and will be readily verified for all further examples – that as long as the operators q_i and $(\hbar/i)\partial/\partial q_i$ are Hermitean, one can always associate by the correspondence rule (*a*) a Hermitean operator with any real dynamical variable. It is precisely with this in mind that the prescription for "symmetrization" indicated in Chapter II, § 15 was given.

The properties of Hermitean operators will be studied systematically in § 5 and following sections. Let us anticipate here an important property of these operators. If A is Hermitean, the mean value of $\mathscr{A}$ calculated for a linear combination $\Phi + \lambda\Psi$ of two functions Φ and Ψ of the function space to which A is applied, is a real quantity. Consequently the quantity

$$\langle \Phi + \lambda\Psi, A(\Phi + \lambda\Psi) \rangle = \langle \Phi, A\Phi \rangle + \lambda \langle \Phi, A\Psi \rangle + \lambda^* \langle \Psi, A\Phi \rangle + |\lambda|^2 \langle \Psi, A\Psi \rangle$$

is real. This must be true for any value of the complex number λ. Since $\langle \Phi, A\Phi \rangle$ and $\langle \Psi, A\Psi \rangle$ are real quantities, we conclude therefrom, denoting the phase of λ by α, that

$$e^{i\alpha} \langle \Phi, A\Psi \rangle + e^{-i\alpha} \langle \Psi, A\Phi \rangle$$

is real, or stated differently that

$$e^{i\alpha}(\langle \Phi, A\Psi \rangle - \langle A\Phi, \Psi \rangle) = e^{-i\alpha}(\langle A\Psi, \Phi \rangle - \langle \Psi, A\Phi \rangle).$$

Since this equation must hold for any value of α, the quantities

between parentheses on both sides of the equation must vanish. In other words, if Ψ and Φ are any two functions of the function space in which the Hermitean operator A acts,

$$\langle \Phi, A\Psi \rangle = \langle A\Phi, \Psi \rangle \tag{V.8}$$

or, what is equivalent,

$$\langle \Phi, A\Psi \rangle = \langle \Psi, A\Phi \rangle^*. \tag{V.8'}$$

Equation (V.8') is often taken as the definition of hermiticity.

From the double postulate (a) and (b) one can deduce the statistical distribution of the values of the physical quantity $\mathscr{A}$. The end of this section and the two following sections are devoted to this problem.

4. Absence of Fluctuations and the Eigenvalue Problem

The fluctuations of the statistical distribution we are looking for are expressed by the root-mean-square deviation ΔA:

$$(\Delta A)^2 = \langle (A - \langle A \rangle)^2 \rangle = \langle A^2 \rangle - \langle A \rangle^2 \qquad (\geqslant 0).$$

A^2 being a dynamical variable by the same token as A, its mean value is given by postulate (b). When the deviation ΔA is zero, the fluctuations are absent and one may assert with certainty that $\mathscr{A}$ takes on a well-defined value; the latter is obviously equal to $\langle A \rangle$.

Let us see what condition the relation $\Delta A = 0$ imposes upon the function Ψ. By applying the definition (V.6) to the mean values of the operators A and A^2, one can put this relation into the form

$$\langle \Psi, A^2\Psi \rangle \langle \Psi, \Psi \rangle = \langle \Psi, A\Psi \rangle^2.$$

However, the quantity $\langle \Psi, A^2\Psi \rangle \equiv \langle \Psi, A(A\Psi) \rangle$ is equal to $\langle A\Psi, A\Psi \rangle$, as can easily be shown by applying property (V.8) of the Hermitean operator A to the functions Ψ and $A\Psi$. We therefore have

$$\langle \Psi, A\Psi \rangle^2 = \langle \Psi, \Psi \rangle \langle A\Psi, A\Psi \rangle.$$

We have the case where the Schwarz inequality reduces to an equality. Consequently the functions Ψ and $A\Psi$ are proportional. The fluctuations of the statistical distribution of $\mathscr{A}$ vanish for the dynamical states Ψ_a such that

$$A\Psi_a = a\Psi_a, \tag{V.9}$$

where a is a constant.

Equation (V.9) is an eigenvalue equation; the time-independent Schrödinger equation has already furnished an example of equations of this type. We have thus reached a first conclusion:

The physical quantity $\mathscr{A}$ possesses with certainty (that is to say with probability equal to 1) *a well-defined value if, and only if, the dynamical state of the physical system is represented by an eigenfunction Ψ_a of the Hermitean operator associated with $\mathscr{A}$, and the value assumed by this quantity is the eigenvalue a associated with that function.*

What has just been stated must hold true in particular for the energy $H(q_i; p_i)$ of the system. The corresponding operator, the Schrödinger Hamiltonian, was actually associated with the energy of the system when we established the time-independent Schrödinger equation (Ch. II, Sec. III). We assumed then that the energy of the system takes a well-defined value E when the system is in a stationary state, and that its wave function is an eigenfunction of the operator H corresponding to the eigenvalue E. These assumptions agree with the general postulates (*a*) and (*b*) introduced in the previous section.

The arguments which led to equation (V.9) assume that the functions $\Psi, A\Psi, A^2\Psi$ which occur in the various scalar products belong to Hilbert space. Eq. (V.9) itself is an equation in which, until stated otherwise, the unknown function Ψ_a must be square-integrable.

Stated in this way, the eigenvalue problem may very well have no solution. Such is the case of operators as common as q_i and $(\hbar/\mathrm{i})\partial/\partial q_i$. Consider more closely the eigenvalue problem for these two operators in the case of one-dimensional systems. For q, we write

$$(q-q')\,\psi(q)=0,$$

which is possible only if $\psi(q)$ is zero everywhere except when $q=q'$. Not only is there no square-integrable solution satisfying such conditions, but the only possible solution is a particularly strange function, zero everywhere except at one point. We shall discuss this matter more fully in § 8. For $(\hbar/\mathrm{i})\mathrm{d}/\mathrm{d}q$ the eigenvalue problem is written

$$\frac{\hbar}{\mathrm{i}}\frac{\mathrm{d}}{\mathrm{d}q}\psi(q)=p'\psi(q).$$

It has a solution, defined to within a constant, for any value of p':

the function $\exp(ip'q/\hbar)$. This eigenfunction is not square-integrable.

For the two preceding operators the result stated above is useless. Clearly, in order to arrive at a statement of general interest, one must also consider the solutions of the eigenvalue problem (V.9) which are not square-integrable. We have already dealt with a particular eigenvalue equation, the Schrödinger equation of one-dimensional systems (cf. Ch. III), and we can be guided by the results of that study in arriving at a general statement. The eigenvalue spectrum of the Schrödinger Hamiltonian generally consists of two parts: a series of discrete eigenvalues whose eigenfunctions have a finite norm, and a continuous spectrum of eigenvalues whose eigenfunctions are bounded in all space but have infinite norm. By superposition of eigenvalues of the continuous spectrum corresponding to neighboring values of the energy, it is intuitively clear that one can form square-integrable functions whose energy is determined, if not in precise fashion, at any rate with a root-mean-square deviation which can be made as small as desired.

Let us elaborate on this point by taking up the notation of Chapter III, § 13. Starting from the eigenfunction $y(x;\varepsilon)$ corresponding to the eigenvalue ε of the continuous spectrum of H:

$$Hy(x;\varepsilon)=\varepsilon y(x;\varepsilon) \tag{V.10}$$

one constructs the "eigendifferential" [1]):

$$Y_\varepsilon(x;\delta\varepsilon)=(\delta\varepsilon)^{-\frac{1}{2}}\int_{\varepsilon}^{\varepsilon+\delta\varepsilon} y(x;\varepsilon')\,d\varepsilon'. \tag{V.11}$$

The latter is a square-integrable function and the quantities $\langle H\rangle$ and ΔH calculated with such a wave function have a well-defined meaning. Using eqs. (V.10) and (V.11) it is easy to see that

$$\langle H\rangle=\frac{\langle Y_\varepsilon, HY_\varepsilon\rangle}{\langle Y_\varepsilon, Y_\varepsilon\rangle}\underset{\delta\varepsilon\to 0}{\sim}\varepsilon+O(\delta\varepsilon).$$

$$\langle H^2\rangle=\frac{\langle Y_\varepsilon, H^2Y_\varepsilon\rangle}{\langle Y_\varepsilon, Y_\varepsilon\rangle}\underset{\delta\varepsilon\to 0}{\sim}\varepsilon^2+\varepsilon\cdot O(\delta\varepsilon).$$

Therefore, the root-mean-square deviation $\Delta H\equiv\sqrt{\langle H^2\rangle-\langle H\rangle^2}$ tends

---

[1]) The factor $(\delta\varepsilon)^{-\frac{1}{2}}$ was introduced in the definition of the "eigendifferential" in order that its normalization remain finite in the limit where $\delta\varepsilon\to 0$.

to zero like $(\varepsilon\delta\varepsilon)^{\frac{1}{2}}$ when $\delta\varepsilon$ tends to zero. It can be made arbitrarily small. This result can be expressed in the following manner: by superposition of wave functions (of infinite norm) corresponding to eigenvalues located in a restricted region $(\alpha, \alpha+\delta\alpha)$ of the continuous spectrum of A (if it exists), one can form square-integrable functions, and the root-mean-square deviation of the distribution of the values of A about its mean value $[\approx \alpha+O(\delta\alpha)]$ may be made as small as desired by choosing the size $\delta\alpha$ of the region sufficiently small.

In conclusion, it is clear that the eigenvalue problem defined by equation (V.9) must play a fundamental role not only in the domain of the discrete spectrum, but also in the domain of the continuous spectrum, domain in which the eigenfunctions no longer belong to Hilbert space. We shall now make a systematic study of this eigenvalue problem.

II. STUDY OF THE DISCRETE SPECTRUM

5. Eigenvalues and Eigenfunctions of a Hermitean Operator

Consider the eigenvalue equation

$$A\psi_a = a\psi_a. \tag{V.9}$$

Throughout this section, we consider only eigensolutions ψ_a *located in Hilbert space. This restricts us to the discrete spectrum.* The general study, including the continuous spectrum (if it exists), will be made in Section III.

Since A is a linear operator:

1. If ψ_a is an eigenfunction, $c\psi_a$, c being an arbitrary constant, is also an eigenfunction of the same eigenvalue. To fix this constant, it is customary to *normalize* the eigenfunctions to unity:

$$\langle \psi_a, \psi_a \rangle = 1.$$

By doing so, ψ_a is defined to within an arbitrary phase.

2. If two linearly independent [1]) functions $\psi_a^{(1)}$ and $\psi_a^{(2)}$ correspond to the same eigenvalue, the same holds true for any linear

[1]) Two functions ψ_1 and ψ_2 are linearly independent if it is impossible to find two non-zero constants λ_1, λ_2 such that $\lambda_1\psi_1 + \lambda_2\psi_2 = 0$.

combination of these functions. One says in this case that there is *degeneracy*. The maximum number of linearly independent functions of a given eigenvalue is called the *order* of the degeneracy (we have encountered examples of degeneracy of order 2 in Chapter III while studying the continuous spectrum).

From the definition of hermiticity and the property (V.8) which follows from it, there result two fundamental properties.

1. *Every eigenvalue a is real.* Indeed, taking the scalar product of both sides of eq. (V.9) by ψ_a from the left, one shows that a is equal to the mean value of A with respect to the dynamical state ψ_a:

$$a = \frac{\langle \psi_a, A\psi_a \rangle}{\langle \psi_a, \psi_a \rangle}$$

and it is real, since A is Hermitean.

2. *Two eigenfunctions corresponding to different eigenvalues are orthogonal* (cf. Ch. III, § 13). Indeed, let us assume that

$$A\psi_1 = a_1\psi_1, \qquad A\psi_2 = a_2\psi_2.$$

Taking the scalar product from the left by ψ_2 in the first equation and the scalar product from the right by ψ_1 in the second equation, and subtracting term by term, one obtains, taking into account the property (V.8)

$$0 = \langle \psi_2, A\psi_1 \rangle - \langle A\psi_2, \psi_1 \rangle = (a_1 - a_2)\langle \psi_2, \psi_1 \rangle.$$

Therefore, if $a_1 \neq a_2$, there necessarily follows

$$\langle \psi_2, \psi_1 \rangle = 0.$$

Consequently, *two eigenfunctions ψ_1 and ψ_2 belonging to different eigenvalues are linearly independent.* Indeed, let us assume that one could find two numbers λ_1, λ_2 such that

$$\lambda_1\psi_1 + \lambda_2\psi_2 = 0.$$

Upon taking the scalar product of each term from the left by ψ_1, we have, taking into account the orthogonality,

$$\lambda_1\langle \psi_1, \psi_1 \rangle = 0,$$

hence λ_1 is necessarily zero. One shows in the same manner that $\lambda_2 = 0$.

If a is a degenerate eigenvalue of order n, each of its eigenvalues can always be put in the form of a linear combination of n particular, linearly independent eigenfunctions $\psi^{(1)}, \psi^{(2)}, \ldots, \psi^{(n)}$. There is a large arbitrariness in the choice of these n basis functions. But one can always choose them such that they are normalized to unity, and orthogonal in pairs. Starting from a set $\psi^{(1)}, \ldots, \psi^{(n)}$ which does not possess these properties one can for instance perform the following manipulations (*Schmidt's process of orthogonalization*). One defines $\varphi^{(1)}$ by

$$c_1\varphi^{(1)} = \psi^{(1)}$$

and one adjusts the constant c_1 so that $\langle\varphi^{(1)}, \varphi^{(1)}\rangle = 1$, namely

$$|c_1|^2 = \langle\psi^{(1)}, \psi^{(1)}\rangle.$$

One defines $\varphi^{(2)}$ by

$$c_2\varphi^{(2)} = \psi^{(2)} - \varphi^{(1)}\langle\varphi^{(1)}, \psi^{(2)}\rangle.$$

The left-hand side is certainly not zero since $\psi^{(1)}$ and $\psi^{(2)}$ are linearly independent. It is clear that $\langle\varphi^{(2)}, \varphi^{(1)}\rangle = 0$. c_2 is adjusted in such a way that $\langle\varphi^{(2)}, \varphi^{(2)}\rangle = 1$. $\varphi^{(3)}$ is defined by

$$c_3\varphi^{(3)} = \psi^{(3)} - \varphi^{(1)}\langle\varphi^{(1)}, \psi^{(3)}\rangle - \varphi^{(2)}\langle\varphi^{(2)}, \psi^{(3)}\rangle.$$

This function is certainly not zero; it is evidently orthogonal to $\varphi^{(1)}$ and $\varphi^{(2)}$, and can be normalized by a suitable choice of c_3. One continues in this way. The n functions $\varphi^{(1)}, \ldots, \varphi^{(n)}$ thus formed satisfy the $n(n+1)/2$ relations

$$\langle\varphi^{(l)}, \varphi^{(m)}\rangle = \delta_{lm} \qquad (l, m = 1, 2, \ldots, n),$$

where δ_{lm} is the Kronecker symbol:

$$\delta_{lm} = \begin{cases} 1 & \text{if} \quad l = m \\ 0 & \text{if} \quad l \neq m. \end{cases}$$

One says that they form a set of *orthonormal* functions.

The study of degenerate eigenvalues must be completed by the following property, which we shall state without proof (cf. footnote p. 166). If the order of degeneracy is infinite, i.e. if one can find an arbitrarily large number of linearly independent eigenfunctions, one can always form a (denumerably infinite) set $\varphi^{(1)}, \varphi^{(2)}, \ldots, \varphi^{(r)}, \ldots$

of orthonormal eigenfunctions such that any eigenfunction corresponding to the eigenvalue a can be expanded in a series of these functions.

One can also show (cf. footnote p. 166) that the eigenvalues form a *discrete set* (finite or denumerably infinite) $a_1, a_2, \ldots, a_p, \ldots$. This property is characteristic of the eigenvalues corresponding to eigenfunctions *located in Hilbert space*.

6. Expansion of a Wave Function in a Series of Orthonormal Eigenfunctions

As was shown above, one can associate with each eigenvalue a_p of A a set of orthonormal eigenfunctions

$$\varphi_p^{(1)}, \varphi_p^{(2)}, \ldots, \varphi_p^{(r)}, \ldots$$

containing one element, a finite number of elements, or an infinite number of elements according to whether a_p is non-degenerate, degenerate of finite order, or infinitely degenerate. We shall denote by $\{\varphi_p^{(r)}\}$ the set formed by all these functions. Any function of the set satisfies the relations

$$A\varphi_p^{(r)} = a_p \varphi_p^{(r)} \tag{V.12}$$

$$\langle \varphi_p^{(r)}, \varphi_q^{(s)} \rangle = \delta_{pq}\delta_{rs}. \tag{V.13}$$

The question arises now whether an arbitrary wave function ψ, belonging to Hilbert space, can be represented by a series of functions of the set $\{\varphi_p^{(r)}\}$. This series certainly exists if ψ is an eigenfunction of A; in that case, the only non-vanishing terms of the expansion are the coefficients of the functions φ corresponding to the same eigenvalue. If any ψ can be expanded in a series of eigenfunctions, one says that the set $\{\varphi_p^{(r)}\}$ is *complete*.

We shall indicate without proof (cf. footnote p. 166) some properties of expansions in series of orthonormal functions.

Let $u_1, u_2, \ldots, u_n, \ldots$ be a set of orthonormal functions:

(*i*) If ψ can be expanded in a series of these functions:

$$\psi = \sum_n c_n u_n,$$

the expansion coefficients are

$$c_n = \langle u_n, \psi \rangle$$

and satisfy the Parseval relation:

$$\sum_n |c_n|^2 = \langle \psi, \psi \rangle.$$

(*ii*) Conversely, if the numerical series $\sum_n |c_n|^2$ converges to a finite value, say N, the expansion $\sum_n c_n u_n$ converges (in quadratic mean) to a function ψ of norm N.

(*iii*) If the expansions $\sum_n c_n u_n$ and $\sum_n d_n u_n$ converge to ψ and φ respectively, the series $\sum_n d_n^* c_n$ converges to the scalar product of ψ by φ:

$$\langle \varphi, \psi \rangle = \sum_n d_n^* c_n.$$

(*iv*) For any square-integrable ψ the series

$$\hat{\psi} = \sum_n u_n \langle u_n, \psi \rangle$$

always converges. The difference $\psi - \hat{\psi}$ is orthogonal to all the functions u_n, and its norm is equal to $\langle \psi, \psi \rangle - \langle \hat{\psi}, \hat{\psi} \rangle$. One therefore always has

$$\langle \psi, \psi \rangle \geqslant \langle \hat{\psi}, \hat{\psi} \rangle$$

and when the equality holds, $\psi = \hat{\psi}$.

All these properties remain valid if the functions u are characterized by several discrete indices instead of a single one. They therefore apply to the set $\{\varphi_p^{(r)}\}$. In particular, *if the set* $\{\varphi_p^{(r)}\}$ *is complete*, any wave function Ψ can be represented by the series

$$\Psi = \sum_{p,r} c_p^{(r)} \varphi_p^{(r)} \tag{V.14}$$

whose coefficients are given by

$$c_p^{(r)} = \langle \varphi_p^{(r)}, \Psi \rangle \tag{V.15}$$

and satisfy the Parseval relation

$$\sum_{p,r} |c_p^{(r)}|^2 = \langle \Psi, \Psi \rangle. \tag{V.16}$$

Moreover, the scalar product of two wave functions Ψ_1 and Ψ_2 is

$$\langle \Psi_1, \Psi_2 \rangle = \sum_{p,r} \langle \Psi_1, \varphi_p^{(r)} \rangle \langle \varphi_p^{(r)}, \Psi_2 \rangle. \tag{V.17}$$

In fact, if one is not too concerned about mathematical rigor,

one easily verifies equation (V.15) and (V.16) by substituting for Ψ the expression (V.14) in the right-hand side of these equations and making use of the orthonormality relations (V.13). Equation (V.17) is obtained by an analogous procedure.

The analogy with ordinary vector space is striking. The complete set of orthonormal functions plays the role of a set of basis vectors of unit length being pairwise orthogonal. The function Ψ is a vector of this space (with an infinite number of dimensions), the coefficients $\langle \varphi_p^{(r)}, \Psi \rangle$ being its components along these basis vectors [eq. (V.15)], and the norm of this vector, that is to say the square of its length, is equal to the sum of the squares of the absolute values of its components [eq. (V.16)]. The scalar product of Ψ_2 by Ψ_1 is equal to the sum of the products of each component of Ψ_2 by the complex conjugate of the corresponding component of Ψ_1.

7. Statistical Distribution of the Results of Measurement of a Quantity Associated with an Operator having a Complete Set of Eigenfunctions with Finite Norm

The possibility of representing any wave function Ψ by an expansion of the type (V.14) greatly facilitates the study of all problems involving A. Let us therefore suppose that the operator possesses a complete set of orthonormal eigenfunctions [the Hamiltonian of the harmonic oscillator (Ch. XII) being an example]. This set is certainly not unique; one can always change the phases of the functions, or else replace orthonormal functions corresponding to the same eigenvalue by orthonormal linear combinations of these functions. However, the properties given below are independent of the choice of the system.

A priori, a function such as $A\Psi$ is not necessarily square-integrable. However, according to (V.14),

$$A\Psi = \sum_{p,r} c_p^{(r)} A\varphi_p^{(r)} = \sum_{p,r} a_p c_p^{(r)} \varphi_p^{(r)}.$$

This expansion converges [§ 6, properties (*i*) and (*ii*)] if and only if $\sum_{p,r} a_p^2 |c_p^{(r)}|^2$ converges and, in that case, the sum of this numerical series is equal to the norm of $A\Psi$. In this way we obtain a criterion for deciding if $A\Psi$ belongs to Hilbert space.

One arrives at analogous conclusions for the function $A^2\Psi$. More generally, starting from a function $F(x)$, one can, by means of the

expansion (V.14) define the operator $F(A)$, a function of the operator A. Its action on Ψ is given by

$$F(A)\Psi = \sum_{p,r} F(a_p)\, c_p{}^{(r)}\, \varphi_p{}^{(r)}.$$

The operator is well-defined if the series converges, i.e. for all functions Ψ such that the numerical series

$$\sum_{p,r} |F(a_p)|^2\, |c_p{}^{(r)}|^2$$

is convergent. In that case — as can be verified easily by the reader — the function $F(A)\Psi$ thus defined is the same whatever the set $\{\varphi_p{}^{(r)}\}$ which served for its definition.

In particular, the operator $\exp(i\xi A)$, where ξ is a given parameter, is defined for all functions of Hilbert space. Indeed

$$e^{i\xi A}\Psi = \sum_{p,r} e^{i\xi a_p}\, c_p{}^{(r)}\, \varphi_p{}^{(r)} \tag{V.18}$$

and the convergence criterion for the series, namely the convergence of the numerical series $\sum_{p,r} |c_p{}^{(r)}|^2$, is always fulfilled.

We are now in a position to establish the statistical distribution of the quantity $\mathscr{A}$ for any dynamical state of the system. Indeed, the characteristic function $f(\xi)$ of this distribution [1]), the knowledge of which suffices to determine the latter completely, is by definition

---

[1]) To within a constant, it is the Fourier transform of this distribution. Let X be a random variable able to take on all values contained in the interval $(-\infty, +\infty)$ and let $P(x)$ be the probability of finding X in the interval $(x, x + dx)$; the characteristic function $f(\xi)$ of its statistical distribution is the average value of $\exp(i\xi x)$, that is to say

$$f(\xi) = \int_{-\infty}^{+\infty} e^{i\xi x}\, P(x)\, dx.$$

If X can only take certain discrete values $x_1, \ldots, x_n, \ldots$ and if $w_1, \ldots, w_n, \ldots$ are the respective probabilities of these values,

$$f(\xi) = \sum_n w_n\, e^{i\xi x_n}.$$

More generally, if $F(x)$ is the probability that $X \leqslant x$, one has, upon introducing the Stieltjes integral

$$f(\xi) = \int_{-\infty}^{+\infty} e^{i\xi x}\, dF(x).$$

the mean value of the quantity $\exp(i\xi A)$ in that state. By a very natural extension of postulate (*b*) of § 3, we define this average value by the expression

$$f(\xi) = \frac{\langle \Psi, e^{i\xi A}\Psi \rangle}{\langle \Psi, \Psi \rangle} \tag{V.19}$$

(which is always meaningful, even if the mean value of A is not defined).

Let Ψ be the wave function representing the dynamical state under consideration. Making use of the expansions (V.14) and (V.18) and expression (V.17) for the scalar product, one has

$$f(\xi) = \sum_p w_p e^{i\xi a_p},$$

with the notation

$$w_p = \frac{\sum_r |c_p^{(r)}|^2}{\langle \Psi, \Psi \rangle} = \frac{\sum_r |\langle \varphi_p^{(r)}, \Psi \rangle|^2}{\langle \Psi, \Psi \rangle}.$$

By inspection of the characteristic function, we conclude that:

1. *The only values the quantity $\mathscr{A}$ can take on are the eigenvalues $a_1, a_2, \ldots, a_p, \ldots$ of the operator associated with it;*

2. *The probability that $\mathscr{A}$ takes one of these values, say a_p, is equal to w_p.*

One easily verifies that $\sum_p w_p = 1$ (Parseval's relation), and that the average value of $\mathscr{A}$ is given, provided that it converges, by

$$\langle A \rangle = \sum_p w_p a_p,$$

also, more generally, the mean value of a function $f(\mathscr{A})$ is given, if it exists, by

$$\langle f(A) \rangle = \sum_p w_p f(a_p).$$

In particular, in order that $\mathscr{A}$ take on a given value with certainty, it is necessary and sufficient that Ψ be an eigenfunction corresponding to that eigenvalue, in accordance with the deductions of § 4.

One can express the general result above in a way which reveals even more clearly the fact that it does not depend upon the choice of the system $\{\varphi_p^{(r)}\}$. The function Ψ_p defined by

$$\Psi_p = \sum_r \varphi_p^{(r)} \langle \varphi_p^{(r)}, \Psi \rangle,$$

is indeed independent of this choice (cf. Problem V.4). The expansion (V.14) can also be written

$$\Psi = \sum_p \Psi_p. \tag{V.20}$$

Stated differently, one can in a unique manner put Ψ in the form of a sum of eigenfunctions of A, each of which corresponds to a different eigenvalue. Then, the probability w_p of finding the particular value a_p is equal to the ratio of the norm of Ψ_p to that of Ψ:

$$w_p = \frac{\langle \Psi_p, \Psi_p \rangle}{\langle \Psi, \Psi \rangle}. \tag{V.21}$$

III. STATISTICS OF MEASUREMENT IN THE GENERAL CASE

8. Difficulties of the Continuous Spectrum. Introduction of the Dirac δ-Functions

All the conclusions which we have reached are vitiated if the set $\{\varphi_p^{(r)}\}$ is not complete. That case is by no means exceptional. However, the discussions of § 4 suggest a possible extension of the foregoing theory, in which one would still start from equation (V.9) but impose on the eigensolutions conditions less restrictive than to belong to Hilbert space. These more general eigensolutions may have an infinite norm. Our first step is to extend to them the concepts of orthogonality and normalization so important in this whole theory.

For purposes of orientation, consider two examples relating to one-dimensional systems, namely the statistical distributions of position and momentum. In that case, the statistical distributions are known; the formal extension of the results of the preceding section is thereby simplified. The position coordinate q may take on all possible values in the interval $(-\infty, +\infty)$ and the probability of finding q in the interval $(q', q'+\mathrm{d}q')$ is equal to

$$P(q')\,\mathrm{d}q' = |\psi(q')|^2\,\mathrm{d}q' \tag{V.22}$$

where $\psi(q)$ is the wave function – assumed normalized to unity – representing the dynamical state of the system. Likewise, the momentum p represented by the operator $(\hbar/\mathrm{i})\mathrm{d}/\mathrm{d}q$ may take on all possible values in the interval $(-\infty, +\infty)$ and the probability of finding p in the interval $(p', p'+\mathrm{d}p')$ is equal to

$$\Pi(p')\,\mathrm{d}p' = |\varphi(p')|^2\,\mathrm{d}p' \tag{V.23}$$

where $\varphi(p)$ is the Fourier transform, suitably normalized, of the wave function $\psi(q)$.

In both cases, the possible values of the quantities considered form a *continuous spectrum*. Therein lies the principal difference from the foregoing situation characterized by a discrete spectrum and by the possibility of representing any wave function ψ by a series [cf. eq. (V.14) or (V.20)] each term of which is associated with one of the possible values of that spectrum. The natural extension to the case of the continuous spectrum consists in representing ψ no longer by a series but by an integral.

Let us recall briefly the theory of the entirely discrete spectrum.

The Hermitean operator A possesses a discrete set of eigenvalues, assumed non-degenerate for the sake of simplicity. To each of these, a_n, corresponds an eigenfunction φ_n (defined to within a phase) and we have

$$A\varphi_n = a_n\varphi_n \tag{V.24}$$

$$\langle\varphi_n, \varphi_{n'}\rangle = \delta_{nn'}. \tag{V.25}$$

The orthonormal set $\{\varphi_n\}$ being complete, any function assumed normalized to unity can be represented by the series

$$\psi = \sum_n c_n\varphi_n \tag{V.26}$$

and we have from the orthonormality relations (V.25),

$$\langle\varphi_n, \psi\rangle = \sum_{n'} c_{n'}\langle\varphi_n, \varphi_{n'}\rangle = c_n. \tag{V.27}$$

By applying the same relations, one obtains for the expression of the characteristic function

$$\langle\psi, e^{i\xi A}\psi\rangle = \sum_{n,n'} c_n^* c_{n'} e^{i\xi a_{n'}}\langle\varphi_n, \varphi_{n'}\rangle = \sum_n |c_n|^2 e^{i\xi a_n}, \tag{V.28}$$

whence one deduces that the probability that A take the value a_n is equal to the square $|c_n|^2$ of the modulus of the coefficient of φ_n in the expansion (V.26).

By analogy, let us denote by $u(p'; q)$ the eigensolution of the operator $p = (\hbar/i)d/dq$ relative to the eigenvalue p',

$$pu(p'; q) \equiv \frac{\hbar}{i}\frac{d}{dq}u(p'; q) = p'u(p'; q). \tag{V.24'}$$

Pursuing the formal development, we represent the wave function $\psi(q)$ by an integral of the form

$$\psi(q) = \int_{-\infty}^{+\infty} c(p')\, u(p'; q)\, \mathrm{d}p'.$$

If the formal analogy can be carried to the end, $|c(p')|^2\, \mathrm{d}p'$ is the probability that p is found in the interval $(p', p'+\mathrm{d}p')$. One must therefore have $|c(p')|^2 = |\varphi(p')|^2$; in other words $c(p')$ is equal to $\varphi(p')$ to within a phase factor. Since the eigensolution $u(p'; q)$ is itself defined to within a constant, one can always choose the latter in such a way that the analogue of eq. (V.26) can be written

$$\psi(q) = \int_{-\infty}^{+\infty} \varphi(p')\, u(p'; q)\, \mathrm{d}p'. \qquad \text{(V.26')}$$

The coefficient $\varphi(p')$ corresponding to each eigenvalue p' must be, by generalization of eq. (V.27), equal to

$$\varphi(p') = \langle u(p'; q), \psi(q) \rangle \qquad \text{(V.27')}$$

which gives, after substituting for $\psi(q)$ its integral representation (V.26') in the scalar product,

$$\varphi(p') = \int_{-\infty}^{+\infty} \varphi(p'') \langle u_{p'}, u_{p''} \rangle\, \mathrm{d}p'', \qquad \text{(V.29)}$$

an expression in which we have used the abbreviated notation $u_{p'}$ to denote the eigensolution corresponding to the eigenvalue p'. This property of the scalar product $\langle u_{p'}, u_{p''} \rangle$ must hold for any wave function $\varphi(p')$ of momentum space (the only condition imposed upon φ is to be square-integrable); this property thus generalizes the orthonormality relations (V.25).

In fact, no function of p' and p'' possesses the desired property. Nevertheless, if one does not bother too much about mathematical rigor [1]), one can, following Dirac, make use of the "singular function"

---

[1]) We have already relaxed the mathematical rigor by writing down equation (V.29). The only correct equation is

$$\varphi(p') = \langle u_{p'}, \int \varphi(p'')\, u_{p''}\, \mathrm{d}p'' \rangle$$

from which one obtains (V.29) by inverting the order of the integration over q and p'' This manipulation is certainly not justified since the scalar product $\langle u_{p'}, u_{p''} \rangle$ diverges.

$\delta(x)$ defined by the property

$$\int_a^b f(x)\,\delta(x-x_0)\,dx = \begin{cases} f(x_0) & \text{for } x_0 \text{ in the interval } (a,b) \\ 0 & \text{for } x_0 \text{ outside the interval } (a,b) \end{cases} \tag{V.30}$$

for any function that is continuous at the point $x=x_0$.

Equation (V.29) is satisfied if

$$\langle u_{p'},\, u_{p''}\rangle = \delta(p'-p''). \tag{V.25'}$$

This relation is the generalization to the continuous spectrum of the orthonormality relation (V.25).

One can visualize the "Dirac function" $\delta(x)$ as the limit of a function that is zero everywhere except in a very small interval surrounding the point $x=0$ where it has a very narrow and very high peak such that its integral over the entire interval is equal to 1. In the limit where the width of the peak goes to zero, one has

$$\delta(x) = \begin{cases} 0 & \text{if } x \neq 0 \\ +\infty & \text{if } x = 0 \end{cases} \quad \text{and} \quad \int_{-\infty}^{+\infty} \delta(x)\,dx = 1. \tag{V.31}$$

$\delta(x)$ is certainly not a function in the usual sense of the word since the integral, if it exists, of a function which is zero everywhere except at one point, necessarily vanishes. We shall not dwell here on the mathematical justification of the use of the "Dirac function". It requires the introduction of a new concept, that of distribution, of which the usual functions (more precisely the locally integrable functions) are particular cases. The mathematicians do not speak of the function $\delta(x-x_0)$ but of the distribution $\delta_{x_0}[f]$, defined as the functional of $f(x)$ equal to $f(x_0)$. In other words, the definition (V.30) must be replaced by the definition:

$$\delta_{x_0}[f] = f(x_0).$$

As the concept of distribution may be unfamiliar to the reader, we shall refer to it as little as possible and we shall continue to make use of the (incorrect) notation $\delta(x-x_0)$ which has in fact some indisputable formal advantages. Its main rules of calculation are listed in Appendix A, where one will also find a brief outline of Distribution Theory.

Let us now take up the problem of the measurement of p. The eigensolution $u(p';q)$ of equation (V.24') is the function $c\exp(ip'q/\hbar)$.

The generalized orthonormality relation (V.25′) is satisfied if one takes $c = 1/\sqrt{2\pi\hbar}$, namely

$$u(p'; q) = \frac{1}{\sqrt{2\pi\hbar}} e^{ip'q/\hbar}.$$

With this, and using eq. (A.22), we have

$$\langle u_{p'}, u_{p''} \rangle = \frac{1}{2\pi\hbar} \int_{-\infty}^{+\infty} e^{i(p''-p')q/\hbar} \, dq = \delta(p'-p'').$$

In addition, the eigenfunctions $u_{p'}$ form a complete set since any square-integrable function $\psi(q)$ can be put in the form (properties of the Fourier integral)

$$\psi(q) = \int_{-\infty}^{+\infty} \varphi(p') \frac{e^{ip'q/\hbar}}{\sqrt{2\pi\hbar}} \, dp'.$$

The coefficient $\varphi(p')$ of this "expansion in a series of eigenfunctions" is actually equal to the scalar product $\langle u_{p'}, \psi \rangle$. Indeed

$$\langle u_{p'}, \psi \rangle = \langle u_{p'}, \int_{-\infty}^{+\infty} \varphi(p'') \, u_{p''} \, dp'' \rangle = \int_{-\infty}^{+\infty} \varphi(p'') \langle u_{p'}, u_{p''} \rangle \, dp''$$

$$= \int_{-\infty}^{+\infty} \varphi(p'') \, \delta(p'-p'') \, dp'' = \varphi(p').$$

We recognize the reciprocity property of the Fourier integral, making use of the generalized orthonormality relation.

Pursuing further the analogy with the case of the discrete spectrum, we introduce the operator $\exp(i\xi p)$. We have

$$e^{i\xi p} \, u(p'; q) = e^{i\xi p'} \, u(p'; q),$$

hence

$$e^{i\xi p} \, \psi(q) = \int_{-\infty}^{+\infty} \varphi(p') \, e^{i\xi p'} \, u(p'; q) \, dp';$$

from this the characteristic function

$$f(\xi) = \langle \psi, e^{i\xi p} \, \psi \rangle = \int_{-\infty}^{+\infty} \varphi^*(p'') \, dp'' \int_{-\infty}^{+\infty} e^{i\xi p'} \varphi(p') \, dp' \times \langle u_{p''}, u_{p'} \rangle$$

$$= \int_{-\infty}^{+\infty} \varphi^*(p'') \, dp'' \int_{-\infty}^{+\infty} e^{i\xi p'} \varphi(p') \, dp' \times \delta(p'-p'')$$

$$= \int_{-\infty}^{+\infty} |\varphi(p')|^2 \, e^{i\xi p'} \, dp'.$$

The statistical distribution which corresponds to it (cf. footnote p. 177) is just the expected distribution (V.23).

The study of the position measurement can be made by the same scheme. The eigensolution, suitably normalized, corresponding to the eigenvalue q' of operator q, is $\delta(q'-q)$. Indeed, according to equation (A.19),

$$q\delta(q'-q)=q'\delta(q'-q).$$

The ensemble of functions $\delta(q'-q)$, where q' can take all possible values from $-\infty$ to $+\infty$, forms an orthonormal set since [eq. (A.21)]

$$\int_{-\infty}^{+\infty} \delta(q-q')\,\delta(q-q'')\,\mathrm{d}q=\delta(q'-q''),$$

and this system is complete since any wave function $\psi(q)$ has the integral representation

$$\psi(q)=\int_{-\infty}^{+\infty} \psi(q')\,\delta(q'-q)\,\mathrm{d}q. \tag{V.32}$$

One easily verifies that the coefficient $\psi(q')$ is equal to the scalar product $\langle\delta(q'-q),\psi(q)\rangle$. Following the same prescription as for the measurement of p, one arrives at the conclusion that the square of the modulus of that coefficient $|\psi(q')|^2$ is equal to the probability density that $q=q'$, in accordance with eq. (V.22).

9. Expansion in a Series of Eigenfunctions in the General Case. Closure Relation

Let us return to the eigenvalue problem

$$A\psi=a\psi. \tag{V.9}$$

We no longer require that the eigensolutions have a finite norm but merely that their scalar product by any wave function (i.e. by any function of finite norm) be convergent.

Most generally, the eigenvalue spectrum consists of:

1. a discrete part, i.e. a finite or denumerably infinite set of values a_n, labelled by the integral index n;

2. a continuous part, i.e. a set of values $a(\nu)$ which can be conveniently defined as the values taken by a uniform, continuous and monotonic function of some index ν.

The eigenfunctions of the discrete part of the spectrum have finite

norm. All properties of the discrete spectrum were studied in § 5, and do not need to be repeated here.

Let $\psi(\nu; q_1, \ldots, q_R)$ be an eigenfunction of the continuous spectrum corresponding to the eigenvalue $a(\nu)$. It is a continuous function of the parameter ν whose normalization integral certainly diverges (it is a "vector of infinite length" in function space). However, the eigendifferential

$$(\Delta\nu)^{-\frac{1}{2}} \int_{\nu}^{\nu+\Delta\nu} \psi(\nu'; q_1, \ldots, q_R)\, \mathrm{d}\nu'$$

is supposed to be a function whose norm is finite and tends to a constant value when $\Delta\nu$ goes to zero. We know that the eigenfunctions of the continuous spectrum of the Hamiltonian of a one-dimensional system possess this property (cf. Ch. III); the same holds true, as is easily verified, for the operators q and $(\hbar/\mathrm{i})\mathrm{d}/\mathrm{d}q$ of the preceding paragraph.

One often says that a function is *normalizable* when its norm is finite. By a very natural extension of the terminology, we shall likewise apply this term to fun tions of infinite norm whose eigendifferential has a finite norm. Thus the eigenfunctions of the continuous spectrum are normalizable in this sense, as well as those of the discrete spectrum, although they do not belong to Hilbert space.

Using property (V.8) (which has a meaning only for square-integrable functions Φ and Ψ) suitably applied to eigendifferentials rather than to eigenfunctions, one obtains the fundamental properties concerning the continuous spectrum which are the counterpart to the properties of the discrete spectrum shown in Chapter V, § 5:

1. every eigenvalue $a(\nu)$ is real;

2. two eigenfunctions corresponding to distinct eigenvalues are orthogonal. To be precise the orthogonality property is the generalization of equation (III.42′). It is not correct to write

$$\langle \psi_\nu, \psi_{\nu'} \rangle = 0$$

since the scalar product $\langle \psi_\nu, \psi_{\nu'} \rangle$ diverges in general. But one has

$$\langle \psi_\nu, (\Delta\nu')^{-\frac{1}{2}} \int_{\nu'}^{\nu'+\Delta\nu'} \psi_{\nu''}\, \mathrm{d}\nu'' \rangle = 0$$

when ν lies outside the interval $(\nu', \nu'+\Delta\nu')$. The proof of these properties is not difficult and will be left to the reader.

If the eigenvalues of the continuous spectrum are not degenerate, one can always normalize them in such a way that

$$\langle \psi_\nu, \psi_{\nu'} \rangle = \delta(\nu - \nu').$$

The extension to degenerate eigenvalues is straightforward. We shall merely quote the results. To a given eigenvalue there corresponds, depending upon the nature of the degeneracy, either a finite, or an infinite (denumerable or non-denumerable) number of linearly independent eigenfunctions. One can label such functions either with the aid of an index taking on a finite number of values, or by one or more indices taking an infinite number of discrete values, or else by one or more indices varying continuously; or even by a number of discrete indices and a number of continuous indices. Assume, to be definite, that the labelling is done with a discrete index r and a continuous index ϱ.

One can always arrange matters in such a way that the eigenfunctions $\varphi^{(r)}(\nu, \varrho)$ are orthonormal, in the generalized sense of the term, that is to say, such that

$$\langle \varphi^{(r)}(\nu, \varrho), \varphi^{(r')}(\nu', \varrho') \rangle = \delta_{rr'} \delta(\varrho - \varrho')\, \delta(\nu - \nu'). \tag{V.33}$$

Together with the orthonormal eigenfunctions $\varphi_n{}^{(r)}$ of the discrete spectrum, these functions form an orthonormal set of eigenfunctions of the Hermitean operator A; any eigenfunction of A can be written as a linear combination of the functions of this set.

Let us assume that a wave function Ψ can be expanded in a series of these functions, that is:

$$\Psi = \sum_{nr} c_n{}^{(r)} \varphi_n{}^{(r)} + \sum_r \int \gamma^{(r)}(\nu, \varrho)\, \varphi^{(r)}(\nu, \varrho)\, \mathrm{d}\nu\, \mathrm{d}\varrho. \tag{V.34}$$

The coefficients of each eigenfunction in this expansion are obtained by taking the scalar product of both sides of this equation from the left by the eigenfunction in question. One obtains, taking into account the orthonormality relations

$$c_n{}^{(r)} = \langle \varphi_n{}^{(r)}, \Psi \rangle \tag{V.35a}$$

$$\gamma^{(r)}(\nu, \varrho) = \langle \varphi^{(r)}(\nu, \varrho), \Psi \rangle. \tag{V.35b}$$

From these same relations, we arrive at the generalized Parseval relation

$$\langle \Psi, \Psi \rangle = \sum_{n,r} |c_n{}^{(r)}|^2 + \sum_r \int |\gamma^{(r)}(\nu, \varrho)|^2\, \mathrm{d}\nu\, \mathrm{d}\varrho. \tag{V.36}$$

If every wave function (i.e. every square-integrable function) can be expanded in a series of type (V.34), the set $\{\varphi\}$ is then, by definition, a *complete* orthonormal set.

There is a very simple way of writing down that an orthonormal set is complete: it consists in writing the expansion (V.34) for the function

$$\delta(q-q') \equiv \delta(q_1-q_1')\,\delta(q_2-q_2')\,\ldots\,\delta(q_R-q_R').$$

The coefficients of this expansion are obtained by substituting in expressions (V.35) and (V.36) the function $\delta(q-q')$ for the function $\Psi(q) \equiv \Psi(q_1, q_2, \ldots, q_R)$. One is led to the so-called *closure relation*

$$\delta(q-q') = \sum_{nr} \varphi_n^{*(r)}(q')\,\varphi_n^{(r)}(q) + \sum_r \int \varphi^{*(r)}(\nu, \varrho; q')\,\varphi^{(r)}(\nu, \varrho; q)\,\mathrm{d}\nu\,\mathrm{d}\varrho. \quad \text{(V.37)}$$

Together with the orthonormality relations

$$\langle \varphi_n^{(r)},\, \varphi_{n'}^{(r')} \rangle = \delta_{nn'}\,\delta_{rr'} \quad \text{(V.38}a\text{)}$$

$$\langle \varphi_n^{(r)},\, \varphi^{(r')}(\nu, \varrho) \rangle = 0 \quad \text{(V.38}b\text{)}$$

$$\langle \varphi^{(r)}(\nu, \varrho),\, \varphi^{(r')}(\nu', \varrho') \rangle = \delta_{rr'}\,\delta(\varrho-\varrho')\,\delta(\nu-\nu') \quad \text{(V.38}c\text{)}$$

it forms a set of necessary and sufficient conditions for the set $\{\varphi\}$ to be orthonormal and complete.

Expansion (V.34) is obtained, with the correct values (V.35) of the coefficients, by writing that

$$\Psi(q) = \int \delta(q-q')\,\Psi(q')\,\mathrm{d}\tau$$

and substituting for $\delta(q-q')$ the right-hand side of eq. (V.37).

Note that the complete, orthonormal set, if it exists, is certainly not unique. Indeed, one can, just as in the case when the spectrum is entirely discrete,

1) arbitrarily change the phases of each of the eigenfunctions;

2) choose in an infinite number of different ways the set of orthonormal functions corresponding to the same degenerate eigenvalue.

In addition:

3) there is some arbitrariness in the normalization of the eigenfunctions of the continuous spectrum. Indeed, one can replace any continuous index ν by the index $\mu \equiv \mu(\nu)$, where $\mu(\nu)$ is a continuous, differentiable monotonic function of ν, but otherwise arbitrary. The

normalization condition (V.38c) is replaced by an analogous condition with index μ instead of ν; the latter is satisfied if one takes as a new eigenfunction

$$\varphi^{(r)}(\mu, \varrho; q) = \left|\frac{\mathrm{d}\mu}{\mathrm{d}\nu}\right|^{-\frac{1}{2}} \varphi^{(r)}(\nu, \varrho; q). \tag{V.39}$$

All Hermitean operators do not possess a complete, orthonormal set of eigenfunctions [1]). However, the Hermitean operators capable of representing physical quantities possess such a set. For this reason, we give the name *observables* to such operators. To prove that a specific Hermitean operator is an observable is, often, a difficult mathematical problem. The proof has actually been given for simple cases such as, for instance, the position or momentum coordinates, the Hamiltonian of one-dimensional quantum systems, angular momentum, etc. In the following, we shall always take for granted that all the operators associated with physical quantities possess a complete orthonormal set of eigenfunctions. In fact, the completeness property is so closely related to the physical interpretation of these operators that the whole theory would have to be profoundly revised if it did not hold true.

10. Statistical Distribution of the Results of Measurement in the General Case

We assume that A is an observable. Since the expansion (V.34) exists for any square-integrable Ψ it is possible to define, except for possible divergence of the series, the action of an operator of the form $F(A)$ on any such function.

Assume for simplicity that the spectrum of A possesses no degeneracy. Thus

$$\Psi = \sum_n c_n \varphi_n + \int \gamma(\nu)\, \varphi(\nu)\, \mathrm{d}\nu. \tag{V.40}$$

---

[1]) Let us consider the operator $\mathrm{i}(\mathrm{d}/\mathrm{d}x)$ acting on the square-integrable functions $\psi(x)$ defined on the semi-axis $(0, +\infty)$. This operator is Hermitean if one restricts it to functions which vanish for $x = 0$. In fact, it is under these conditions that one has

$$\int_0^\infty \psi_1^* \left(\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}x}\psi_2\right) \mathrm{d}x - \int_0^\infty \left(\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}x}\psi_1\right)^* \psi_2\, \mathrm{d}x = \mathrm{i}\,\psi_1^*\,\psi_2\Big|_0^\infty = 0.$$

However, it possesses no eigensolution at all. The only possible eigensolutions are in fact of the form $\exp(-ikx)$ (eigenvalue k); they not do vanish at $x = 0$.

One can show that the necessary and sufficient condition for the convergence of the expansion of the right-hand side is that the series $\sum_n |c_n|^2$ and the integral $\int |\gamma(\nu)|^2 \, d\nu$ converge.

In all generality

$$F(A)\Psi = \sum_n c_n F(a_n) \varphi_n + \int \gamma(\nu) F(a_\nu) \varphi(\nu) \, d\nu$$

and this definition has a meaning if

$$\sum_n |c_n|^2 |F(a_n)|^2 \quad \text{and} \quad \int |\gamma(\nu)|^2 |F(a_\nu)|^2 \, d\nu$$

converge. In particular, the action of the operator $\exp(i\xi A)$ is always well defined since the expression

$$e^{i\xi A}\Psi = \sum_n c_n e^{i\xi a_n} \varphi_n + \int \gamma(\nu) e^{i\xi a_\nu} \varphi(\nu) \, d\nu \tag{V.41}$$

always converges.

In order to arrive at the characteristic function of the distribution of A, we make use of the relation

$$\langle \Psi, e^{i\xi A}\Psi \rangle = \sum_n |c_n|^2 e^{i\xi a_n} + \int |\gamma(\nu)|^2 e^{i\xi a_\nu} \, d\nu$$

obtained by using the expansions (V.40) and (V.41) and the orthonormality relations. The characteristic function $f(\xi)$ can be put in the form:

$$f(\xi) = \frac{\langle \Psi, e^{i\xi A}\Psi \rangle}{\langle \Psi, \Psi \rangle} = \sum_n w_n e^{i\xi a_n} + \int \varpi(\nu) e^{i\xi a_\nu} \, d\nu$$

with the notation

$$w_n = \frac{|c_n|^2}{\langle \Psi, \Psi \rangle}, \qquad \varpi(\nu) = \frac{|\gamma(\nu)|^2}{\langle \Psi, \Psi \rangle}. \tag{V.42}$$

Inspection of this characteristic function (cf. footnote p. 177) leads to the following conclusions:

1) *the only values the quantity $\mathscr{A}$ can take are the eigenvalues of the operator associated with it;*

2) *the probability that $\mathscr{A}$ takes one of the values a_n of the discrete spectrum is equal to w_n;*

3) *the probability that $\mathscr{A}$ takes one of the values of the continuous spectrum contained in the interval* $[a(\nu), a(\nu+\mathrm{d}\nu)]$ *is equal to* $\varpi(\nu)\,\mathrm{d}\nu$.

The sum of all these probabilities $\sum w_n + \int \varpi(\nu)\,\mathrm{d}\nu$ is actually equal to 1 (Parseval relation). One can also verify that the average value of A, if it exists, is actually equal to $\langle \Psi, A\Psi\rangle / \langle \Psi, \Psi\rangle$, in accordance with the fundamental postulate from which we started out.

If the spectrum is degenerate, one obtains the same result with a suitable modification of the quantities w_n and $\varpi(\nu)$. To be definite, let us assume that expansion (V.40) is replaced by expansion (V.34); the new definitions of w_n and $\varpi(\nu)$ are:

$$w_n = \frac{\sum_r |c_n^{(r)}|^2}{\langle \Psi, \Psi\rangle} \tag{V.43}$$

$$\varpi(\nu) = \frac{\sum_r \int \mathrm{d}\varrho\, |\gamma^{(r)}(\nu, \varrho)|^2}{\langle \Psi, \Psi\rangle}. \tag{V.44}$$

As it stands here, the desired probability law involves explicitly a particular set of eigenfunctions of A. There is a high degree of arbitrariness in the choice of that set. It is clear, however, that the probability law, just as its characteristic function, must be independent of that choice. This property is easily verified directly in expressions (V.43) and (V.44) (Problem V.5).

11. Other Ways of Treating the Continuous Spectrum

The main advantage of the foregoing treatment lies in its great formal simplicity. This advantage compensates to a large extent for the lack of mathematical rigor when using the δ-"function". In fact, all operations carried out with the δ-"function" can be made rigorous with the help of Distribution Theory (cf. Appendix A).

However, it is possible to overcome the difficulties of the continuous spectrum by having recourse to more classical mathematical procedures. Rather than base this entire analysis upon the eigenvalue problem and to make use of eigensolutions which may not belong to Hilbert space, one can, following a method due to von Neumann, treat the problem rigorously without ever leaving Hilbert space. The method consists in seeking what mathematicians call the decompositions of unity in Hilbert space, and to show that with every observable of

Wave Mechanics there is associated a particular decomposition of unity. This treatment is strictly equivalent to the treatment given above. We mention it here only for completeness [1]).

Another way of treating problems connected with the continuous spectrum consists in replacing the eigenvalue problem (V.9) by another problem where the sequence of eigenvalues is entirely discrete, and of which the initial problem may be considered as a limiting case when the conditions are suitably modified. Although such a procedure cannot claim to be rigorous, it has the advantage of being intuitively very simple to grasp. We shall treat the operators q and $(\hbar/\mathrm{i})\mathrm{d}/\mathrm{d}q$ by this method. The treatment should be compared to that of § 8.

To treat the problem of the position measurement, we divide the interval $(-\infty, +\infty)$ into equal segments of length η, and replace the wave functions $\psi(q)$ by approximate wave functions $\psi^{\times}(q)$ which are constant in each segment and are defined by the condition

$$\psi^{\times}(q) = \psi(n_q \eta),$$

where n_q denotes the largest integer contained in q/η; in other words: $q - \eta < n_q \eta \leqslant q$. Similarly, we replace the operator q by the operator

$$q^{\times} = \text{multiplication by } n_q \eta.$$

In the limit where $\eta \to 0$, $q^{\times} \to q$ and $\psi^{\times}(q) \to \psi(q)$.

The set of functions $\psi^{\times}$ forms a Hilbert space in which the operator $q^{\times}$ is well defined and possesses a discrete eigenvalue spectrum

$$n\eta \qquad (n = 0, \pm 1, \pm 2, \pm 3, \ldots, \pm\infty).$$

With each eigenvalue $n\eta$ is associated the eigenfunction u_n normalized to unity

$$u_n = \begin{cases} \eta^{-\frac{1}{2}} & \text{if} \quad n\eta \leqslant q < (n+1)\eta \\ 0 & \text{in the opposite case.} \end{cases}$$

The eigenfunctions u_n satisfy the orthonormality relations

$$\langle u_n, u_{n'} \rangle = \delta_{nn'}.$$

---

[1]) The reader who is interested in the mathematical aspects of Quantum Theory, will find a complete presentation thereof in J. von Neumann, *loc. cit.* (footnote p. 166).

Furthermore, they form a complete set since any function can be represented by an expansion in a series of the u_n:

$$\psi^{\times} = \sum_{n=-\infty}^{+\infty} \eta^{\frac{1}{2}}\, \psi(n\eta)\, u_n. \tag{V.45}$$

One can therefore apply the theory of §§ 5 and 6 with the result that the probability that $q^{\times} = n\eta$ is equal to $\eta|\psi(n\eta)|^2$.

In the limit $\eta \to 0$, the distance between neighboring eigenvalues goes to zero; the spectrum becomes continuous. The value n corresponding to a given non-zero value $q' = n\eta$ goes to infinity; however, the probability of finding this precise value of the position is proportional to η and therefore tends to zero. Actually, this probability is of no concern since the position spectrum is continuous. One actually looks for the probability $P(q')\,\delta q'$ of finding the particle in the interval $(q', q'+\delta q')$, namely

$$P(q')\,\delta q' = \sum^{(q',\, q'+\delta q')} \eta|\psi(n\eta)|^2,$$

the summation being extended over all n such that $n\eta$ be contained in the interval $(q', q'+\delta q')$. As $\delta q'$ remains constant but sufficiently small, the terms of that sum, $\delta q'/\eta$ in number, are practically all equal to $\eta|\psi(q')|^2$. Hence, in the limit $\eta \to 0$,

$$P(q')\,\delta q' = |\psi(q')|^2\,\delta q'.$$

Note that expansion (V.45) can also be written

$$\begin{aligned}\psi^{\times} &= \sum_{n=-\infty}^{+\infty} \psi(n\eta)\, \frac{u_n}{\sqrt{\eta}} \times \eta \\ &= \sum_{q'} \psi(q')\, v_\eta(q') \times \eta,\end{aligned}$$

where $\sum_{q'}$ indicates a summation over the discrete sequence of values $q' = n\eta$ and where

$$v_n(q') = \eta^{-\frac{1}{2}}\, u_n = \begin{cases} \eta^{-1} & \text{if} \quad q-\eta < q' \leqslant q \\ 0 & \text{otherwise.} \end{cases}$$

When $\eta \to 0$, the series tends to the integral of the product of $\psi(q')$ by the limit of the function $v_\eta(q')$; this limit is precisely $\delta(q'-q)$. One recognizes the expansion (V.32).

In an anologous way the momentum measurement may be treated by restricting the interval of variation of q to the domain $(-L/2,$

$+L/2$), where L is a quantity which is finally allowed to tend to infinity. In order that the operator $p=(\hbar/\mathrm{i})\mathrm{d}/\mathrm{d}q$ be Hermitean in this finite space, one must impose suitable conditions upon the functions $\psi(q)$ at the boundaries of the interval. The hermiticity condition reads

$$\int_{-L/2}^{+L/2} \varphi^* \left(\frac{\hbar}{\mathrm{i}}\frac{\mathrm{d}\psi}{\mathrm{d}q}\right) \mathrm{d}q - \int_{-L/2}^{+L/2} \left(\frac{\hbar}{\mathrm{i}}\frac{\mathrm{d}\varphi}{\mathrm{d}q}\right)^* \psi \, \mathrm{d}q \equiv \frac{\hbar}{\mathrm{i}} \varphi^*\psi \Big|_{-L/2}^{+L/2} = 0,$$

for any functions $\varphi(q)$ and $\psi(q)$, or still

$$\frac{\psi(L/2)}{\psi(-L/2)} = \frac{\varphi^*(-L/2)}{\varphi^*(L/2)} = \text{constant, independent of } \psi \text{ and } \varphi.$$

In other words, it is necessary that for any function $\psi(q)$:

$$\psi(L/2) = \mathrm{e}^{\mathrm{i}\alpha}\psi(-L/2),$$

where $\exp(\mathrm{i}\alpha)$ is a phase factor chosen once and for all. Let us make it equal to 1, hence the *condition of periodicity*

$$\psi(L/2) = \psi(-L/2).$$

The eigenvalue problem of the operator $(\hbar/\mathrm{i})\mathrm{d}/\mathrm{d}q$ is then easily solved. The eigenvalue spectrum is discrete:

$$p_n = \frac{2\pi\hbar}{L} n \qquad (n = 0, \pm 1, \pm 2, \ldots, \pm\infty).$$

To the eigenvalue p_n corresponds the eigenfunction normalized to unity

$$u_n = L^{-\frac{1}{2}} \mathrm{e}^{\mathrm{i}p_n q/\hbar}.$$

The functions u_n are mutually orthogonal; furthermore, they form a complete set since, according to the theory of Fourier series, any square-integrable function $\psi(q)$ of the interval $(-L/2, +L/2)$ can be represented by the series

$$\psi(q) = \sum_{n=-\infty}^{+\infty} c_n \, u_n, \tag{V.46}$$

with

$$c_n = \langle u_n, \psi(q)\rangle = L^{-\frac{1}{2}} \int_{-L/2}^{+L/2} \mathrm{e}^{-\mathrm{i}p_n q/\hbar} \psi(q) \, \mathrm{d}q.$$

Applying the theory of §§ 5 and 6 it is found that the probability of finding $p=p_n$ is equal to $|c_n|^2$.

In the limit $L \to \infty$, the distance

$$\varepsilon = \frac{2\pi\hbar}{L}$$

between neighboring eigenvalues goes to zero and the momentum spectrum $p_n = n\varepsilon$ becomes continuous. The study of the limiting process is entirely analogous to that carried out for q. When $\varepsilon \to 0$, $p' = n\varepsilon$ remaining constant, $\varepsilon^{-\frac{1}{2}} c_n$ tends toward the Fourier transform

$$\varphi(p') = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} e^{-\frac{i}{\hbar}p'q/\hbar} \psi(q)\, dq.$$

We leave it to the reader to find, by this limiting process, the statistical distribution of the momentum measurements, and to show how the representation of $\psi(q)$ by the Fourier series (V.46) tends toward the Fourier integral

$$\psi(q) = \int_{-\infty}^{+\infty} \varphi(p')\, u(p'; q)\, dp'$$

where $u(p'; q)$ is the limit of $\varepsilon^{-\frac{1}{2}} u_n$, that is to say

$$u(p'; q) = \frac{e^{\frac{i}{\hbar}p'q/\hbar}}{\sqrt{2\pi\hbar}}.$$

12. Comments and Examples

To summarize, we have shown how the statistical distribution of the results of measurement of a dynamical variable can be deduced from the knowledge of its wave function. It is particularly satisfying that this result can be stated in an entirely general manner. With each dynamical variable, one associates a Hermitean operator possessing a complete orthonormal set of eigenfunctions. Starting from a very natural and formally very simple postulate about average values, we have shown that the only possible results of measurement are the eigenvalues of the observable A, and that the desired probability law is directly related to the squares of the moduli of the expansion coefficients of the wave function in a series of eigenfunctions of A.

Among the dynamical variables most commonly considered one should mention, besides the position and momentum coordinates, the energy represented by the Schrödinger Hamiltonian, and the angular momentum.

The *energy* spectrum can be, depending upon the case, either purely discrete (cf. Ch. III, § 5), purely continuous (cf. Ch. III, § 3), or mixed (cf. Ch. III, § 6). The eigenvalue problem of the Hamiltonian H is important in Quantum Theory not only in relation to the definition of the energy, but also because of the role it plays in the solution of the Schrödinger equation. *When H is time-independent*, the only case where the concept of energy has a real meaning, the wave function $\Psi(t)$ at time t is deduced from the wave function $\Psi(t_0)$ at the initial time t_0 by the operation

$$\Psi(t) = e^{-iH(t-t_0)/\hbar}\, \Psi(t_0).$$

One knows how to calculate the right-hand side of this equation by using the expansion of $\Psi(t_0)$ in a series of eigenfunctions of H [this is a special case of eq. (V.41)]; it is easily shown, making use of this expansion, that

$$\frac{\partial}{\partial t}\Psi(t) = -\frac{i}{\hbar} H\, e^{-iH(t-t_0)/\hbar}\, \Psi(t_0) = -\frac{i}{\hbar} H\Psi(t),$$

hence that $\Psi(t)$ satisfies the Schrödinger equation, and furthermore that the initial conditions are actually fulfilled when $t = t_0$.

Suppose, for simplicity, that the spectrum of H is entirely discrete and non-degenerate; denote by E_n ($n = 1, 2, \ldots$) the eigenvalues of H, and by ψ_n the corresponding eigenfunctions. The expansion of $\Psi(t_0)$ is written

$$\Psi(t_0) = \sum_n c_n \psi_n, \qquad c_n = \langle \psi_n, \Psi(t_0) \rangle.$$

The function $\Psi(t)$ is then given by

$$\Psi(t) = \sum_n c_n\, e^{-iE(t-t_0)/\hbar}\, \psi_n. \tag{V.47}$$

Note that the absolute value of the coefficient of ψ_n in this expansion is independent of t. Hence the interesting property: *the statistical distribution of the energy of a system* (whose Hamiltonian is time-independent) *is constant in time.*

The *angular momentum* of a particle in Classical Mechanics is the

vector $\boldsymbol{r} \times \boldsymbol{p}$. In Quantum Theory, there corresponds to it the vector operator

$$\boldsymbol{l} \equiv \frac{\hbar}{\mathrm{i}} (\boldsymbol{r} \times \nabla). \tag{V.48}$$

Let us write one of its components explicitly:

$$l_z = \frac{\hbar}{\mathrm{i}} \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right).$$

Take Oz as polar axis and denote by (r, θ, φ) the polar coordinates of the particle; one easily verifies that $\partial/\partial\varphi = x\partial/\partial y - y\partial/\partial x$. Consequently,

$$l_z = \frac{\hbar}{\mathrm{i}} \frac{\partial}{\partial \varphi}. \tag{V.49}$$

The eigenvalue problem of l_z appears particularly simple if one expresses the eigenfunction in polar coordinates

$$l_z \psi(r, \theta, \varphi) \equiv \frac{\hbar}{\mathrm{i}} \frac{\partial}{\partial \varphi} \psi(r, \theta, \varphi) = l_z' \psi(r, \theta, \varphi).$$

We have

$$\psi(r, \theta, \varphi) = F(r, \theta)\, \mathrm{e}^{\mathrm{i} l_z' \varphi / \hbar},$$

where $F(r, \theta)$ is an arbitrary function of r and θ. Since the eigenfunction is a single-valued function of $\boldsymbol{r}$, $\psi(r, \theta, \varphi)$ must take on the same value when one changes φ to $\varphi + 2\pi$, hence

$$l_z' = m\hbar \qquad (m \text{ integer}). \tag{V.50}$$

Consequently, the eigenvalue spectrum of any component of the angular momentum of a particle is an entirely discrete spectrum. This result is easily extended to the components of the total angular momentum of a system of particles, in good agreement with the experimental fact of space quantization.

IV. DETERMINATION OF THE WAVE FUNCTION

13. Measuring Process and "Filtering" of the Wave Packet. Ideal Measurements

The statistical distributions defined in the preceding sections are directly amenable to experiment. Thus the distribution associated with a given dynamical variable $\mathscr{A}$ is the distribution of the results

obtained when one carries out the measurement of $\mathscr{A}$ on a very large number $\mathscr{N}$ of identical systems, independent of each other and in the same dynamical state at the instant of measurement. Each system is represented at that instant [1]) by the same wave function Ψ (defined to within a constant) to which there corresponds a well-defined theoretical distribution. The latter may be compared with the distribution observed experimentally.

For the physical interpretation of the theory to be complete, one must further specify:

(*i*) how previous observations, made on a system, allow us to know its dynamical state, and, in particular, how one can be sure that the $\mathscr{N}$ systems considered above are actually all in the dynamical state represented by Ψ;

(*ii*) what becomes of each of these systems once the measurement is completed.

These two questions are closely related. We shall deal with the second one first.

In contrast to what happens during the process of measurement [2]), the system, once the measurement is completed, can again be treated as an entity completely separated from the measuring apparatus. It becomes possible again to describe it by means of a wave function involving only its own dynamical variables. This wave function of the system after measurement is certainly different from the wave function immediately before the measurement, except perhaps if the latter is an eigenfunction of the observable A associated with the measured quantity. We shall call this (*non-causal*) change of the wave function by the measurement process the *filtering of the wave packet.*

We know that this non-causal change appears as an uncontrollable perturbation in the motion of the system by its interaction with the measuring device; its effect is roughly to make the complementary

[1]) The instant in question here is the *instant at which the measurement begins.* Actually, as soon as the measurement has started, the system is in interaction with the measuring device and any description of its evolution by means of its wave function alone becomes impossible.

[2]) The detailed study of the mechanism of measurement will not be made in this book. On that subject, see the references cited in the footnote of p. 157.

variable to the one measured the less well-defined the more precise the measurement that is carried out. This uncontrollable perturbation should not be confused with any modifications – in principle exactly calculable – which the system might undergo during the measuring process. In particular, it often happens that the measured variable is modified during the measuring process. The two momentum measurements of Chapter IV are examples of such a situation. The modification of the measured quantity evidently depends upon the type of measuring device used. Thus in the momentum measurement by Compton collision contemplated in Chapter IV, p. 147, the difference $p' - p$ is smaller the lower the frequency ν of the photon; it vanishes if ν is chosen infinitely small. Clearly, no general statement can be made concerning the modification of the dynamical state of the system during the measuring operation, since this modification depends in each case upon the particular conditions under which the measurement was carried out. One can, however, imagine idealized conditions of measurement, in which all exactly calculable modifications, which we have just mentioned, are strictly compensated, and where only the specific, uncontrollable perturbation of the quantum phenomena enters. We shall assume that such *ideal measurements* can actually be realized, or, at the very least, that they are limiting cases of actually realizable measurements (for instance: position measurements discussed in Chapter IV in the limit where the duration of the measurement is zero; momentum measurement by Compton collision in the limit where $\nu = 0$).

Consider therefore an ideal measurement of the quantity $\mathscr{A}$, and suppose at the start that the value found a_i is a non-degenerate eigenvalue. According to our hypothesis, we know with certainty that $\mathscr{A} = a_i$ once the measurement is completed, hence that the wave function of the system is the eigenfunction ψ_i (to within a constant) corresponding to the eigenvalue a_i. The arbitrary constant has no physical significance since the statistical distribution of the results of any subsequent measurement is independent of the choice of that constant. The wave function of the system after measurement is thus known without ambiguity. The measuring device works in some sense like a "perfect filter". The wave function before measurement is a function $\Psi = \sum_n c_n \psi_n$. There is a probability $|c_i|^2$ that the result of measurement is a_i. Assuming that the measurement gives a_i, the net effect of the measuring process is to "pass" (without distortion)

only the term $c_i\psi_i$ of the expansion of Ψ in a series of eigenfunctions of A.

In the more general case where the eigenvalues of the observable A, and especially the value a_i, are degenerate, the wave function before measurement (assumed normalized to unity) can be put in the form [cf. eq. (V.20)]

$$\Psi = \sum_p \Psi_p, \qquad \Psi_p = \sum_r c_p^{(r)} \psi_p^{(r)}. \tag{V.51}$$

There is a probability $\langle \Psi_i, \Psi_i \rangle = \sum_r |c_p^{(r)}|^2$ of obtaining the result a_i. In the case of an ideal measurement, the wave function after measurement is an eigenfunction of A corresponding to the eigenvalue a_i: it is a linear combination of the functions $\psi_i^{(r)}$ (r is variable). However, this information is not sufficient to specify it completely; it is less so the higher the degree of degeneracy. The ideal measurement is one where the measuring device acts as a "perfect filter" and "passes" without distortion and to the exclusion of all the rest, the portion of the expansion (V.51) of Ψ pertaining to eigenvalue a_i, i.e. the function

$$\Psi_i = \sum_r c_i^{(r)} \psi_i^{(r)}.$$

When the measurement is not ideal, the "passing" of these terms is accompanied by some distortion. That distortion is in principle exactly calculable and depends upon the measuring device used.

In the remainder of this chapter, all contemplated measurements will be assumed ideal. All discussions are thereby greatly simplified. Actually, such a restriction is not essential to the argument, and one could relax it without bringing about fundamental changes in the physical interpretation of the theory.

14. Commuting Observables and Compatible Variables

Consider two observables A and B. For simplicity the eigenvalue spectra are assumed to be entirely discrete, although the properties stated below are quite general. Let us suppose that they possess a common eigenfunction Ψ_0

$$A\Psi_0 = a\Psi_0$$
$$B\Psi_0 = b\Psi_0.$$

The physical meaning of these two equations is the following: if the

physical system is in state Ψ_0 at a given time t, a precise measurement of the quantities $\mathscr{A}$ and $\mathscr{B}$ will yield with certainty the results a and b, respectively. A necessary condition for these two equations to hold simultaneously is that

$$(AB-BA)\Psi_0=[A,\,B]\Psi_0=0; \tag{V.52}$$

that is to say, that the commutator of A and B have Ψ_0 as eigenfunction corresponding to the eigenvalue 0.

An example in which this condition never holds is given by the observables x and p_x, since their commutator is a non-zero constant. More precisely [cf. eq. (II.10)]

$$[x,\,p_x]=\frac{\hbar}{\mathrm{i}}\left[x,\frac{\partial}{\partial x}\right]=\mathrm{i}\hbar\neq 0, \tag{V.53}$$

and it is indeed well known that these two quantities can never be defined simultaneously with infinite precision.

On the other hand, equation (V.52) is automatically satisfied when the observables A and B commute. In that case one has the important theorem:

If two observables commute, they possess a complete orthonormal set of common eigenfunctions, and conversely.

Physically, this means that the dynamical variables represented by these two observables may be defined simultaneously in a precise way: they are *compatible variables*. In particular, it is possible to carry out simultaneously an ideal measurement of the variables $\mathscr{A}$ and $\mathscr{B}$; in that case, the wave function after measurement is an eigenfunction common to A and B.

The proof of the direct theorem is as follows. Assume that the observables A and B commute:

$$[A,\,B]=0.$$

Let ψ_a be an eigenfunction of A, and a its eigenvalue. ψ_a can be expanded into a complete orthonormal set of eigenfunctions of B. It can therefore be written in the form

$$\psi_a=\sum_m \varphi(a;\,b_m),$$

where $\varphi(a;\,b_m)$ is an eigenfunction of B corresponding to the eigenvalue b_m. One can always arrange matters in such a way that every function

occurring in that sum corresponds to a different eigenvalue [cf. eq. (V.20)]. Let us show that

$$\hat{\varphi}_m \equiv (A-a)\,\varphi(a;b_m)=0.$$

Since A and B commute, we have

$$B\hat{\varphi}_m=(A-a)\,B\varphi(a;b_m)=(A-a)\,b_m\varphi(a;b_m)=b_m\hat{\varphi}_m.$$

The functions $\hat{\varphi}_m$ are therefore eigenfunctions of B; since the corresponding eigenvalues are all different, these functions are linearly independent. However, one has

$$\sum_m \hat{\varphi}_m=(A-a)\psi_a=0.$$

This is possible only if each of the functions $\hat{\varphi}_m$ vanishes. In other words, the functions $\varphi(a;b_m)$ are simultaneously eigenfunctions of A and of B.

Consider now a *complete* orthonormal set $\{\psi_n^{(r)}\}$ of eigenfunctions of A:

$$A\psi_n^{(r)}=a_n\psi_n^{(r)}.$$

According to the above argument, these functions can be put in the form

$$\psi_n^{(r)}=\sum_m \varphi^{(r)}(a_n;b_m). \tag{V.54}$$

The functions $\varphi^{(r)}(a_n;b_m)$ are eigenfunctions common to A and B. The ensemble of functions $\varphi^{(r)}(a_n;b_m)$ corresponding to the same pair of eigenvalues a_n, b_m may not be linearly independent. But it is always possible to choose (by means of the orthogonalization process of Schmidt, for instance) a set of orthonormal functions $\chi^{(s)}(a_n;b_m)$ corresponding to the same pair of eigenvalues and such that the functions $\varphi^{(r)}(a_n;b_m)$ are linear combinations of these functions

$$\varphi^{(r)}(a_n;b_m)=\sum_s c_{rs}\,\chi^{(s)}(a_n;b_m). \tag{V.55}$$

The ensemble $\{\chi\}$ of all these functions constitutes an orthonormal set of eigenfunctions common to A and B. Moreover, it is a complete set, since any wave function Ψ can be expanded in a series of χ; to form this expansion, it suffices to expand Ψ in a series of functions of the complete set $\{\psi_n^{(r)}\}$, and then to substitute for the functions

$\psi_n^{(r)}$ in this expansion, their expressions as function of the $\chi^{(s)}(a_n; b_m)$ obtained with the aid of eqs. (V.54) and (V.55). Q.E.D.

Conversely, if A and B possess a complete orthonormal set of common eigenfunctions $\chi^{(s)}(a_n; b_m)$, one has

$$AB\chi^{(s)}(a_n; b_m) = a_n b_m \chi^{(s)}(a_n; b_m) = BA\chi^{(s)}(a_n; b_m),$$

hence

$$[A, B]\, \chi^{(s)}(a_n; b_m) = 0.$$

The action of the commutator $[A, B]$ on all functions of the set $\{\chi\}$ yields zero. Since by hypothesis all wave functions Ψ can be developed in a series of the χ, one has $[A, B]\, \Psi = 0$, for any Ψ. Therefore

$$[A, B] = 0.$$

Using the commuting observables A, B one can build up new observables of the form $f(A, B)$, $f(x, y)$ being an arbitrarily chosen real function. By definition, the action of $f(A, B)$ on an eigenfunction $\chi(a; b)$ common to A and B yields

$$f(A, B)\, \chi(a; b) = f(a, b)\, \chi(a; b).$$

Its action on an arbitrary function Ψ is obtained by expanding Ψ in a series of the χ's and applying the operator $f(A, B)$ to each of the terms of the expansion. This procedure is valid provided the series converges; otherwise the function $f(A, B)\, \Psi$ does not exist. It is evident from the way in which it was defined, that $f(A, B)$ possesses, along with A and B, a complete orthonormal set of common eigenfunctions, namely the set $\{\chi\}$; hence $f(A, B)$ commutes with A and B.

All these results are easily extended to an arbitrary number R of pairwise commuting observables. If R observables commute in pairs, they possess (at least) a complete orthonormal set of common eigenfunctions, and conversely. Furthermore, any (real) function of these observables is an observable which commutes with each of them and possesses the same set of eigenfunctions.

15. Complete Sets of Commuting Observables

Consider an observable A. With its eigenfunctions one can build up a complete orthonormal set of eigenfunctions; we shall henceforth

call it the *basis of* A. In general, this basis is not unique. The arbitrary part in the choice of this basis was discussed in § 9. We shall by convention consider two bases as identical if their functions differ only in the phase and (in the case of the continuous spectrum) in the norm. With this convention, the basis of A is unique if none of its eigenvalues are degenerate. On the other hand, suppose, to be specific, that the eigenvalue a is doubly degenerate, and let ψ_1, ψ_2 be two orthonormal eigenfunctions corresponding to that eigenvalue. As can be easily verified, the functions

$$\varphi_1 = \psi_1 \cos \alpha + \psi_2 \sin \alpha$$
$$\varphi_2 = -\psi_1 \sin \alpha + \psi_2 \cos \alpha$$

also have that property. Thus one can form the basis of A just as well with the pair (φ_1, φ_2) as with the pair (ψ_1, ψ_2), Now, assume that the basis of A is not unique. Let B be an observable that commutes with A. It may happen that the basis common to A and B, whose existence we have demonstrated in § 14, is unique. One then says that the observables A and B form a complete set of commuting observables.

If A and B do not share a unique common basis, one is led to associate with them a third observable C which commutes with both of them, and so forth.

More generally, one says that *the observables* $A, B, \ldots, L$ *form a complete set of commuting observables if they possess one and only one common basis.* In that case, any observable which commutes with each of the observables of the set, necessarily has this set as its basis. Its eigenvalues are therefore well-defined functions of the eigenvalues $a, b, \ldots, l$ of the observables $A, B, \ldots, L$. In other words, this observable can be considered as a function of the observables of the set.

The dynamical variables represented by the observables of a complete set of commuting observables, can always be all simultaneously defined with precision; moreover, they form a *complete set of compatible variables* (Ch. IV, § 17). If one carries out simultaneously a precise measurement of the values assumed by these variables, one can be sure that the wave function of the system is an eigenfunction of the observables $A, B, \ldots, L$ corresponding to the eigenvalues $a, b, \ldots, l$ found in the operation of measurement. *Since there exists but one eigenfunction possessing this property, the specification of these*

measurements completely defines the wave function of the physical system. One says that *the dynamical state of the system is completely specified by giving the quantum numbers a, b, ..., l.* In fact, this function is only determined to within a constant. Since the only physically measurable quantities, that is to say the statistical distributions of the results of various possible measurements, are independent of the choice of this constant, one can fix the latter as one wishes without changing the physical meaning of the wave function. If, as is often done, the function is normalized to unity, there still remains an arbitrary phase without physical significance.

The time sequence of an experiment in physics can be viewed in the following manner. At the initial time t_0 one *prepares* the system by performing on it the simultaneous measurement of a complete set of compatible variables. Its dynamical state is thus completely determined at time t_0.

Once this preparation is completed, the wave function of the system evolves in time in a manner exactly determined by the Schrödinger equation. At all later instants, the dynamical state of the system is thus perfectly known, at least as long as it is not perturbed by the intervention of a measuring device. Eventually, at a later time t, one carries out a given measurement. Since one knows the wave function $\Psi(t)$ at the instant where the measurement is performed, one can exactly predict the statistical distribution of the results of the measurement. Repetition of this experiment on a very large number $\mathcal{N}$ of identical systems, yields an experimental distribution which can be compared with the theoretical distribution.

16. Pure States and Mixtures

In practice, a complete "preparation" of the system like the one contemplated above, is rarely achieved. Most frequently, the dynamical variables measured during the preparation do not constitute a complete set. As a consequence, the dynamical state of the system is known incompletely and one must have recourse to the methods of statistics. Instead of a given dynamical state, one is dealing with a statistical mixture of states; instead of assigning to the system a well-defined wave function, one assigns to it a statistical mixture of wave functions each having a suitable statistical weight. There exists a *Quantum Statistical Mechanics*, just as there is a Classical Statistical Mechanics.

When the preparation is complete, and consequently the dynamical state of the system is exactly known, one says that one is dealing with a *pure state*, in contrast to the statistical mixtures which characterize incomplete preparations.

In the prediction of the results of measurements carried out on a mixture, statistics enters in two ways: first for specific quantum reasons connected with the uncontrollable perturbation of the system in the measuring operation, and secondly because the dynamical state of the system is incompletely known.

Suppose that at the time of preparation t_0, the system could be represented by the set of wave functions $\Psi^{(1)}(t_0), \ldots, \Psi^{(k)}(t_0), \ldots$, with statistical weights $p_1, \ldots, p_k, \ldots$ ($\sum_k p_k = 1$), respectively. Let $\Psi^{(1)}(t), \ldots, \Psi^{(k)}(t), \ldots$, be the solutions of the Schrödinger equation corresponding to the initial conditions $\Psi^{(1)}(t_0), \ldots, \Psi^{(k)}(t_0), \ldots$ respectively. At time t, the system is represented by the set of these functions $\Psi^{(1)}(t), \ldots, \Psi^{(k)}(t), \ldots$, with the same statistical weights $p_1, \ldots, p_k, \ldots$. Let $\langle A \rangle_k$ be the average value of the results of measurement of a given quantity $\mathscr{A}$, performed on the system when it is in the dynamical state $\Psi^{(k)}(t)$:

$$\langle A \rangle_k = \frac{\langle \Psi^{(k)}(t), A\Psi^{(k)}(t) \rangle}{\langle \Psi^{(k)}(t), \Psi^{(k)}(t) \rangle}.$$

The average value of the results of measurement of $\mathscr{A}$ carried out on the statistical mixture at time t is given by

$$\langle A \rangle = \sum_k p_k \langle A \rangle_k.$$

Similarly, if $w_i^{(k)}$ is the probability of finding the result a_i when the dynamical state of the system is represented by $\Psi^{(k)}(t)$, the probability of finding this result when carrying out the same measurement on the mixture at time t is

$$w_i = \sum_k p_k w_i^{(k)}. \tag{V.56}$$

A very important case of quantum-statistical mixture is that of a system in thermodynamic equilibrium with a heat reservoir at temperature T. The various possible dynamical states are the eigenstates of the Hamiltonian H of the system. The statistical weight of a given eigenstate depends only upon the corresponding eigenvalue of H; it is proportional to the Boltzmann factor $\exp(-E/kT)$, in which E is the eigenvalue of H and k is the Boltzmann constant.

V. COMMUTATOR ALGEBRA AND ITS APPLICATIONS

17. Commutator Algebra and Properties of Basic Commutators

As long as one deals only with commuting observables the rules of ordinary algebra may be used without restriction. However, the observables of a given quantum system do not all commute. More precisely, the observables of a quantum system in R dimensions are functions of the position observables q_i $(i=1, 2, \ldots, R)$ and the momentum observables p_i $(i=1, 2, \ldots, R)$ [1]), all pairs of which do not commute. The commutators of the q's and the p's play a fundamental role in the theory. One has:

$$[q_i, q_j]=0, \qquad [p_i, p_j]=0 \tag{V.57}$$

$$[q_i, p_j]=i\hbar\,\delta_{ij}. \tag{V.58}$$

Relations (V.57) are obvious; in particular the second merely states that operations of differentiation commute with each other. Relation (V.58) is a generalization of eq. (V.53); it is readily obtained by using the explicit form of the operators p:

$$p_i=\frac{\hbar}{i}\frac{\partial}{\partial q_i}.$$

From the fact that the q's and the p's do not commute in pairs, the precise definition of a dynamical variable $\mathscr{A}\equiv A(q_1, \ldots, q_R; p_1, \ldots, p_R)$ requires that one properly specifies the order of the q's and the p's in the explicit expression of the function $A(q_1, \ldots, q_R; p_1, \ldots, p_R)$. In practice, A is put in the form of a polynomial in p – or possibly in the form of a power series in p – whose coefficients are functions of q. Each term is a product of components p_i and functions of the q arranged in a certain order. The function A, considered as an operator, is well defined only when the order in each of its terms is specified.

It is interesting to know the commutators of the q's and the p's with a given function A. For functions of the q's alone, or of the p's

---

[1]) Actually, this is true only to the extent that the quantum system has a classical analogue. In what follows we shall come to introduce supplementary variables, namely the spin variables which have no classical analogue.

alone, one obtains the relations

$$[q_i, F(q_1, \ldots, q_R)] = 0 \tag{V.59}$$

$$[p_i, G(p_1, \ldots, p_R)] = 0 \tag{V.60}$$

$$[p_i, F(q_1, \ldots, q_R)] = \frac{\hbar}{\mathrm{i}} \frac{\partial F}{\partial q_i} \tag{V.61}$$

$$[q_i, G(p_1, \ldots, p_R)] = \mathrm{i}\hbar \frac{\partial G}{\partial p_i}. \tag{V.62}$$

The relations (V.59) and (V.60) are particular cases of the property stated at the end of § 14. To prove equation (V.61), it suffices to write down the operator p_i explicitly and to verify that the action of each side of the equation on an arbitrary wave function gives the same result [cf. eq. (II.9)]. Equation (V.62) is proved by making an analogous verification in momentum space; let us recall that if $\Phi(p_1, \ldots, p_R)$ is the wave function of momentum space corresponding to $\Psi(q_1, \ldots, q_R)$, the function of momentum space corresponding to $q_i\Psi(q_1, \ldots, q_R)$ is

$$\mathrm{i}\hbar \frac{\partial}{\partial p_i} \Phi(p_1, \ldots, p_R).$$

One arrives at the same result using the rules of *commutator algebra*. Let us give here the four principal rules. These rules are direct consequences of the definition of commutators. Their proofs are left to the reader. If A, B, and C denote three arbitrary linear operators, one has

$$[A, B] = -[B, A] \tag{V.63}$$

$$[A, B+C] = [A, B] + [A, C] \tag{V.64}$$

$$[A, BC] = [A, B]C + B[A, C] \tag{V.65}$$

$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0. \tag{V.66}$$

By repeated application of rule (V.65), one has

$$[A, B^n] = \sum_{s=0}^{n-1} B^s [A, B] B^{n-s-1}.$$

In particular, for a one-dimensional system one has

$$[q, p^n] = n\,\mathrm{i}\hbar\, p^{n-1}.$$

Equation (V.62) is thus verified when F is an arbitrary power of the p; it is thus also verified [rule (V.64)] when F is a polynomial, or else a convergent power series in p.

For general functions of the q's and p's, one can also write

$$[p_i, A] = \frac{\hbar}{i} \frac{\partial A}{\partial q_i} \tag{V.67}$$

$$[q_i, A] = i\hbar \frac{\partial A}{\partial p_i}, \tag{V.68}$$

$\partial A/\partial q_i$, $\partial A/\partial p_i$ being defined by partial differentiation of A, it being understood that the order of the p's and q's in their explicit expression has been suitably chosen.

Let us illustrate this by an example dealing with one-dimensional quantum systems. Let $f(q)$ be a function of q. The three commutators of q and of each of the functions $p^2 f(q)$, $pf(q)p$, and $f(q)p^2$ may all be identified (to within the factor $i\hbar$) with the derivative with respect to p of these functions, but they are not the same operators. Indeed, by repeated application of rule (V.62),

$$[q, p^2 f(q)] = 2i\hbar\, pf(q)$$
$$[q, pfp] = i\hbar(fp + pf)$$
$$[q, fp^2] = 2i\hbar\, fp.$$

In the same way

$$[p, p^2 f] = \frac{\hbar}{i} p^2 f'$$

$$[p, pfp] = \frac{\hbar}{i} pf'p$$

$$[p, fp^2] = \frac{\hbar}{i} f'p^2.$$

18. Commutation Relations of Angular Momentum

As an application of rules (V.63) to (V.65) of commutator algebra, let us calculate the commutators of the components of the angular momentum of a particle:

$$\boldsymbol{l} \equiv \boldsymbol{r} \times \boldsymbol{p}.$$

One has

$$\begin{aligned}[l_x, l_y] &= [yp_z - zp_y, zp_x - xp_z] \\ &= [yp_z, zp_x] + [zp_y, xp_z] && \text{[rule (V.64)]} \\ &= y[p_z, z]p_x + p_y[z, p_z]x && \text{[rule (V.65)]} \\ &= i\hbar(xp_y - yp_x) \\ &= i\hbar\, l_z.\end{aligned}$$

The other two commutators are calculated by cyclic permutation. Thus

$$[l_x, l_y] = i\hbar\, l_z, \qquad [l_y, l_z] = i\hbar\, l_x, \qquad [l_z, l_x] = i\hbar\, l_y. \tag{V.69}$$

The three components of the angular momentum do not commute in pairs. There is no complete orthonormal set common to any two of them. In other words, two components of angular momentum cannot, in general [1]), be defined simultaneously with infinite precision.

Note that [rule (V.65)]

$$\begin{aligned}[l_z, l_x^2] &= i\hbar(l_y l_x + l_x l_y) \\ [l_z, l_y^2] &= -i\hbar(l_y l_x + l_x l_y) \\ [l_z, l_z^2] &= 0.\end{aligned}$$

Adding term by term [rule (V.64)], we obtain

$$[l_z, \boldsymbol{l}^2] = 0, \tag{V.70}$$

where the operator

$$\boldsymbol{l}^2 = l_x^2 + l_y^2 + l_z^2 \tag{V.71}$$

is the square of the length of the vector $\boldsymbol{l}$.

The operators $\boldsymbol{l}^2$ and l_z commute: they can therefore be simultaneously defined with infinite precision. The pairs $(\boldsymbol{l}^2, l_x)$ and $(\boldsymbol{l}^2, l_y)$ obviously possess the same property.

---

[1]) The words "in general" are important. The three components have no common basis, but have common eigenfunctions, those for which $l_x = l_y = l_z = 0$. These are the functions which depend only upon the length $r = \sqrt{x^2 + y^2 + z^2}$ and not upon the direction of the vector $\boldsymbol{r}$.

19. Time Dependence of the Statistical Distribution. Constants of the Motion

Consider the Schrödinger equation and the complex conjugate equation:

$$i\hbar \frac{\partial \Psi}{\partial t} = H\Psi, \qquad i\hbar \frac{\partial \Psi^*}{\partial t} = -(H\Psi)^*.$$

If Ψ is normalized to unity at the initial instant, it remains normalized at any later time. The mean value of a given observable A is equal at every instant to the scalar product

$$\langle A \rangle = \langle \Psi, A\Psi \rangle = \int \Psi^* A \Psi \, d\tau,$$

and one has

$$\frac{d}{dt}\langle A \rangle = \left\langle \frac{\partial \Psi}{\partial t}, A\Psi \right\rangle + \left\langle \Psi, A \frac{\partial \Psi}{\partial t} \right\rangle + \left\langle \Psi, \frac{\partial A}{\partial t} \Psi \right\rangle.$$

The last term of the right-hand side, $\langle \partial A/\partial t \rangle$, is zero if A does not depend upon the time explicitly.

Taking into account the Schrödinger equation and the hermiticity of the Hamiltonian, one has

$$\begin{aligned} \frac{d}{dt}\langle A \rangle &= -\frac{1}{i\hbar}\langle H\Psi, A\Psi \rangle + \frac{1}{i\hbar}\langle \Psi, AH\Psi \rangle + \left\langle \frac{\partial A}{\partial t} \right\rangle \\ &= \frac{1}{i\hbar}\langle \Psi, [A, H]\Psi \rangle + \left\langle \frac{\partial A}{\partial t} \right\rangle. \end{aligned}$$

Hence we obtain the general equation giving the time-dependence of the mean value of A:

$$i\hbar \frac{d}{dt}\langle A \rangle = \langle [A, H] \rangle + i\hbar \left\langle \frac{\partial A}{\partial t} \right\rangle. \tag{V.72}$$

When we replace A by the operator $\exp(i\xi A)$, we obtain an analogous equation for the time-dependence of the characteristic function of the statistical distribution of A.

In particular, for any variable C which commutes with the Hamiltonian

$$[C, H] = 0$$

and which does not depend explicitly upon the time, one has the result

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle C\rangle = 0.$$

The mean value of C remains constant in time. More generally, if C commutes with H, the function $\exp(\mathrm{i}\xi C)$ also commutes with H, and, consequently

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle \mathrm{e}^{\mathrm{i}\xi C}\rangle = 0.$$

The characteristic function, and hence the statistical distribution of the observable C, remain constant in time.

By analogy with Classical Analytical Mechanics, C is called a *constant of the motion*. In particular, if at the initial instant the wave function is an eigenfunction of C corresponding to a given eigenvalue c, this property continues to hold in the course of time. One says that c is a "good quantum number". If, in particular, H does not explicitly depend upon the time, and if the dynamical state of the system is represented at time t_0 by an eigenfunction common to H and C, the wave function remains unchanged in the course of time, to within a phase factor. The energy and the variable C remain well defined and constant in time.

20. Examples of Constants of the Motion. Energy. Parity

There exists an observable which always commutes with the Hamiltonian: the Hamiltonian itself. The energy is therefore a constant of the motion of all systems whose Hamiltonian does not depend explicitly upon the time. This result was already shown in § 12.

As another possible constant of the motion, let us mention *parity* (cf. Ch. III, § 14). We denote under the name of parity the observable P defined by

$$P\psi(q) = \psi(-q).$$

It is easily verified that P is Hermitean. Moreover, $P^2 = 1$ and, consequently, the only possible eigenvalues of P are $+1$ and -1; even functions are associated with $+1$, and odd functions with -1.

When the Hamiltonian is invariant under the substitution of $-q$ *for* q, we obviously have

$$[P, H]=0.$$

Indeed, if

$$H\left(\frac{\hbar}{\mathrm{i}}\frac{\mathrm{d}}{\mathrm{d}q}, q\right) = H\left(-\frac{\hbar}{\mathrm{i}}\frac{\mathrm{d}}{\mathrm{d}q}, -q\right),$$

one has, for any $\psi(q)$,

$$PH\psi = H\left(-\frac{\hbar}{\mathrm{i}}\frac{\mathrm{d}}{\mathrm{d}q}, -q\right)\psi(-q) = H\left(\frac{\hbar}{\mathrm{i}}\frac{\mathrm{d}}{\mathrm{d}q}, q\right)\psi(-q) = HP\psi.$$

Under these conditions, if the wave function has a definite parity at a given initial instant of time, it conserves the same parity in the course of time.

This property is easily extended to a system having an arbitrary number of dimensions; in particular, it applies to systems of particles for which the parity operation amounts to a reflection in space ($\boldsymbol{r}_i \to -\boldsymbol{r}_i$) and for which the observable parity is defined by

$$P\Psi(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots) = \Psi(-\boldsymbol{r}_1, -\boldsymbol{r}_2, \ldots).$$

EXERCISES AND PROBLEMS

1. Prove the Schwarz inequality $|\langle\varphi, \psi\rangle| \leqslant \sqrt{\langle\varphi, \varphi\rangle\langle\psi, \psi\rangle}$ from the properties (*a*), (*b*) and (*c*) of the scalar product (§ 2). [Write down that the norm of an arbitrary linear combination of φ and ψ is necessarily positive or zero]. Show that the equality between the two sides holds if, and only if the functions φ and ψ are multiples of each other.

2. Consider the eigenvalue problem of the operator $p = (\hbar/\mathrm{i})\mathrm{d}/\mathrm{d}q$ acting on the functions $\psi(q)$ defined in the interval $(-\infty, +\infty)$. Verify that the spectrum is continuous and that the eigenfunctions have an infinite norm. Show that by superposition of eigenfunctions corresponding to neighboring eigenvalues, one can build up functions of finite norm (eigendifferentials) for which the root-mean-square deviation Δp can be made as small as desired. Consider the same question for the operator

$$\frac{p^2}{2m} = -\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}q^2}.$$

3. Consider the operator q acting on the functions $\psi(q)$ of finite norm in the interval $(-\infty, +\infty)$. (*a*) Form a sequence of continuous functions depending

on a parameter η such that their norm and the mean value $\langle q \rangle$ remain independent of η, and that the root-mean-square deviation Δq vanishes in the limit $\eta \to 0$ (there exists a large number of sequences of this type). One notes that this sequence does not converge toward a function of Hilbert space. (*b*) In the same way, build the sequence of eigendifferentials (they are functions of finite norm, but are not continuous) depending on the parameter δq and obtained by superposing the eigen-"functions" $\delta(q - q_0)$ corresponding to eigenvalues q_0 located in the interval δq; verify that the root-mean-square deviation Δq corresponding to these eigendifferentials also vanishes in the limit $\delta q \to 0$.

4. Verify that the statistical distribution of the measurements of a given quantity is uniquely defined, despite the large degree of arbitrariness in the choice of the complete orthonormal set of eigenfunctions associated with the operator which represents this physical quantity (arbitrariness in the choice of the phases of the eigenfunctions; in the choice of the eigenfunctions of the same degenerate eigenvalue; in the normalization of the functions of the continuous spectrum).

5. How can one extend equations (V.20) and (V.21) of this chapter to the case of the continuous spectrum?

6. In polar coordinates (r, θ, φ) the component l_z of the angular momentum of a particle takes the form $(\hbar/i)\partial/\partial\varphi$. This suggests that l_z and φ form a pair of complementary variables. However, the uncertainty relation $\Delta l_z \Delta\varphi \gtrsim \hbar$ has no meaning. Explain why. Show that when $\Delta l_z = 0$, the angle φ is completely indeterminate. Expand the wave packet

$$\psi(\varphi) = \sum_{n=-\infty}^{+\infty} \exp\left[-\sigma(\varphi - \varphi_0 + 2n\pi)^2\right]$$

in a series of eigenfunctions of l_z. Compare the uncertainty in φ and the uncertainty in l_z in this example and discuss the complementary character of these two variables.

CHAPTER VI

CLASSICAL APPROXIMATION AND THE WKB METHOD

I. THE CLASSICAL LIMIT OF WAVE MECHANICS

1. General Remarks

In the limit where $\hbar \to 0$, the laws of Quantum Mechanics must reduce to those of Classical Mechanics. The correspondence principle, which played such an important role in the establishment of the theory, had as its exact goal the fulfilment of this fundamental requirement.

Classical Mechanics must furnish a good description of phenomena in all those circumstances where the quantum of action may be considered infinitely small. The purpose of the present chapter is to indicate under what circumstances and to what extent this classical approximation is justified.

One of the manifestations of the finite character of $\hbar$ is the existence of a discrete spectrum of eigenvalues for certain observables; the spacing between neighboring eigenvalues is of the order of $\hbar$ [1]). In order that the approximation be justified, it is necessary that this spacing could be considered negligible; that is the case if *large quantum numbers* are involved. This condition, based on the correspondence principle, has already been used in the Old Quantum Theory to calculate the Rydberg constant and to establish the Bohr-Sommerfeld quantization rules (Ch. I, §§ 13 and 15). The condition is certainly not sufficient; thus, some purely quantum-mechanical effects such as the uncertainty relations are not related to the discreteness of certain spectra. More generally, the conditions of validity of the classical approximation are those of *geometrical optics*.

---

[1]) Thus, in the hydrogen atom,

$$E_n = -\tfrac{1}{2}\left(\frac{e^2}{\hbar c}\right)^2 \frac{mc^2}{n^2}, \qquad \Delta E = E_{n+1} - E_n = -\frac{2n+1}{(n+1)^2} E_n.$$

In the limit where $\hbar \to 0$, the distance from a given energy level $E_n = E$ to its nearest neighbor is $\frac{2n+1}{(n+1)^2}E$; in the limiting process, $n^2\hbar^2 =$ constant, hence $n \to \infty$ as $1/\hbar$ and $\Delta E \approx O(1/n) = O(\hbar)$.

The classical approximation can be formulated in two different ways.

The first, which is most readily grasped intuitively, consists in defining the dynamical state of each particle at a given instant by its position and its velocity. If $\hbar$ were zero, this would be rigorously justified, since the components of the position and the momentum of each particle would be represented by commuting observables. In fact, the existence of non-zero commutators

$$[q_i, p_j] = i\hbar\delta_{ij} \tag{VI.1}$$

sets a limit to the precision to which position and momentum can be defined simultaneously, a limit which is given by the uncertainty relations. The dynamical state of the system is rigorously represented by its wave function, and the best one can do is to form the minimum wave packet for which $\Delta q_i \cdot \Delta p_i = \frac{1}{2}\hbar$. The classical picture consists in attributing to each particle a position and a momentum equal to the respective mean values of these quantities in the corresponding quantum state. It systematically ignores any fluctuation about these mean values. In order that this picture be satisfactory, it is necessary that:

(*a*) the mean values follow the classical laws of motion to a good approximation;

(*b*) the dimensions of the wave packet be small with respect to the characteristic dimensions of the problem, and that they remain small in the course of time.

These points will be considered in the two succeeding paragraphs. In particular, we shall see that, except under very special circumstances, any wave packet "spreads" indefinitely in the course of time, and may occupy as large a portion of space as one wishes provided one waits a sufficiently long time. The above classical picture therefore only holds during a finite interval of time.

The second way to formulate the classical approximation consists in likening the system to a statistical mixture of classical systems. More precisely, one defines by means of the wave function a classical statistical mixture whose density at each point of configuration space is equal to the probability density of the presence of the quantum system at that point; one can show that in the limit $\hbar \to 0$, the evolution of this mixture is just the one predicted by Classical

Mechanics. This will be done in § 4. From the mathematical point of view, this formulation [1]) is more satisfactory than the first, because it views the equations of Classical Mechanics as limits of the Schrödinger equation. The conditions of validity of this classical approximation are exactly those of geometrical optics.

This approximation is closely related to an approximate method of solution of the Schrödinger equation, known under the name of WKB method [2]), which is applicable when the Schrödinger equation can be replaced by its classical limit except for limited regions of space surrounding singular points. The WKB method is outlined in the second section of this chapter.

2. Ehrenfest's Theorem

The Ehrenfest Theorem gives the law of motion of the mean values of the coordinates q and the conjugate momenta p of a quantum system. It stipulates that the equations of motion of these mean values are formally identical to the Hamilton equations of Classical Mechanics, except that the quantities which occur on both sides of the classical equations must be replaced by their average values.

This theorem results directly from the application of the general equation (V.72) to the position and momentum coordinates. Let $q_1, \ldots, q_R$ be the (cartesian) coordinates of position, $p_1, \ldots, p_R$ their conjugate momenta, and $H(q_1, \ldots, q_R; p_1, \ldots, p_R)$ the Hamiltonian of the system.

In accordance with eq. (V.72),

$$i\hbar \frac{d}{dt}\langle q_i \rangle = \langle [q_i, H] \rangle$$

$$i\hbar \frac{d}{dt}\langle p_j \rangle = \langle [p_j, H] \rangle.$$

The calculation of the commutators of the right-hand sides was

---

[1]) Actually, there exist several variants, since the definition of the classical statistical mixture from the wave function is not unique (cf. Problem VI.4). The simplest one is given in § 4 below.

[2]) The WKB technique was introduced by Lord Rayleigh (1912) in the solution of problems on wave propagation. Its first application to quantum mechanics is due to H. Jeffreys (1923); it was then developed simultaneously by G. Wentzel, H. A. Kramers, and L. Brillouin (1926).

carried out in Chapter V [eqs. (V.67) and (V.68)]. Whence the announced results [1])

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\langle q_i\rangle &= \left\langle \frac{\partial H}{\partial p_i} \right\rangle, \qquad (i=1, 2, \ldots, R) \\ \frac{\mathrm{d}}{\mathrm{d}t}\langle p_j\rangle &= -\left\langle \frac{\partial H}{\partial q_j} \right\rangle, \qquad (i=1, 2, \ldots, R). \end{aligned} \tag{I}$$

One must clearly understand the connection between the system of equations (I) and the canonical equations of Hamilton. It is generally not correct to state that the mean values $\langle q_i\rangle$ and $\langle p_i\rangle$ follow the laws of Classical Mechanics. The derivatives with respect to time of the *classical* quantities q_i and p_j are well-defined functions $\partial H/\partial p_i$, $-\partial H/\partial q_j$ of these quantities. The values taken by these quantities in the course of time can be exactly derived from their initial values. According to equations (I) on the other hand, the derivatives $\mathrm{d}\langle q_i\rangle/\mathrm{d}t$, $\mathrm{d}\langle p_j\rangle/\mathrm{d}t$ are equal to certain average values whose calculation generally necessitates the knowledge of the wave function $\Psi(t)$. The mean values $\langle q_i\rangle$, $\langle p_j\rangle$ do not follow the classical laws of motion unless one can replace the mean values in (I) of the functions of the right-hand sides by the functions of the mean values, namely

$$\left\langle \frac{\partial}{\partial p_i} H(q_1, \ldots, q_R; p_1, \ldots, p_R) \right\rangle \text{ by } \frac{\partial}{\partial p_i} H(\langle q_1\rangle \ldots \langle q_R\rangle; \langle p_1\rangle \ldots \langle p_R\rangle) \tag{VI.2a}$$

$$\left\langle \frac{\partial}{\partial q_j} H(q_1, \ldots, q_R; p_1, \ldots, p_R) \right\rangle \text{ by } \frac{\partial}{\partial q_j} H(\langle q_1\rangle \ldots \langle q_R\rangle; \langle p_1\rangle \ldots \langle p_R\rangle). \tag{VI.2b}$$

This replacement is rigorously justified only if the Hamiltonian is a polynomial of the second degree with respect to the q's and the p's (free particle, harmonic oscillator, charged particle in a constant electric or magnetic field, cf. Problems VI.1 and VI.2). Apart from these particular cases, this can only happen if the fluctuations of the q's and the p's about their average values are negligible.

By way of an example, let us consider the case of a particle in a potential

$$H = \frac{p^2}{2m} + V(\boldsymbol{r}).$$

---

[1]) Provided one exercises caution in the definition of the operators $\partial H/\partial p_i$, $\partial H/\partial q_i$ (cf. Ch. V, § 17).

Let us introduce the force

$$\boldsymbol{F} = -\text{grad } V(\boldsymbol{r}).$$

The Ehrenfest equations are in that case

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle \boldsymbol{r} \rangle = \frac{\langle \boldsymbol{p} \rangle}{m}, \qquad \frac{\mathrm{d}}{\mathrm{d}t}\langle \boldsymbol{p} \rangle = \langle \boldsymbol{F} \rangle,$$

or else

$$\langle \boldsymbol{F} \rangle = m \frac{\mathrm{d}^2}{\mathrm{d}t^2}\langle \boldsymbol{r} \rangle, \tag{VI.3}$$

the quantum analogue of Newton's law.

In order that the average position

$$\langle \boldsymbol{r} \rangle \equiv \int \Psi^*(\boldsymbol{r}, t)\, \boldsymbol{r} \Psi(\boldsymbol{r}, t)\, \mathrm{d}\boldsymbol{r}$$

actually follow Newton's classical equation, one must be able to replace in eq. (VI.3) the mean value of the force

$$\langle \boldsymbol{F} \rangle = \int \Psi^*(\boldsymbol{r}, t)\, \boldsymbol{F}(\boldsymbol{r})\, \Psi(\boldsymbol{r}, t)\, \mathrm{d}\boldsymbol{r}$$

by its value $\boldsymbol{F}(\langle \boldsymbol{r} \rangle)$ at the point $\boldsymbol{r}$. When the force vanishes (free particle) or depends linearly upon the coordinates of $\boldsymbol{r}$ (harmonic oscillator), one has rigorously $\boldsymbol{F} = \boldsymbol{F}(\langle \boldsymbol{r} \rangle)$. In other cases the substitution is justified only if the wave function remains localized in a sufficiently small region of space so that the force has a practically constant value over that entire region.

3. Motion and Spreading of Wave Packets

In order that the motion of a wave packet may be likened to the motion of a classical particle, it is first of all necessary that its position and momentum follow the laws of classical mechanics; but the dimensions of this packet must also remain sufficiently small at all times. In fact, as Ehrenfest's Theorem suggests, the first requirement is rarely satisfied without the second. We shall especially focus our attention on the latter.

In order to see the essential features, it suffices to study the motion of a one-dimensional wave packet $\psi(q, t)$. Let H be the Hamiltonian:

$$H = \frac{p^2}{2m} + V(q).$$

The quantities whose evolution in time we wish to study are the

mean values $\langle q\rangle$ and $\langle p\rangle$ and the mean-square deviations

$$\chi \equiv (\Delta q)^2 = \langle q^2\rangle - \langle q\rangle^2, \qquad \varpi \equiv (\Delta p)^2 = \langle p^2\rangle - \langle p\rangle^2.$$

In the classical approximation, the packet represents a particle of position and momentum

$$q_{\text{cl.}} = \langle q\rangle, \qquad p_{\text{cl.}} = \langle p\rangle$$

respectively [1]). Note that the energy of this classical particle

$$E_{\text{cl.}} = \frac{\langle p\rangle^2}{2m} + V(\langle q\rangle)$$

is not equal to the average value $\langle H\rangle$. Actually, if the classical approximation is justified, $E_{\text{cl.}}$ is constant in time, as well as the difference

$$\varepsilon = \langle H\rangle - E_{\text{cl.}}. \tag{VI.4}$$

Since the extension $\sqrt{\chi}$ of the wave packet is to remain small, it is natural to replace functions such as $V(q)$, $V'(q)$ by their Taylor expansion about $\langle q\rangle$, namely

$$V(q) = V_{\text{cl.}} + (q - \langle q\rangle)\, V'_{\text{cl.}} + \tfrac{1}{2}(q - \langle q\rangle)^2\, V''_{\text{cl.}} + \dots \tag{VI.5}$$

$$V'(q) = V'_{\text{cl.}} + (q - \langle q\rangle)\, V''_{\text{cl.}} + \tfrac{1}{2}(q - \langle q\rangle)^2\, V'''_{\text{cl.}} + \dots \tag{VI.6}$$

In these expressions, $V_{\text{cl.}}$, $V'_{\text{cl.}}$, ... denote the values assumed by the functions V, V', ... at the point $q = \langle q\rangle$. Taking the average values of these quantities, we obtain, so to speak, their expansions in series of powers of χ, namely

$$\langle V\rangle = V_{\text{cl.}} + \tfrac{1}{2}\chi V''_{\text{cl.}} + \dots \tag{VI.7}$$

$$\langle V'\rangle = V'_{\text{cl.}} + \tfrac{1}{2}\chi V'''_{\text{cl.}} + \dots \tag{VI.8}$$

With the help of these expansions, we are in a position to obtain entirely general results, valid for any $V(q)$.

---

[1]) In the study of the one-dimensional wave packet of Chapter II, the center of the wave packet was defined by a condition of stationary phase, while the definition adopted here is the mean value $\langle q\rangle$. In the classical approximation, the difference between these two definitions is negligible.

The quantities $\langle q \rangle$ and $\langle p \rangle$ follow Ehrenfest's equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle q \rangle = \frac{\langle p \rangle}{m} \tag{VI.9 a}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle p \rangle = -\langle V' \rangle. \tag{VI.9 b}$$

They reduce to the classical equations if the right-hand side of eq. (VI.9*b*) is replaced by the first term of its expansion (VI.8). This replacement is justified if the term $\frac{1}{2}\chi V'''_{\mathrm{cl.}}$ and terms of higher order are negligible. This holds rigorously true when $V'''(q)$ is zero everywhere, i.e. when $V(q)$ is a polynomial of at most the second degree in q, and especially when $V(q) = cq^2$ (harmonic oscillator) and when $V(q) = 0$ (free particle). Otherwise, $V(q)$ must vary sufficiently slowly over a distance of the order of the extension $\chi^{\frac{1}{2}}$ of the wave packet, so that the effect of V''' and higher order derivatives in the expansion (VI.8) is sufficiently small.

Assuming these conditions to be fulfilled [which amounts to assuming the expansions (VI.7) and (VI.8) to be rapidly converging] we obtain for the constant ε [cf. eq. (VI.4)] the expression

$$\varepsilon \equiv \frac{\varpi}{2m} + \langle V \rangle - V_{\mathrm{cl.}} \simeq \frac{1}{2m}(\varpi + m V''_{\mathrm{cl.}}\,\chi) = \text{constant} \tag{VI.10}$$

connecting the mean-square deviations ϖ and χ.

We shall restrict ourselves therefore to a study of χ as a function of time. χ is the mean value of the operator $q^2 - \langle q \rangle^2$ (depending explicitly upon the time since $\langle q \rangle$ is a function of time). Applying relation (V.72) to this operator, one obtains after some calculation

$$\frac{\mathrm{d}}{\mathrm{d}t}\chi = \frac{1}{m}(\langle pq + qp \rangle - 2\langle p \rangle \langle q \rangle).$$

By an analogous procedure for $\mathrm{d}\chi/\mathrm{d}t$, one has

$$\frac{\mathrm{d}^2\chi}{\mathrm{d}t^2} = \frac{2\varpi}{m^2} - \frac{1}{m}(\langle V'q + qV' \rangle - 2\langle q \rangle \langle V' \rangle).$$

By replacing the operator V' in the bracket of the right-hand side by the first two terms of expansion (VI.6), we obtain the approximate equation

$$\frac{\mathrm{d}^2\chi}{\mathrm{d}t^2} \simeq \frac{2}{m^2}(\varpi - m V''_{\mathrm{cl.}}\,\chi) \tag{VI.11}$$

which we can write, taking into account equation (VI.10)

$$\frac{d^2\chi}{dt^2} \approx \frac{4}{m}(\varepsilon - V''_{cl.}\chi). \tag{VI.12}$$

Knowing the deviations χ_0, ϖ_0, and $\dot{\chi}_0 \equiv d\chi_0/dt$ at time t_0, one obtains χ at any later time by solving the equation (VI.12) – noting carefully that $V''_{cl.}$ may be time-dependent – and ϖ by means of equation (VI.10). In view of this fact, the error introduced by replacing $\langle V' \rangle$ by $V'_{cl.}$ in equation (VI.9*b*) can be evaluated. All the elements are therefore at hand for deciding if the assimilation of the wave packet to a classical particle is justified.

The two most interesting cases are those of the harmonic oscillator and the free particle, cases for which the motion of the center of the packet is rigorously identical to that of a classical particle. In the case of the harmonic oscillator ($V = \frac{1}{2}m\omega^2 q^2$), $\langle q \rangle$ oscillates about zero with the frequency $\omega/2\pi$ and χ oscillates about $\varepsilon/m\omega^2$ with twice that frequency. (Cf. Problem VI.1.)

If one is dealing with a free particle ($V = 0$), $\langle q \rangle$ undergoes uniform rectilinear motion of velocity $\langle p \rangle/m$, the deviation ϖ remains strictly constant ($\varpi = \varpi_0 = 2m\varepsilon$) and since one has rigorously $d^2\chi/dt^2 = 2\varpi_0/m^2$ [the equations (VI.11) and (VI.12) are exact in that case]

$$\chi = \chi_0 + \dot{\chi}_0 t + \frac{\varpi_0}{m^2} t^2. \tag{VI.13}$$

Consequently, after a sufficiently long time, χ becomes as large as desired: the free wave packet "spreads" indefinitely.

The appearance of an indefinite spreading is an important fact because it sets a limit for the time interval during which the wave packet may be likened to a classical particle. Indeed, apart from very special cases such as that of harmonic oscillation, the wave packet always spreads, especially in collision problems where the laws of motion for the packet approach those of a free wave packet as soon as the colliding particle is sufficiently far from the scattering center.

The spreading law for a free wave packet turns out to be quite simple if the wave packet is taken to be minimum at time t_0 (cf. Problem IV.4), i.e. $\varpi_0\chi_0 = \frac{1}{2}\hbar^2$ (in which case $\dot{\chi}_0 = 0$); then one has

$$\chi = \chi_0 + \frac{\varpi_0}{m^2} t^2$$

or else

$$\Delta q = \left[(\Delta q_0)^2 + \left(\frac{\Delta p_0}{m} t \right)^2 \right]^{\frac{1}{2}} . \tag{VI.14}$$

The very form of the "spreading term" $\Delta p_0 t/m$ suggests a simple classical picture for this free wave packet, that of a swarm of projectiles initially bunched within an interval Δq_0 about the average value $\langle q_0 \rangle$, the velocities of these projectiles being distributed over an interval

$$\Delta v = \Delta p_0/m$$

about the group velocity of the packet $v = \langle p_0 \rangle/m$. Owing to their dispersion in velocity, projectiles initially located at the same point, come to be uniformly distributed over a band $\Delta v t$ at time t; consequently, the projectiles do not retain their original clustering, and the width of the swarm increases appreciably according to the law (VI.14).

This same law can be put into other forms. In particular, it may be written

$$\Delta q = \Delta q_0 \left[1 + \frac{1}{4} \left(\frac{D \lambda\!\!\!^{-}}{(\Delta q_0)^2} \right)^2 \right]^{\frac{1}{2}}$$

where $D \equiv vt$ is the distance covered by the wave packet during the time t and $\lambda\!\!\!^{-} \equiv \hbar/mv$ is the average wave length. The spread of the free wave packet is thus negligible provided

$$\sqrt{D \lambda\!\!\!^{-}} \ll \Delta q_0. \tag{VI.15}$$

In fact, it is easy to show that the extension Δq always exceeds the length $\sqrt{D \lambda\!\!\!^{-}}$. (Note that: $D \lambda\!\!\!^{-} = \hbar t/m$.)

4. Classical Limit of the Schrödinger Equation

We shall now study the second formulation of the classical approximation mentioned in the introduction of this chapter.

To be definite, we treat the case of a particle in a potential $V(\mathbf{r})$. Let us separate modulus and phase in its wave function:

$$\Psi(\mathbf{r}) = A(\mathbf{r}) \exp \left(\frac{\mathrm{i}}{\hbar} S(\mathbf{r}) \right). \tag{VI.16}$$

Substituting expression (VI.16) in the Schrödinger equation and

separating real and imaginary parts, we obtain the two equations

$$\frac{\partial S}{\partial t} + \frac{(\nabla S)^2}{2m} + V = \frac{\hbar^2}{2m}\frac{\triangle A}{A} \tag{VI.17}$$

$$m\frac{\partial A}{\partial t} + (\nabla A \cdot \nabla S) + \frac{A}{2}\triangle S = 0. \tag{VI.18}$$

These two equations are strictly equivalent to the Schrödinger equation. In fact, equation (VI.18) is just the continuity equation (IV.11). Indeed, the probability density of presence $P(\mathbf{r})$ and the current density $\mathbf{J}(\mathbf{r})$ (cf. Ch. IV, §§ 2 and 4) are respectively given by

$$P = A^2, \qquad \mathbf{J} = A^2\frac{\nabla S}{m}.$$

After multiplying term by term by $2A$, eq. (VI.18) can be put in the form

$$m\frac{\partial}{\partial t}A^2 + \operatorname{div}(A^2\nabla S) = 0. \tag{VI.19}$$

The classical approximation consists in setting $\hbar$ equal to zero in eq. (VI.17), that is

$$\frac{\partial S}{\partial t} + \frac{(\nabla S)^2}{2m} + V = 0. \tag{VI.20}$$

We deduce the following result:

In the classical approximation, Ψ describes a fluid of non-interacting classical particles of mass m, (statistical mixture), *and subject to the potential $V(\mathbf{r})$: the density and current density of this fluid at each point of space are at all times respectively equal to the probability density P and the probability current density $\mathbf{J}$ of the quantum particle at that point* [1]).

Indeed, since the continuity equation of this fluid is satisfied [eq. (VI.19)], it suffices to show that the velocity field

$$\mathbf{v} = \frac{\mathbf{J}}{P} = \frac{\nabla S}{m} \tag{VI.21}$$

---

[1]) The density in phase space of this classical statistical mixture is somewhat peculiar since to each point of configuration space corresponds a well-defined momentum ΔS. The solution S of equation (VI.20) is the "principal function" of Hamilton used in the Hamilton-Jacobi formulation of Classical Mechanics.

of this fluid actually follows the law of motion of the classical fluid in question. Now, with the definition (VI.21), eq. (VI.20) is written

$$\frac{\partial S}{\partial t} + \frac{mv^2}{2} + V = 0.$$

Writing down that the gradient of the left-hand side is zero, one has, taking into account eq. (VI.21),

$$\left(\frac{\partial}{\partial t} + (\mathbf{v} \cdot \text{grad})\right) m\mathbf{v} + \text{grad } V = 0,$$

from which one concludes that the particles of the fluid obey the equation of motion

$$m \frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = -\text{grad } V. \qquad \text{Q.E.D.}$$

It is important to stress the very great generality of this result which remains valid for systems with any number of dimensions. The density $P = |\Psi|^2$ is a well-defined function of configuration space; similarly, the current $\boldsymbol{J}$ is a well-defined vector field of that space. The proof for the general case will be left to the reader.

When Ψ represents a stationary state of energy E,

$$\frac{\partial A}{\partial t} = 0, \qquad \frac{\partial S}{\partial t} = -E.$$

Equations (VI.17) and (VI.18) reduce to

$$(\nabla S)^2 - 2m(E - V) = \hbar^2 \frac{\Delta A}{A} \qquad \text{(VI.22)}$$

$$\text{div } (A^2 \nabla S) = 0. \qquad \text{(VI.23)}$$

In the classical approximation, the right-hand side of eq. (VI.22) is neglected and the result proved above still holds. Ψ then represents a stationary flow of the fluid of classical particles.

Optical analogy is even more suggestive than this hydrodynamical analogy, especially for stationary solutions. Since the velocities of the particles are proportional to the gradient of S, the trajectories of these particles are orthogonal to the surfaces of equal phase $S = const.$ In the language of optics, the latter are the wave fronts and the

trajectories of the particles are the rays [1]). Hence, the classical approximation is equivalent to the geometrical optics approximation: we find once again, as a consequence of the Schrödinger equation, the basic postulate of the theory of matter waves.

The optical analogy is extremely useful whenever one wishes to determine if the conditions of validity of the classical approximation are actually fulfilled. Let $\lambda\!\!\!^{-}$ be the reduced wavelength:

$$\lambda\!\!\!^{-} = \frac{\hbar}{\sqrt{2m\,[E - V(\boldsymbol{r})]}}.$$

It is a well-defined function of $\boldsymbol{r}$. Eq. (VI.22) can be written in the form

$$(\nabla S)^2 = \frac{\hbar^2}{\lambda\!\!\!^{-2}}\left(1 + \lambda\!\!\!^{-2}\frac{\Delta A}{A}\right). \qquad \text{(VI.22')}$$

In the classical approximation

$$\lambda\!\!\!^{-2}\frac{\Delta A}{A} \ll 1 \qquad \text{(VI.24)}$$

over all space; more precisely, the regions in which this condition is not fulfilled must be sufficiently small so that eq. (VI.22') can be replaced by the approximate equation

$$(\nabla S)^2 = \frac{\hbar^2}{\lambda\!\!\!^{-2}}. \qquad \text{(VI.25)}$$

This is the equation of the wave fronts of geometrical optics.

When $V = 0$ over all space, $\lambda\!\!\!^{-}$ is constant and the function $S = \boldsymbol{p}\cdot\boldsymbol{r} + \textit{constant}$, where $\boldsymbol{p}$ is a given vector of length $\hbar/\lambda\!\!\!^{-}$, is a particular solution of eq. (VI.25) [the ensemble of functions of this type forms a complete integral of this first-order partial differential equation]; the wave fronts corresponding to this solution are planes perpen-

---

[1]) In the presence of a vector potential A, the definition (VI.21) of the classical velocity must be replaced by

$$\boldsymbol{v} = \frac{1}{m}\left(\nabla S - \frac{e}{c}A\right).$$

With this new definition, the above theorem remains true. The notions of wave front and ray remain valid, but the rays are no longer orthogonal to the wave fronts; this situation is analogous to that of anisotropic media in geometrical optics.

dicular to $\boldsymbol{p}$, and the rays are straight lines parallel to $\boldsymbol{p}$. In the general case, the wave fronts and the rays are curved.

Once the energy E is fixed, the specification of a wave front $S(x, y, z) = S_0$ determines one and only one solution of eq. (VI.25). This solution can be obtained in the following manner. With the surface (S_0) is associated a (two parameter) family of trajectories of the classical particle; these trajectories correspond to the energy E and are orthogonal to (S_0); the other wave fronts are surfaces orthogonal to these trajectories (Fig. VI.1). To find the value taken

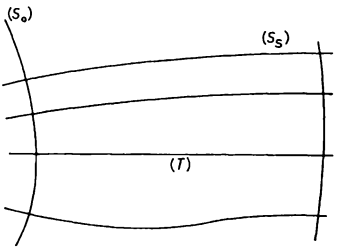


Fig. VI.1. Family of trajectories (of given energy E) associated with a wave front (S_0) in the approximation of geometrical optics: $(\nabla S)^2 = \hbar^2/\lambda\!\!\!^{-2}$.

by S on each of these wave fronts, one considers one of the trajectories (T). Each point of (T) is labelled by its curvilinear coordinate s; we locate the origin $s = 0$ at the intersection of (T) with (S_0). One has, according to eq. (VI.25),

$$S(s) = S_0 + \int_0^s \frac{\hbar}{\lambda\!\!\!^{-}}\, \mathrm{d}s\,;$$

S being thus determined over all space, specifying A on the surface (S_0) uniquely determines the function A over all space. Actually eq. (VI.23) relates only the values $A(s)$ taken by A along the same trajectory; indeed it may be written

$$\frac{\hbar}{\lambda\!\!\!^{-}} \frac{\mathrm{d}}{\mathrm{d}s} (A^2) + A^2\, \Delta S = 0.$$

Since the functions $\lambda\!\!\!^{-}$ and ΔS are well-defined functions of s on that trajectory, this equation defines $A(s)$ in a unique manner when its value $A(0)$ at the intersection of (T) and (S_0) is known.

Knowing the solutions A and S in the classical approximation, it is possible to evaluate the perturbation due to the term $\lambda\!\!\!^{-2} \Delta A/A$ in the rigorous equation (VI.22′). The effect depends not only upon the "optical properties" of the medium through which the wave

propagates, but also upon the particular solution of the wave equation As a general rule, the classical approximation is justified if the "transverse" dimensions of the wave are everywhere large compared to $\lambda\!\!\!^{-}$ and if one has

$$|\text{grad } \lambda\!\!\!^{-}| \ll 1 \tag{VI.26}$$

if not over all space, then at least in that portion of space where the density A^2 takes on non-negligible values.

These conditions of validity are deduced from the following semi-quantitative arguments. As suggested by the optical analogy, the curvature of the light rays, i.e. of the particle trajectories, must be small compared to the wavelength. Now, the radius of curvature R is related to the velocity v of the particles and to the transverse component of the force $-(\text{grad } V)_\perp$ by the relation

$$\frac{mv^2}{R} = |(\text{grad } V)_\perp|.$$

It is therefore necessary that

$$\frac{\lambda\!\!\!^{-}}{R} = \frac{\lambda\!\!\!^{-}|(\text{grad } V)_\perp|}{mv^2} = \frac{m\lambda\!\!\!^{-3}}{\hbar^2}|(\text{grad } V)_\perp| \ll 1.$$

Taking into account the expression for $\lambda\!\!\!^{-}$ as a function of the potential V, this condition is written

$$|(\text{grad } \lambda\!\!\!^{-})_\perp| \ll 1.$$

Now, the curvature of the wave fronts must be small with respect to $1/\lambda\!\!\!^{-}$ (except possibly in certain restricted regions of space, in particular near the focal surfaces) and this may in general be realized if the trajectories also have small curvatures, provided we make an appropriate choice of the surface (S_0). Similarly the relative variation of A on each wave front must be negligible in a region of order $\lambda\!\!\!^{-}$; in other words, the "transverse" dimensions of the wave must be large compared to $\lambda\!\!\!^{-}$. If we assume that these conditions on the solution are fulfilled, the function $A(s)$ along a trajectory (T) is approximately given by

$$A(s) \approx A(0)\sqrt{\frac{\lambda\!\!\!^{-}(s)}{\lambda\!\!\!^{-}(0)}},$$

which yields, after a short calculation,

$$\lambda\!\!\!^{-2}\frac{\Delta A}{A} \approx \frac{\lambda\!\!\!^{-2}}{A}\frac{d^2A}{ds^2} \approx \frac{1}{4}\left[2\lambda\!\!\!^{-}\frac{d^2\lambda\!\!\!^{-}}{ds^2} - \left(\frac{d\lambda\!\!\!^{-}}{ds}\right)^2\right].$$

In order that this term be small it is necessary that $\mathrm{d}\lambda/\mathrm{d}s$ and $\lambda \mathrm{d}^2\lambda/\mathrm{d}s^2$ be much less than 1. In practice, the second condition is always fulfilled if the first one holds; that is to say if the component of grad λ along the trajectory is very much less than 1, that is

$$|(\mathrm{grad}\,\lambda)_{\parallel}| \ll 1.$$

5. Application to Coulomb Scattering. The Rutherford Formula

By way of an application, we shall present a brief outline of the classical theory of Coulomb scattering and its conditions of validity.

Consider the scattering of a particle of mass m by a Coulomb potential

$$V = \frac{Ze^2}{r},$$

where r is the distance of the particle from the center of force C. This particle is, for instance, a proton of charge e in the Coulomb field of an atomic nucleus of charge Ze (repulsive potential). The theory applies equally well if the charges are of opposite sign (attractive potential). We shall thus treat the constant Ze^2 as an algebraic quantity capable of taking on either sign.

Since the energy E of the particle is fixed once and for all by the relation

$$E = \frac{{p_0}^2}{2m} = \tfrac{1}{2}mv_0^2,$$

and the direction of propagation of the incident particle is also given, specification of the impact parameter b completes the determination of the initial conditions of the motion. Let

$$a = \frac{1}{2}\frac{Ze^2}{E}. \tag{VI.27}$$

As is well known, the trajectory is a branch of an hyperbola (Fig. VI.2) with focus C, semi-axis $OA = |a|$, and focal distance $OC = \sqrt{a^2+b^2}$. The deflection angle is given by the relation

$$b = |a| \cot\frac{\theta}{2} \tag{VI.28}$$

(the direction of the deflection depends on the sign of the potential,

but the absolute value of the angle is the same for two potentials equal in absolute value).

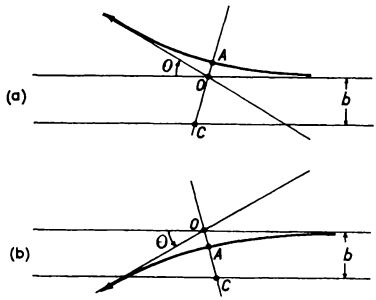


Fig. VI.2. Trajectory followed (heavy line) by a particle in a (*a*) repulsive; (*b*) attractive Coulomb field.

The important quantity one must know in practice is the differential cross section $d\sigma/d\Omega$. Consider a parallel beam of incident monoërgic particles directed toward the center of force; one observes the number of particles scattered into a given solid angle $(\Omega, \Omega + d\Omega)$. $d\sigma/d\Omega$ is by definition the number of particles scattered in that direction, per unit solid angle and per unit time, when the particles are uniformly distributed over the incident beam with a flux that is constant in time and equal to unity; that is to say, when any surface located very far from C in the incident direction of propagation, and perpendicular to that direction, is crossed by the particles at the rate of one particle per unit time and per unit area.

Take the incident direction as polar axis, and call θ and φ the polar angles of the direction of propagation of the scattered particle. θ is the deflection angle introduced above; it is connected with the impact parameter b by relation (VI.28). The solid angle $d\Omega$ in the direction (θ, φ) is equal to

$$d\Omega = \sin\theta \, d\theta \, d\varphi.$$

The number $d\sigma$ of particles scattered per unit time into this solid angle is equal to the number of incident particles crossing the surface $b \, db \, d\varphi$ per unit time; that is, since the incident flux is equal to 1,

$$d\sigma = b \, db \, d\varphi = b \frac{db}{d\theta} d\theta \, d\varphi = \frac{b}{\sin\theta} \frac{db}{d\theta} d\Omega.$$

Replacing in this equation the quantities b and $db/d\theta$ by their

expressions in terms of θ calculated from relation (VI.28), one obtains the *Rutherford formula* [1])

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{a^2}{4\sin^4\frac{\theta}{2}} = \frac{(Ze^2)^2}{16E^2\sin^4\frac{\theta}{2}}. \tag{VI.29}$$

We shall now discuss the validity of the classical approximation. Note first that $|a|$ is a characteristic length of this collision problem. Moreover, the wavelength of the particle

$$\lambda\!\!\!^{-} = \frac{\hbar}{p} = \frac{\hbar}{p_0}\left(1 - \frac{2a}{r}\right)^{-\frac{1}{2}}$$

is of the order of magnitude of its initial value $\lambda\!\!\!^{-}_0 = \hbar/mv_0$. The ratio of the two quantities $|a|$ and $\lambda\!\!\!^{-}$ is therefore of order

$$\gamma = \frac{|a|}{\lambda\!\!\!^{-}_0} = \frac{|Ze^2|}{\hbar v_0}. \tag{VI.30}$$

One expects the classical approximation to be justified in the limit where

$$\gamma \gg 1. \tag{VI.31}$$

Let us examine to what extent condition (VI.26) is fulfilled. One has

$$|\operatorname{grad}\lambda\!\!\!^{-}| = \left|\frac{\mathrm{d}\lambda\!\!\!^{-}}{\mathrm{d}r}\right| = \frac{1}{\gamma}\frac{a^2}{\sqrt{r(r-2a)^3}} \ll 1;$$

$|\operatorname{grad}\lambda\!\!\!^{-}|$ is larger the smaller r. Since the classical trajectory approaches the scattering center the closer the larger the angle of deflection, we conclude that the classical approximation holds for small deflection angles (that is to say for large impact parameters) and ceases to be valid for large angles. In order to reach a more quantitative conclusion, consider the expression for the maximum of $|\operatorname{grad}\lambda\!\!\!^{-}|$ (i.e. its value at the vertex of the hyperbolic trajectory) as a function of θ; after some calculation

$$|\operatorname{grad}\lambda\!\!\!^{-}|_{\max.} = \frac{F(\theta)}{\gamma}, \qquad F(\theta) = \tan\tfrac{1}{2}\theta\,\frac{\sin\frac{1}{2}\theta}{1-(\operatorname{sgn}a)\sin\frac{1}{2}\theta}.$$

[1]) This formula, due to Rutherford, is of great historical importance; upon it rested the entire interpretation of Rutherford's famous experiments on α-particle scattering.

[By definition, sgn $a \equiv |a|/a$.] When θ increases from 0 to π, $F(\theta)$ increases from 0 to $+\infty$. Let θ_c be an angle such that

$$F(\theta_c)=\gamma.$$

The classical approximation is justified as long as $\theta<\theta_c$ and ceases to be valid when $\theta>\theta_c$. Note that θ_c lies the closer to π the larger γ, in agreement with the rough predictions made earlier.

II. THE WKB METHOD [1]

6. Principle of the Method

The method consists in introducing an expansion in powers of $\hbar$ and in neglecting terms of higher order than $\hbar^2$. One thus replaces (at least in some regions of space) the Schrödinger equation by its classical limit. However, the method has a wider range of applicability than the classical approximation proper because this procedure can be carried out even in regions of space where the classical interpretation is meaningless (regions $E<V$ which are inaccessible to classical particles). In order to include these regions in our treatment, we must slightly modify the definitions of A and S of § 4 and write

$$\Psi(\boldsymbol{r}) = \exp\left(\frac{\mathrm{i}}{\hbar} W(\boldsymbol{r})\right) \tag{VI.32}$$

$$W(\boldsymbol{r}) = S(\boldsymbol{r}) + \frac{\hbar}{\mathrm{i}} T(\boldsymbol{r}) \tag{VI.33}$$

$$A(\boldsymbol{r}) = \exp\left[T(\boldsymbol{r}).\right] \tag{VI.34}$$

Requiring S and T to be even functions of $\hbar$ defines A and S in unique fashion. With these new definitions, equations (VI.17) and (VI.18) as well as (VI.22) and (VI.23) remain valid although A and S are no longer necessarily real. The WKB approximation consists in expanding $W(\boldsymbol{r})$ in powers of $\hbar$ and in neglecting in the Schrödinger equation terms of order equal to or greater than $\hbar^2$.

---

[1] Cf. § 1, footnote 2, p. 216. For more details, see R. E. Langer, Phys. Rev. 51 (1937) 669; W. H. Furry, Phys. Rev. 71 (1947) 360. See also P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (New York, McGraw-Hill, 1953) pp. 1092 ff.

7. One-Dimensional WKB Solutions

The WKB method can be easily applied only to *one-dimensional problems*. We shall therefore restrict ourselves to treating one-dimensional problems and we shall look for the stationary solutions of the time-independent Schrödinger equation [eqs. (VI.22) and (VI.23)]. The method developed here may eventually serve to solve the three-dimensional Schrödinger equation, since the latter can be reduced in many practical cases to the solution of a one-dimensional equation by separating angular and radial variables (cf. Ch. IX).

Let $y(x)$ be the wave function satisfying the Schrödinger equation

$$y'' + \frac{2m}{\hbar^2}[E - V(x)]y = 0.$$

Setting

$$y = e^{iw/\hbar}, \qquad w = S + \frac{\hbar}{i}\ln A$$

(S and $\ln A$ are even functions of $\hbar$) one obtains the system of equivalent equations

$$S'^2 - 2m(E - V) = \hbar^2 \frac{A''}{A} \tag{VI.35}$$

$$2A'S' + AS'' = 0. \tag{VI.36}$$

The equation of continuity (VI.36) can be integrated to yield

$$A = \text{const.}\,(S')^{-\frac{1}{2}}. \tag{VI.37}$$

Substituting this expression for A into eq. (VI.35), one obtains the equation

$$S'^2 = 2m(E - V) + \hbar^2\left[\frac{3}{4}\left(\frac{S''}{S'}\right)^2 - \frac{1}{2}\frac{S'''}{S'}\right]. \tag{VI.38}$$

This differential equation of the third order in S is *rigorously* equivalent to the initial Schrödinger equation.

In the WKB *approximation* one expands S in a power series in $\hbar^2$:

$$S = S_0 + \hbar^2 S_1 + \ldots, \tag{VI.39}$$

then one substitutes this expansion into eq. (VI.38) and keeps only zero-order terms:

$$S'^2 \approx S_0'^2 = 2m[E - V(x)]. \tag{VI.40}$$

This approximate equation is easily integrated.

One must distinguish two cases according to whether $E \gtrless V(x)$.

First Case: $E > V(x)$.

We define the wavelength

$$\lambda\!\!\!^{-}(x) = \frac{\hbar}{\sqrt{2m[E - V(x)]}}. \tag{VI.41}$$

Equation (VI.40) is satisfied if $S' \approx \pm\hbar/\lambda\!\!\!^{-}$. The WKB solution is a linear combination of oscillating functions

$$y(x) = \alpha\sqrt{\lambda\!\!\!^{-}}\cos\left(\int^{x}\frac{dx}{\lambda\!\!\!^{-}} + \varphi\right) \tag{VI.42}$$

(α and φ are arbitrary constants).

Second Case: $E < V(x)$ (forbidden region for classical particles).

$$l(x) = \frac{\hbar}{\sqrt{2m[V(x) - E]}}. \tag{VI.43}$$

Eq. (VI.40) is satisfied if $S' \approx \pm i\hbar/l$. The WKB solution is a linear combination of real exponentials

$$y(x) = \sqrt{l}\left[\gamma\exp\left(+\int^{x}\frac{dx}{l}\right) + \delta\exp\left(-\int^{x}\frac{dx}{l}\right)\right]. \tag{VI.44}$$

8. Conditions for the Validity of the WKB Approximation

The theory of the WKB approximation is rather involved. Let us merely indicate here without proof that the expansion (VI.39) in powers of $\hbar^2$ does not in general converge, but that it is an asymptotic expansion which, if broken off after a finite number of terms, gives S to a good approximation if $\hbar$ is sufficiently small.

In order to find a criterion for the validity of the WKB approximation, one can calculate the second term $\hbar^2 S_1$ of the expansion (VI.39). The correction of order $\hbar^2$ consists in multiplying the WKB solution by the factor $\exp(i\hbar S_1)$. The effect is negligible if $\hbar S_1 \ll 1$.

By substituting expansion (VI.39) into equation (VI.38) and equating the coefficients of $\hbar^2$ on the two sides, one obtains the differential equation for S_1

$$2S_0' S_1' = \frac{[(S_0')^{-\frac{1}{2}}]''}{(S_0')^{-\frac{1}{2}}} \equiv \frac{3}{4}\left(\frac{S_0''}{S_0'}\right)^2 - \frac{1}{2}\frac{S_0'''}{S_0'}.$$

When $E > V$, $S_0' = \pm\hbar/\lambda\!\!\!^{-}$. After some calculation this yields

$$\hbar S_1' = \pm \tfrac{1}{2}\sqrt{\lambda\!\!\!^{-}}(\sqrt{\lambda\!\!\!^{-}})'' = \pm\left(\tfrac{1}{4}\lambda\!\!\!^{-}'' - \frac{1}{8}\frac{\lambda\!\!\!^{-}'^2}{\lambda\!\!\!^{-}}\right),$$

from which one obtains

$$\hbar S_1 = \pm\left(\tfrac{1}{4}\lambda\!\!\!^{-}' - \frac{1}{8}\int^x \frac{\lambda\!\!\!^{-}'^2}{\lambda\!\!\!^{-}}\,\mathrm{d}x\right). \tag{VI.45}$$

When $E > V$, one obtains an identical expression with $l(x)$ instead of $\lambda\!\!\!^{-}(x)$. The condition $\hbar S_1 \ll 1$ is fulfilled if

$$\begin{aligned} \lambda\!\!\!^{-}'(x) &\ll 1 \quad \text{when} \quad E > V(x), \\ l'(x) &\ll 1 \quad \text{when} \quad E < V(x). \end{aligned} \tag{VI.46}$$

One can compare the criterion (VI.46) with condition (VI.26) for the validity of the classical approximation in general.

This criterion may just as well be expressed by the following inequality which involves the potential $V(x)$ and its first derivative

$$\frac{|m\hbar V'|}{|2m(E-V)|^{3/2}} \ll 1. \tag{VI.47}$$

9. Turning Points and Connection Formulae

In most applications of the WKB approximation condition (VI.47) is fulfilled everywhere except in the vicinity of the points for which

$$E = V(x).$$

These are the *turning points* of the classical motion, points where the velocity of the particle vanishes and changes sign.

From the mathematical point of view, the WKB approximation consists in replacing the Schrödinger equation

$$y'' + \frac{y}{\lambda\!\!\!^{-2}} = 0$$

by the equation

$$y'' + \left(\frac{1}{\lambda\!\!\!^{-2}} - \frac{(\sqrt{\lambda\!\!\!^{-}})''}{\sqrt{\lambda\!\!\!^{-}}}\right) y = 0 \tag{VI.48}$$

[both in the region $E > V$ as well as in the region $E < V$, where $\lambda\!\!\!^{-} = il$]. In fact one can easily verify that expressions (VI.42) and (VI.44)

are actually the general solutions of equation (VI.48). This equation has a singular point – a singularity in $(x-a)^{-2}$ – at any point a where the wavelength becomes infinite, i.e. at each of the turning points. In the vicinity of these points, the replacement of the Schrödinger equation by eq. (VI.48) is certainly not justified. To obtain the complete solution, one must therefore solve the Schrödinger equation in a region of suitable extension surrounding the turning point, and then smoothly join this solution with the solutions (VI.42) or (VI.44) which are the wave functions in neighboring regions where the WKB approximation is valid.

In practice, it is of little importance to know the particular form of the solution in the region of the turning point as long as one knows how to join the WKB solutions together on either side of that point. This joining problem is a difficult mathematical problem, a thorough and lucid account of which may be found in Langer's article (see footnote p. 231). The solution proposed by Langer consists in replacing the Schrödinger equation, not by equation (VI.48), but by another equation which is regular at the turning point and which asymptotically approaches (VI.48) on either side. The general solution of this equation asymptotically approaches one or the other of solutions (VI.42) and (VI.44) on either side of this point. We shall merely give the *connection formulae* between the exponential WKB solution and the oscillatory WKB solution on either side of a turning point.

Suppose, to be definite, that $E \gtrless V$ according to whether $x \gtrless a$ (barrier to the left). The general solution is a linear combination of two solutions y_1 and y_2 whose asymptotic forms are

$$x \ll a \qquad\qquad x \gg a$$

$$y_1 \sim \sqrt{l}\exp\left(+\int_x^a \frac{\mathrm{d}x}{l}\right) \qquad y_1 \sim -\lambda^{\frac{1}{2}}\sin\left(\int_a^x \frac{\mathrm{d}x}{\lambda} - \frac{\pi}{4}\right) \qquad \text{(VI.49)}$$

$$y_2 \sim \frac{\sqrt{l}}{2}\exp\left(-\int_x^a \frac{\mathrm{d}x}{l}\right) \qquad y_2 \sim \lambda^{\frac{1}{2}}\cos\left(\int_a^x \frac{\mathrm{d}x}{\lambda} - \frac{\pi}{4}\right). \qquad \text{(VI.50)}$$

Let us define the "number of wavelengths" contained in a given interval (x_1, x_2) by the integral $(1/2\pi)\int_{x_1}^{x_2}(\mathrm{d}x/\lambda)$ or $(1/2\pi)\int_{x_1}^{x_2}(\mathrm{d}x/l)$ according to whether we are to the right or to the left of the turning

point. The conditions for the validity of these connection formulae are as follows:

(*i*) At the turning point the kinetic energy $E-V$ tends to zero as $(x-a)$ and remains to a good approximation proportional to $(x-a)$ in a region extending over at least one, but preferably several "wavelengths" on either side.

(*ii*) Each of these turning regions joins on, on either side of its turning point, to an asymptotic region extending over several "wavelengths", in which the WKB approximation proper is justified.

Care must be taken in using formulae (VI.49) and (VI.50). The difficulty arises from the fact that the solution $Ay_1 + By_2$ has, except in the very special case where $A=0$, the same asymptotic form as Ay_1 in the region $x \ll a$; the exponentially increasing term Ay_1 always predominates over the exponentially decreasing term By_2, no matter how small A compared to B, provided A does not vanish. Consequently, knowledge of the asymptotic form is not sufficient to specify the solution unless this form is of the exponentially decreasing type (type y_2); conversely, if the coefficients A and B are only known approximately and if $|A| \ll |B|$, any determination, even though approximate, of the asymptotic form is impossible.

Now suppose that we know the WKB solution in this asymptotic region ($x \ll a$) and that we ask on to what particular oscillatory WKB solution it joins. The question cannot be answered unless this solution is of the exponentially decreasing type $\frac{1}{2}Bl^{\frac{1}{2}} \exp\left[-\int_x^a (\mathrm{d}x/l)\right]$; in that case, the solution is certainly of the form By_2 in the region of the turning point and its behavior in the region $x \gg a$ is given by formula (VI.50). The result may be written

$$\tfrac{1}{2}\sqrt{l}\,\exp\left(-\int_x^a \frac{\mathrm{d}x}{l}\right) \rightarrow \sqrt{\bar{\lambda}}\,\cos\left(\int_a^x \frac{\mathrm{d}x}{\bar{\lambda}} - \frac{\pi}{4}\right), \qquad \text{(VI.51)}$$

the arrow indicating the direction in which the connection is made.

Suppose, on the other hand, that we know the WKB solution in the oscillatory region ($x \gg a$). This solution is of the form (VI.42), namely

$$C\sqrt{\bar{\lambda}}\,\cos\left(\int_a^x \frac{\mathrm{d}x}{\bar{\lambda}} - \frac{\pi}{4} + \varphi\right)$$

(C and φ are complex constants). According to formulae (VI.49) and (VI.50), this is the asymptotic form of the solution defined by

$$A \approx C \sin \varphi \qquad B \approx C \cos \varphi.$$

We must insist here that the constants A and B can be calculated only approximately from the asymptotic form. Because of this fact, if $|\tan \varphi| \ll 1$, any determination of the asymptotic form of this solution in the region $x \ll a$ is impossible. If $|\tan \varphi|$ is not small, it is given by formula (VI.49). The result may be written

$$\sqrt{\lambda} \cos\left(\int_a^x \frac{\mathrm{d}x}{\lambda} - \frac{\pi}{4} + \varphi\right) \rightarrow \sin \varphi \sqrt{l} \exp\left(\int_x^a \frac{\mathrm{d}x}{l}\right) \qquad \text{(VI.52)}$$

the arrow indicating the direction in which the connection is made.

In the case where the barrier is to the right, i.e. when $E \gtrless V$ according to whether $a \gtrless x$, the connection formulae (VI.51) and (VI.52) remain valid provided that one interchanges x and a in the inequalities and in the limits of integration; in particular, the direction of the arrows must be preserved.

10. Penetration of a Potential Barrier

By way of an illustration we shall apply the WKB method to the calculation of the transmission coefficient through the potential barrier shown in Fig. VI.3. In the region $x<a$ (region I), $V(x)=V_0=\text{const.}$; when $x>a$, $V(x)$ is a positive function decreasing monotonically from the positive value $V_a=V(a)$ to $V(\infty)=0$.

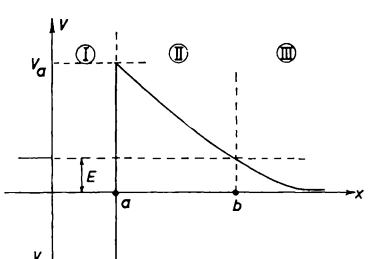


Fig. VI.3. Potential barrier $V(x)$.

Let E be the energy of the particle, b ($>a$) the point of the x axis such that $E=V(b)$. The point of discontinuity a and the turning point

b divide the x axis into three regions, I, II, and III. We assume the WKB method to be applicable in regions II and III.

To find the transmission coefficient, one must construct the solution of the Schrödinger equation whose asymptotic form in region III represents a purely transmitted wave (propagating in the direction of increasing x). In that region, the WKB solution is of the form (VI.42). The condition we impose upon its asymptotic form determines that solution (to within a constant), namely

$$y_{\mathrm{III}} = \sqrt{\lambda}\exp\left(\mathrm{i}\int_b^x \frac{\mathrm{d}x}{\lambda} - \mathrm{i}\frac{\pi}{4}\right) \qquad (x \gg b)$$

(the phase $\pi/4$ has been added to simplify subsequent calculations), or else

$$y_{\mathrm{III}} = \sqrt{\lambda}\left[\cos\left(\int_b^x \frac{\mathrm{d}x}{\lambda} - \frac{\pi}{4}\right) + \mathrm{i}\sin\left(\int_b^x \frac{\mathrm{d}x}{\lambda} - \frac{\pi}{4}\right)\right].$$

According to connection formulae (VI.49) and (VI.50), this particular solution extends into region II in the form

$$y_{\mathrm{II}} = -\mathrm{i}\sqrt{l}\exp\left(\int_x^b \frac{\mathrm{d}x}{l}\right) \qquad (a < x \ll b)$$

$$= -\mathrm{i}\sqrt{l}\,\mathrm{e}^{\tau}\exp\left(-\int_a^x \frac{\mathrm{d}x}{l}\right),$$

an expression in which

$$\tau = \int_a^b \frac{\mathrm{d}x}{l} = \int_a^b \frac{\sqrt{2m[V(x)-E]}}{\hbar}\,\mathrm{d}x.$$

Let us put

$$l_a = l(a) = \frac{\hbar}{\sqrt{2m(V_a - E)}}, \qquad k = \frac{\sqrt{2m(E-V_0)}}{\hbar}.$$

In region I, the solution of the Schrödinger equation is (rigorously) of the form

$$y_{\mathrm{I}} = A\sin[k(x-a)+\delta].$$

The constants A and δ are obtained by applying the continuity

conditions to the wave function and its logarithmic derivative at the point a, namely

$$k \cot \delta = -\frac{1}{l_a}, \qquad A \sin \delta = -\mathrm{i}\sqrt{l_a}\, \mathrm{e}^{\tau}. \tag{VI.53}$$

y_{I} is the sum of an incident and a reflected wave. Since δ is real, the incident wave $-\frac{1}{2}\mathrm{i}A \exp[\mathrm{i}(k(x-a)+\delta)]$ has a flux equal to $\frac{1}{4}|A|^2\hbar k/m$. According to equations (VI.53)

$$\frac{k|A|^2}{4} = \frac{kl_a}{4}\frac{\mathrm{e}^{2\tau}}{\sin^2\delta} = \frac{kl_a}{4}\mathrm{e}^{2\tau}(1+\cot^2\delta) = \mathrm{e}^{2\tau}\frac{1+k^2l_a^2}{4kl_a}.$$

Since the flux of the transmitted wave y_{III} is equal to $\hbar/m$, the transmission coefficient is

$$T = \frac{4kl_a}{1+k^2l_a^2}\mathrm{e}^{-2\tau}$$

$$= 4\frac{\sqrt{(V_a-E)(E-V_0)}}{V_a-V_0}\mathrm{e}^{-2\tau}.$$

This method of calculation is justified if $V(x)$ varies sufficiently slowly in regions II and III where the WKB approximation was made [condition (VI.47)]. Furthermore, in a region of several "wavelengths" about the turning point, $V(x)$ must be correctly approximated by a linear function of x. This requires in particular that the barrier have a "thickness" of at least several "wavelengths", hence that $\tau \gg 2\pi$ and consequently that T be extremely small ($\lesssim 10^{-5}$).

11. Energy Levels of a Potential Well

As a second application, we consider the potential well of Fig. VI.4 and we calculate the energy levels of the discrete spectrum.

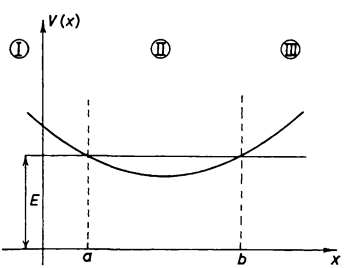


Fig. VI.4. Potential well $V(x)$.

For a given energy E there are two turning points a and b of the classical motion. They divide the x axis into three regions, I, II, and III. We look for the WKB solution which decreases exponentially in regions I and III, namely

$$y_{\mathrm{I}} = \tfrac{1}{2}c\sqrt{l}\,\exp\left(-\int_x^a \frac{\mathrm{d}x}{l}\right) \qquad (x \ll a)$$

$$y_{\mathrm{III}} = \tfrac{1}{2}c'\sqrt{l}\,\exp\left(-\int_b^x \frac{\mathrm{d}x}{l}\right) \qquad (x \gg b)$$

(c and c' are adjustable constants). In accordance with connection formula (VI.50) and the analogous formula corresponding to the barrier to the right, these functions are respectively continued into region II by the functions

$$\begin{aligned} y_a(x) &= c\sqrt{\bar{\lambda}}\,\cos\left(\int_a^x \frac{\mathrm{d}x}{\bar{\lambda}} - \frac{\pi}{4}\right) \\ y_b(x) &= c'\sqrt{\bar{\lambda}}\,\cos\left(\int_x^b \frac{\mathrm{d}x}{\bar{\lambda}} - \frac{\pi}{4}\right). \end{aligned} \qquad (a \ll x \ll b).$$

These functions are equal, $y_a(x) = y_b(x) = y_{\mathrm{II}}$, if

$$I \equiv \int_a^b \frac{\mathrm{d}x}{\bar{\lambda}} \equiv \int_a^b \frac{\sqrt{2m\,|E - V(x)|}}{\hbar}\,\mathrm{d}x = (N + \tfrac{1}{2})\pi, \qquad \text{(VI.54)}$$

where N is an integer ($\geqslant 0$). Rule (VI.54) fixes the energy levels of the discrete spectrum. In that case, one has $c' = (-1)^N c$. Since y_{I}, y_{II} and y_{III} are approximate solutions of the Schrödinger equation, they are, apart from a constant which can be easily calculated, good approximations of the eigensolutions in regions I, II and III, respectively. Thus, the method gives at the same time the energy levels and approximate expressions of the corresponding eigensolutions valid everywhere except in small regions surrounding the two turning points a and b.

The conditions of validity of the method require that both turning regions in which $V(x)$ is a reasonably linear function of x, extend

over several wavelengths; hence the method is applicable only for very large quantum numbers

$$N \gg 1.$$

In certain privileged cases, such as that of the harmonic oscillator, the method yields exact values of all the energy levels up to and including the ground state (cf. Problem VI.6). This must be regarded as an accident.

Quantization rule (VI.54) may be viewed as a condition for a standing wave: the interval (a, b) must contain a "half-integral" (i.e. $\frac{1}{2}$+integer) number of half-wavelengths. In fact, it differs from the Bohr-Sommerfeld quantization rule only by the presence of "half-integral" quantum numbers instead of integral quantum numbers. The integral I is just equal to $2\hbar$ times the action integral $\oint p\ \mathrm{d}q$ of the corresponding classical phase space. On the other hand, the action integral is equal to the extension in phase $\omega(E) \equiv \iint_{H \leqslant E} \mathrm{d}p\ \mathrm{d}q$ of the points whose energy is less than E. Rule (VI.54) can therefore be stated in the form

$$\omega(E) \equiv \oint_{H=E} p\ \mathrm{d}q = (N + \tfrac{1}{2})h. \tag{VI.55}$$

Following formula (VI.55), the extension in phase $\omega(E)$ increases by h when one goes from one energy level to the one immediately above. We deduce therefrom the important result concerning the distribution of energy levels:

The extension in phase of the energy band $(E, E+\delta E)$ measured in units of h *is equal to the number of bound states of the corresponding quantum system, located within that energy band.*

The conditions of validity of this result are actually less restrictive than those of the WKB method; in practice, it is always valid "in the limit of large quantum numbers". One commonly assumes – and it has actually been shown in the simplest cases – that this result also applies to systems whose number of dimensions R is greater than unity provided that one takes h^r as unit of extension in phase.

EXERCISES AND PROBLEMS

1. The Hamiltonian of the harmonic oscillator is $H = (p^2 + m^2\omega^2q^2)/2m$. Show that the mean values $\langle q \rangle$, $\langle p \rangle$ carry out sinusoidal oscillations of frequency

$\omega/2\pi$ about the origin; show that the squares of the deviations ϖ, χ (notation of § 3) oscillate sinusoidally with half the period about some (positive) average values and calculate these values. Under what conditions do ϖ and χ remain constant?

2. Show that the motion of the center of a wave packet representing a charged particle in an electromagnetic field is rigorously the same as that of a classical particle in the two following cases: (*a*) constant electric field; (*b*) constant magnetic field.

3. Show that the "spreading" of the wave packet representing a charged particle in a constant electric field obeys the same laws as that of a free wave packet (ϖ = const., χ = quadratic function of time).

4. With the wave function $\psi(\boldsymbol{r})$ of a particle, one forms the function

$$D(\boldsymbol{R}, \boldsymbol{P}) = \frac{1}{(2\pi\hbar)^3} \int \exp\left(-\frac{\mathrm{i}}{\hbar}\boldsymbol{P}\cdot\boldsymbol{r}\right) \psi^*\left(\boldsymbol{R} - \frac{\boldsymbol{r}}{2}\right) \psi\left(\boldsymbol{R} + \frac{\boldsymbol{r}}{2}\right) \mathrm{d}\boldsymbol{r}$$

which one interprets (after Wigner) as the density in phase of a classical statistical mixture associated with this wave function [1]). Show that:

(*i*) the distributions of position and momentum of this mixture are the same as those of the quantum particle in the state $\psi(\boldsymbol{r})$:

$$\int D(\boldsymbol{R}, \boldsymbol{P})\, \mathrm{d}\boldsymbol{P} = |\psi(\boldsymbol{R})|^2, \qquad \int D(\boldsymbol{R}, \boldsymbol{P})\, \mathrm{d}\boldsymbol{R} = |\varphi(\boldsymbol{P})|^2$$

[$\varphi(\boldsymbol{P})$ is wave function in momentum space];

(*ii*) if the particle is free, the evolution in time of the mixture is strictly that of a statistical mixture of free classical particles of the same mass;

(*iii*) find the "spreading" law of a free wave packet.

5. Calculate by the WKB method the transmission coefficient T of a particle of mass m and energy E through a slowly varying potential barrier $V(x)$ approaching limits less than E at the two boundaries of the interval $(-\infty, +\infty)$; assume that there exist only two turning points a and b. Show that

$$T = \exp\left(-2\int_a^b \frac{\sqrt{2m\,[V(x) - E]}}{\hbar}\,\mathrm{d}x\right).$$

6. Calculate the energy levels of the harmonic oscillator by the WKB method (Hamiltonian of Problem VI.1). Discuss the conditions of validity of the method.

---

[1]) In contrast to the density in phase of a true classical statistical mixture, $D(\boldsymbol{R}, \boldsymbol{P})$ may take on negative values.

GENERAL FORMALISM OF THE QUANTUM THEORY
(A) MATHEMATICAL FRAMEWORK

1. Superposition Principle and Representation of Dynamical States by Vectors

The entire interpretation of Wave Mechanics developed in Chapters IV and V had as its starting point the definition of the probability densities of position and momentum by means of the wave functions Ψ and Φ referring to configuration space and momentum space, respectively. We have already indicated the symmetrical role played in the theory by these two functions. The parallelism can be pushed all the way. Indeed, the fundamental postulates concerning the mean values (Ch. V, § 3) are just as well expressed by means of operations in momentum space. Instead of generalizing, as we have done, expressions (IV.13) and (IV.20) of the mean values of functions of the form $F(\boldsymbol{r})$ and $G(\boldsymbol{p})$, respectively, one can equally well carry out the same generalization in the corresponding expressions (IV.21) and (IV.14) constructed with the functions Φ. For the postulates (a) and (b) of Ch. V, § 3, one substitutes the equivalent postulates:

(a') With any dynamical variable $\mathscr{A} = A(q_1, \ldots, q_R;\ p_1, \ldots, p_R)$ there is associated the linear operator

$$A\left(i\hbar\frac{\partial}{\partial p_1}, \ldots, i\hbar\frac{\partial}{\partial p_R};\ p_1, \ldots, p_R\right).$$

(b') The average value taken by this dynamical variable when the system is in the dynamical state defined by the function $\Phi(p_1, \ldots, p_R)$ is

$$\langle A\rangle = \frac{\langle \Phi, A\Phi\rangle}{\langle \Phi, \Phi\rangle},$$

an expression in which the brackets on the right-hand side are scalar products of momentum space

$$\langle \Phi, A\Phi\rangle = \int \ldots \int \Phi^*(A\Phi)\, \mathrm{d}p_1 \ldots \mathrm{d}p_R, \qquad \langle \Phi, \Phi\rangle = \int \ldots \int \Phi^*\Phi\, \mathrm{d}p_1 \ldots \mathrm{d}p_R.$$

The equivalence of postulates (a), (b), and postulates (a'), (b') is based upon the properties of the Fourier transformation (cf. Appendix A). The proof of this equivalence is not difficult; we shall not undertake it here.

Just as the functions Φ and Ψ are equivalent *representations* of one and the same dynamical state, the observables

$$A\left(i\hbar\frac{\partial}{\partial p_1}, \ldots, i\hbar\frac{\partial}{\partial p_R}; p_1, \ldots, p_R\right) \text{ and } A\left(q_1, \ldots, q_R; \frac{\hbar}{i}\frac{\partial}{\partial q_1}, \ldots, \frac{\hbar}{i}\frac{\partial}{\partial q_R}\right)$$

are equivalent *representations* of the same dynamical variable (cf. Ch. IV, remark at the end of § 5). The entire theory of observables can be developed indifferently in one or the other of these representations. We thus obtain two equivalent formulations of the Quantum Theory.

These considerations become clearer if one imagines the wave functions Φ and Ψ to represent the same vector of a space with infinitely many dimensions. All the concepts introduced in Chapter V, such as function space, scalar product, norm, orthogonality, etc., then have a very simple geometrical interpretation. According to this picture, the values taken by the function $\Psi(q_1, \ldots, q_R)$ at each point of configuration space are the components of this vector along a certain system of orthogonal axes. Similarly, the values taken by the function $\Phi(p_1, \ldots, p_R)$ at each point of momentum space are the components of this same vector along another system of axes. Moreover, the coefficients $c_p^{(r)}$ of the expansion (V.14) of Ψ in a series of eigenfunctions of a given complete orthonormal system are the components of this vector along a third system of axes.

One is thus led to build up the entire Quantum Theory by starting directly from the vector concept, without reference to the particular representation which can be made thereof. In the spirit of such a formulation, one states as a fundamental principle of the theory the *superposition principle of dynamical states*, according to which the possible dynamical states of a given quantized system possess the following property characteristic of any wave in general: they may be superposed linearly and may consequently be represented by vectors of a certain linear space yet to be defined. Therefore, with each dynamical state is associated a vector of this abstract space. Likewise, with each dynamical variable is associated a linear operator of this space. In this presentation, the theory is formally simpler and more elegant than Wave Mechanics; it is also more general, because

it applies equally well to quantum systems having no classical analogues.

We shall now present this general formulation of the Quantum Theory. We essentially follow the presentation of Dirac, making use of the particularly convenient notation he has introduced for this purpose [1]). The present chapter is devoted to the concepts of linear algebra [2]) which constitute the mathematical framework of the theory. The description of physical phenomena by means of this formalism constitutes the subject of Chapter VIII.

I. VECTORS AND OPERATORS

2. Vector Space. "Ket" Vectors

Following the hypotheses of the foregoing paragraph, we associate with every dynamical state a certain type of vector which we call, following Dirac, *ket vector* or *ket* and which we represent by the symbol $|\rangle$. In order to distinguish the kets from each other, we complete each symbol by inserting either a particular letter, or one or even several indices capable of taking, depending upon the case, discrete or continuous values. Thus the ket u is represented by the symbol $|u\rangle$.

The kets form a linear vector space: any linear combination of several ket vectors is also a ket vector. For instance, let us take two kets $|1\rangle$ and $|2\rangle$ and two arbitrary complex numbers λ_1 and λ_2. The linear combination

$$|v\rangle = \lambda_1|1\rangle + \lambda_2|2\rangle \tag{VII.1}$$

is a vector of ket space.

In analogous fashion, if $|\xi\rangle$ depends on a continuous index, ξ, and if $\lambda(\xi)$ is an arbitrary complex function of ξ, the integral

$$|w\rangle = \int_{\xi_1}^{\xi_2} \lambda(\xi)\,|\xi\rangle\,\mathrm{d}\xi \tag{VII.2}$$

is a vector of ket space. We shall say, by an obvious extension of the terminology, that $|w\rangle$ is a linear combination (or linear superposition) of the kets $|\xi\rangle$.

---

[1]) P. A. M. Dirae, *loc. cit.*, footnote 3, p. 47, Ch. II.

[2]) The reader will find a rigorous and complete treatment of these questions in M. H. Stone, *loc. cit.*, footnote 1, p. 166, Ch. V; also, Paul R. Halmos, *Finite Dimensional Vector Spaces* (Princeton, Princeton University Press, 1948).

By definition, the kets of a given ensemble are *linearly independent* if none of them can be expressed in the form of a linear combination of the others [this linear combination can be of type (VII.1), of type (VII.2), or a mixture].

If the vector space contains at most n linearly independent vectors, it is a space with a finite number of dimensions, and the number of dimensions is by definition equal to n. Take n particular, linearly independent vectors in that space; all other vectors are equal to linear combinations of these n particular vectors.

If there is no limit to the number of linearly independent vectors of the vector space under consideration, the latter has an infinite number of dimensions. This is the case of Hilbert space, and as we have already seen, it is also the case of the wave functions of Wave Mechanics. One can nevertheless always choose a (denumerably or non-denumerably infinite) set of linearly independent vectors such that any vector of that space is equal to a linear combination (possibly an infinite series or an integral) of these "basis vectors".

Consider a set of ket vectors belonging to a given space $\mathscr{E}$. The ensemble of the kets of this set and of all linear combinations of these kets form a vector space $\mathscr{E}'$. $\mathscr{E}'$ is by definition the space subtended by the kets of this set. Any ket of $\mathscr{E}'$ belongs to $\mathscr{E}$. One says that $\mathscr{E}'$ is a *subspace* of $\mathscr{E}$. If the space has a finite number of dimensions n, the number of dimensions of $\mathscr{E}'$ is certainly finite and smaller than n. If $\mathscr{E}$ has an infinite number of dimensions, there exists no restriction whatsoever on the number of dimensions of $\mathscr{E}'$.

3. Dual Space. "Bra" Vectors

It is well known in linear algebra that with every vector space can be associated a dual vector space. Indeed, any linear function $\chi(|u\rangle)$ of the kets $|u\rangle$ possesses the superposition property characteristic of vectors [1]) and therefore defines a new type of vector which, following Dirac, we shall call *bra vector* or *bra* and which we shall represent by

---

[1]) The property of χ of being a linear function of the kets $|u\rangle$ implies that

$$\chi(\lambda_1|1\rangle + \lambda_2|2\rangle) = \lambda_1\chi(|1\rangle) + \lambda_2\chi(|2\rangle).$$

It is evident that if two functions χ_1, χ_2 have this property, any linear combination $\mu_1\chi_1 + \mu_2\chi_2$ of these two functions also has this property.

the symbol $\langle|$. Thus the function $\chi(|u\rangle)$ defines the bra $\langle\chi|$; the value taken by this function for a particular ket $|u\rangle$ is a number (generally complex) which we denote by the symbol $\langle\chi|u\rangle$.

By definition, the bra $\langle\Phi|$ vanishes if the function $\Phi(|u\rangle)$ vanishes for any $|u\rangle$:

$$\langle\Phi| = 0, \quad \text{if} \quad \langle\Phi|u\rangle = 0 \quad \text{for any} \quad |u\rangle. \tag{VII.3}$$

Similarly, two bras are equal,

$$\langle\Phi_1| = \langle\Phi_2|, \quad \text{if} \quad \langle\Phi_1|u\rangle = \langle\Phi_2|u\rangle \quad \text{for any} \quad |u\rangle.$$

If the ket space has a finite number of dimensions, the dual space has the same number of dimensions. If the number of dimensions of ket space is infinite, the dual space has the same property.

In order to introduce a metric in the vector space we have just defined, we make the hypothesis that there exists a one-to-one correspondence between the vectors of this space and those of the dual space. Bra and ket thus associated by this one-to-one correspondence are said to be conjugates of each other and are labelled by the same letter (or the same indices). Thus the bra conjugate to the ket $|u\rangle$ is represented by the symbol $\langle u|$.

We assume also that this correspondence is *antilinear*. In other words, the bra conjugate to the ket

$$|v\rangle = \lambda_1|1\rangle + \lambda_2|2\rangle \tag{VII.4}$$

is

$$\langle v| = \lambda_1^*\langle 1| + \lambda_2^*\langle 2|. \tag{VII.5}$$

Likewise, the bra conjugate to the ket

$$|w\rangle = \int_{\xi_1}^{\xi_2} \lambda(\xi)\,|\xi\rangle\,\mathrm{d}\xi \tag{VII.6}$$

is

$$\langle w| = \int_{\xi_1}^{\xi_2} \lambda^*(\xi)\,\langle\xi|\,\mathrm{d}\xi. \tag{VII.7}$$

Thus the correspondence between kets and bras is analogous to the correspondence between the wave functions of wave mechanics and their complex conjugates. Let us note in passing that if a ket vanishes, its conjugate bra is also zero, and conversely.

The ensemble of the bras conjugate to the kets of a subspace $\mathscr{E}'$ of the space $\mathscr{E}$ forms the dual subspace of $\mathscr{E}'$.

4. Scalar Product

By definition, the *scalar product* of the ket $|u\rangle$ by the ket $|v\rangle$ is the number (generally complex) $\langle v|u\rangle$, that is to say the value $v(|u\rangle)$ taken by the linear function associated with the bra conjugate to $|v\rangle$.

As an immediate consequence of the definition, the scalar product $\langle v|u\rangle$ is linear with respect to $|u\rangle$ and antilinear with respect to $|v\rangle$.

We adopt the hypothesis that the scalar product possesses all the other characteristic properties of the scalar product of the wave functions of Wave Mechanics (Ch. V, § 2), namely:

1) The scalar product of $|v\rangle$ by $|u\rangle$ is the complex conjugate of the scalar product of $|u\rangle$ by $|v\rangle$

$$\langle u|v\rangle = \langle v|u\rangle^*. \tag{VII.8}$$

2) Any vector u has a real non-negative norm $N_u \equiv \langle u|u\rangle$

$$\langle u|u\rangle \geqslant 0. \tag{VII.9}$$

It vanishes if and only if the vector u vanishes.

From these properties comes the *Schwarz inequality*: for any $|u\rangle$ and $|v\rangle$

$$|\langle u|v\rangle|^2 \leqslant \langle u|u\rangle \langle v|v\rangle. \tag{VII.10}$$

The equality holds if and only if the vectors $|u\rangle$ and $|v\rangle$ are collinear (i.e. proportional).

The preceding axioms must be complemented by the postulate that the space of the ket vectors $\mathscr{E}$ (as well as its dual space, bra space) is complete and separable (cf. Ch. V, § 2): *it is a Hilbert space.*

By definition two vectors are *orthogonal* if their scalar product vanishes. Two subspaces $\mathscr{E}_1$, $\mathscr{E}_2$ of $\mathscr{E}$ are orthogonal if each of the vectors in one is orthogonal to each of the vectors in the other. It is evident that $\mathscr{E}_1$ and $\mathscr{E}_2$ have no vector in common; indeed, any vector common to $\mathscr{E}_1$ and $\mathscr{E}_2$ can only be zero since, being orthogonal to itself, it has a vanishing norm.

The ensemble of vectors orthogonal to $\mathscr{E}_1$ form a subspace $\mathscr{E}_1^\times$ orthogonal to $\mathscr{E}_1$, the *complementary subspace* to $\mathscr{E}_1$. $\mathscr{E}_1^\times$ reduces to

zero if the subspace $\mathscr{E}_1$ coincides with the space $\mathscr{E}$ itself. One can show (cf. footnote 2, p. 245) that *any vector of* $\mathscr{E}$ *may be written*, and *in a unique manner, as the sum of a vector located in* $\mathscr{E}_1$ *and a vector located in its complementary space*:

$$|u\rangle = |u_1\rangle + |u_1^{\times}\rangle$$

$|u_1\rangle$ is by definition the *projection* of $|u\rangle$ in the subspace $\mathscr{E}_1$. We shall have occasion to return at length to this concept of projection in Section II.

In all our considerations on the scalar product, it is implicitly assumed that the vectors (ket or bra) have finite norm, lest the axiom concerning the norm of vectors lose all meaning. As long as this is understood, the space $\mathscr{E}$ of the ket vectors under consideration is a Hilbert space. We have seen in Chapter V that the vectors capable of representing dynamical states must actually have finite norm, but that the treatment of the continuous spectra in the eigenvalue problems requires the introduction of eigenvectors of infinite length. We must therefore also include in our vector space vectors $|\xi\rangle$ of infinite norm, depending on (at least) one continuous index, and extend the concept of scalar product to this category of vectors.

We assume that $|\xi\rangle$ has a finite scalar product:

$$\langle u|\xi\rangle$$

with any vector $|u\rangle$ of finite norm; moreover, this scalar product is linear in $|\xi\rangle$ and antilinear in $|u\rangle$. One defines the scalar product $\langle\xi|u\rangle$ in analogous fashion and we state as an axiom that

$$\langle\xi|u\rangle = \langle u|\xi\rangle^*.$$

On the other hand, the scalar product of two vectors of the type $|\xi\rangle$ may not converge. In particular, the norm of $|\xi\rangle$ is infinite. But we assume that the eigendifferential

$$|\xi, \delta\xi\rangle \equiv (\delta\xi)^{-\frac{1}{2}} \int_{\xi}^{\xi+\delta\xi} |\xi'\rangle \, \mathrm{d}\xi' \qquad \text{(VII.11)}$$

has a positive definite norm which tends toward a finite limit as $\delta\xi \to 0$. *Rigorously speaking*, the vector $|\xi\rangle$ does not belong to space $\mathscr{E}$; but its eigendifferentials, and more generally, the linear combi-

nations of the type (VII.2) do belong to this space and satisfy all the axioms relating to the vectors of a Hilbert space.

5. Linear Operators

Once we have defined the space of the ket vectors, we can define the linear operators of that space (cf. Ch. II, § 11).

Let us suppose that to each ket $|u\rangle$ of vector space there corresponds a certain ket $|v\rangle$: one says that $|v\rangle$ results from the action of a certain operator on the ket $|u\rangle$. If in addition this correspondence is linear, the operator thus defined is a certain *linear* operator A. One writes

$$|v\rangle = A|u\rangle.$$

Such an operator vanishes if the ket $|v\rangle$ is zero for any $|u\rangle$.

For an operator A to be zero, it is necessary and sufficient that for any $|u\rangle$,

$$\langle u|A|u\rangle = 0.$$

The proof of this property presents no serious difficulty and wil not be given here. One deduces immediately that:

For two operators A and B to be equal, it is necessary and sufficient that for any $|u\rangle$

$$\langle u|A|u\rangle = \langle u|B|u\rangle. \qquad \text{(VII.12)}$$

The main algebraic operations on the operators have already been defined (Ch. II, § 11): multiplication by a constant, sum, and product. The *sum* of linear operators is associative and commutative. The *product* is associative, distributive with respect to the sum, *but it is not commutative* – therein lies the fundamental difference between the algebra of linear operators and ordinary algebra. Let us recall that the commutator of two linear operators A and B is represented by the symbol

$$[A, B] \equiv AB - BA.$$

The main properties of commutator algebra were studied in Ch. V, § 17 [eqs. (V.63) to (V.66)]; they all remain valid and will not be re-examined here.

Note that the multiplication of a ket by a given constant c also defines a linear operator. This operator c commutes with all linear operators: $[A, c] = 0$.

In particular, the multiplication by unity is the "identity" operator.

If the correspondence defined above between $|u\rangle$ and $|v\rangle$ is reciprocal, it defines two linear operators A and B:

$$|v\rangle = A|u\rangle, \qquad |u\rangle = B|v\rangle. \tag{VII.13}$$

These operators are by definition the *inverse* of each other. One may also say that the operators A and B are each other's inverse if they simultaneously satisfy the equations

$$AB = 1, \qquad BA = 1. \tag{VII.14}$$

These two definitions are equivalent.

The inverse of a given operator A does not always exist. When it exists, one commonly denotes it by the symbol A^{-1}. Making use of the defining equations (VIII.14), one can easily prove the following property:

If two operators P and Q each possess an inverse, the product PQ has an inverse and one has

$$(PQ)^{-1} = Q^{-1} P^{-1} \tag{VII.15}$$

[note the reversal of the order of the factors on the right-hand side of (VII.15)].

Once we know the action of the linear operator A in the space of the ket vectors, its action in the dual vector space is defined unambiguously as follows. A bra $\langle\chi|$ having been given, the scalar product $\langle\chi|(A|u\rangle)$ is obviously a linear function of $|u\rangle$, since the operator A is linear. Let $\langle\eta|$ be the bra defined by this function; to each bra $\langle\chi|$ corresponds a bra $\langle\eta|$. It is clear that this correspondence is linear (property of the scalar product). One says that $\langle\eta|$ results from the operation of A on $\langle\chi|$ and one writes

$$\langle\eta| = \langle\chi|A. \tag{VII.16}$$

Following this definition, one has the important identity between scalar products

$$(\langle\chi|A)|u\rangle \equiv \langle\chi|(A|u\rangle). \tag{VII.17}$$

The parentheses are therefore unnecessary in these two expressions and we shall henceforth write $\langle\chi|A|u\rangle$ to denote indifferently one or the other of these two equal scalar products.

By means of the identity (VII.17) one can define the various operations of the algebra of linear operators acting on the bras. In particular, one has for the three fundamental operations:

(*a*) *multiplication of A by a complex constant c*:

$$(cA)|u\rangle = c(A|u\rangle) \quad \text{hence} \quad \langle\chi|(cA) = c(\langle\chi|A);$$

(*b*) *sum* of operators $S = A + B$:

$$S|u\rangle = A|u\rangle + B|u\rangle \quad \text{hence} \quad \langle\chi|S = \langle\chi|A + \langle\chi|B;$$

(*c*) *product* of operators $P = AB$:

$$P|u\rangle = A(B|u\rangle), \quad \text{hence} \quad \langle\chi|P = (\langle\chi|A)B.$$

We shall always adopt the convention of placing the bras to the left and the kets to the right of the operators; the algebraic manipulations on the linear operators are the same in the two cases.

Certain operators are particularly easy to handle when using the foregoing notations: they are the operators of the type $|u\rangle\langle v|$, whose action on any ket $|w\rangle$ yields a ket proportional to $|u\rangle$, namely the ket $|u\rangle\langle v|w\rangle$ (constant of proportionality $\langle v|w\rangle$), and whose action on any bra $\langle w|$ yields a bra proportional to $\langle v|$, namely the bra $\langle w|u\rangle\langle v|$. The operator $|u\rangle\langle v|$ has no inverse.

6. Tensor Product [1]) of Two Vector Spaces

To complete this introduction to vector algebra, we must still define a frequently used operation, namely the tensor product of two vector spaces.

The purpose and interest of this operation are illustrated by the following example.

Let us consider a quantum system of two particles. The product $\Psi_1(\boldsymbol{r}_1)\,\Psi_2(\boldsymbol{r}_2)$ of a wave function $\Psi_1(\boldsymbol{r}_1)$ relative to particle 1 by a wave function $\Psi_2(\boldsymbol{r}_2)$ relative to particle 2 represents a very special dynamical state of this system (Ch. IV, § 6). The most general wave function $\Psi(\boldsymbol{r}_1, \boldsymbol{r}_2)$ is not of this form, but can always be written as a linear combination of wave functions of this form. One of the many ways of doing this consists in expanding Ψ in a series of functions

[1]) This type of product is often called Kronecker product.

of a complete orthonormal system of functions of $\boldsymbol{r}_1$; since the expansion coefficients are functions of $\boldsymbol{r}_2$, each term of the series actually has the required form. The space of the wave functions of the total system is therefore formed by linear combination of products of wave functions relating to each of the partial systems $\Psi_1(\boldsymbol{r}_1)$ and $\Psi_2(\boldsymbol{r}_2)$. One says that the space of functions $\Psi(\boldsymbol{r}_1, \boldsymbol{r}_2)$ is the tensor product of the space of functions $\Psi_1(\boldsymbol{r}_1)$ and the space of functions $\Psi_2(\boldsymbol{r}_2)$.

In fact, the products $\Psi_1(\boldsymbol{r}_1)\,\Psi_2(\boldsymbol{r}_2)$ play a privileged role in the study of the total system. Indeed, the dynamical variables of particle 1 are represented by observables A_1 acting on the wave function $\Psi(\boldsymbol{r}_1, \boldsymbol{r}_2)$ considered as a function or $\boldsymbol{r}_1$; those of particle 2, by observables A_2 acting on this same function considered as a function of $\boldsymbol{r}_2$. Clearly, each of the observables A_1 commutes with each of the observables A_2. When Ψ is of the form $\Psi_1(\boldsymbol{r}_1)\,\Psi_2(\boldsymbol{r}_2)$ the action of observables of this type is especially simple; thus $A_1(\Psi_1\Psi_2)$ is equal to the product of $(A_1\Psi_1)$ by Ψ_2.

The foregoing remarks apply most generally to any quantized system capable of being analyzed into two simpler systems.

In the abstract language which we are developing in this chapter, the tensor product is defined in the following manner. Let $\mathscr{E}_1$, $\mathscr{E}_2$ be two vector spaces. Taking a ket $|u\rangle^{(1)}$ of the one and a ket $|u\rangle^{(2)}$ of the other, one can form the product ket $|u\rangle^{(1)}\,|u\rangle^{(2)}$. This multiplication is commutative and we use the notation

$$|u^{(1)}\,u^{(2)}\rangle \equiv |u\rangle^{(1)}\,|u\rangle^{(2)}. \tag{VII.18}$$

Moreover, we assume it to be distributive with respect to the sum. If

$$|u\rangle^{(1)} = \lambda|v\rangle^{(1)} + \mu|w\rangle^{(1)},$$
$$|u^{(1)}\,u^{(2)}\rangle = \lambda|v^{(1)}\,u^{(2)}\rangle + \mu|w^{(1)}\,u^{(2)}\rangle.$$

Similarly, if

$$|u\rangle^{(2)} = \lambda|v\rangle^{(2)} + \mu|w\rangle^{(2)},$$
$$|u^{(1)}\,u^{(2)}\rangle = \lambda|u^{(1)}\,v^{(2)}\rangle + \mu|u^{(1)}\,w^{(2)}\rangle.$$

The kets $|u^{(1)}\,u^{(2)}\rangle$ span a new vector space, the space $\mathscr{E}^{(1)}\otimes\mathscr{E}^{(2)}$, which one calls the *tensor product of the vector spaces* $\mathscr{E}^{(1)}$ *and* $\mathscr{E}^{(2)}$. If the latter have N_1 and N_2 dimensions, respectively, the number of dimensions of the product space is N_1N_2. However, the tensor product may also be carried out on spaces possessing an infinite

number of dimensions, as shown by the example given at the beginning of this paragraph.

To each linear operator $A^{(1)}$ of the space $\mathscr{E}^{(1)}$ there corresponds a linear operator of the product space which we designate by the same symbol. The action of $A^{(1)}$ on any ket $|u\rangle^{(1)}$ being known, namely

$$A^{(1)}|u\rangle^{(1)} = |v\rangle^{(1)},$$

its action on the particular kets $|u^{(1)}\, u^{(2)}\rangle$ of the product space is defined by

$$A^{(1)}|u^{(1)}\, u^{(2)}\rangle = |v^{(1)}\, u^{(2)}\rangle \tag{VII.19}$$

and its action on the general ket of the product space is deduced from it by linear superposition. Likewise, every linear operator $A^{(2)}$ of the space $\mathscr{E}^{(2)}$ permits the definition of a linear operator in the product space.

Each operator $A^{(1)}$ commutes with each of the operators $A^{(2)}$:

$$[A^{(1)}, A^{(2)}] = 0.$$

Indeed one verifies, using the definition of the operators $A^{(1)}$ and $A^{(2)}$ itself, that the action of $[A^{(1)}, A^{(2)}]$ on each vector $|u^{(1)}\, u^{(2)}\rangle$ spanning the product space yields a vanishing result:

$$A^{(1)}\, A^{(2)}|u^{(1)}\, u^{(2)}\rangle = |v^{(1)}\, v^{(2)}\rangle = A^{(2)}\, A^{(1)}|u^{(1)}\, u^{(2)}\rangle.$$

In the product space one can define the correspondence between kets and bras, the action of linear operators on the bras, and so on. The algebraic rules stated above remain valid for all algebraic operations in the product space. The proof of these results is not difficult and will not be given here.

II. HERMITEAN OPERATORS, PROJECTORS AND OBSERVABLES

7. Adjoint Operators and Conjugation Relations

From the one-to-one correspondence between bras and conjugate kets one can derive an analogous conjugation relation between linear operators.

Let A be a linear operator. Let $|v\rangle$ be the ket conjugate to the bra

$\langle u|A$. $|v\rangle$ depends antilinearly upon the bra $\langle u|$; it is therefore a linear function of $|u\rangle$. This linear correspondence defines a linear operator which goes by the name of *Hermitean conjugate* operator of A or *adjoint* operator of A, and which one denotes by the symbol $A^\dagger$:

$$|v\rangle = A^\dagger|u\rangle.$$

It is clear that $A^\dagger = 0$ if $A = 0$, and conversely.

Since $A^\dagger|u\rangle$ is the ket conjugate to the bra $\langle u|A$, the scalar product of this ket by a given arbitrary bra $\langle t|$ is the complex conjugate of the scalar product of $|t\rangle$ by $\langle u|A$ (property (VII.8)). We therefore have the very important conjugation relation

$$\langle t|A^\dagger|u\rangle = \langle u|A|t\rangle^*. \tag{VII.20}$$

Since this relation holds for any $|u\rangle$ and $|t\rangle$, the ket conjugate to $\langle t|A^\dagger$ is necessarily equal to $A|t\rangle$. Consequently, the Hermitean conjugate of $A^\dagger$ is the operator A itself. Hermitean conjugation is a *reciprocal* operation:

$$(A^\dagger)^\dagger = A. \tag{VII.21}$$

In an analogous manner one obtains the following fundamental relations:

$$(cA)^\dagger = c^* A^\dagger \tag{VII.22}$$

$$(A+B)^\dagger = A^\dagger + B^\dagger \tag{VII.23}$$

$$(AB)^\dagger = B^\dagger A^\dagger. \tag{VII.24}$$

Note the reversal in the order of the operators in the expression (VII.24) for the adjoint of the product AB. Moreover, the adjoint of the operator $|u\rangle\langle v|$ is

$$(|u\rangle\langle v|)^\dagger = |v\rangle\langle u|. \tag{VII.25}$$

Hermitean conjugation is for operators what the conjugation between bras and kets is for vectors, and complex conjugation for ordinary numbers. All these conjugation operations play an important role in the formalism. The notation of Dirac makes it possible to carry them out without difficulty for any algebraic expression. One merely has to obey the following very simple rule: *replace everywhere the numbers by their complex conjugate, bras by the conjugate kets and*

vice versa, *the operators by their Hermitean conjugates, and reverse in each term the order of the various symbols occurring there, that is to say the order of the bras, kets, and operators.*

This rule is the obvious generalization of the properties (VII.20), (VII.24) and (VII.25). Let us give some examples of its application. The Hermitean conjugate of the operator $AB|u\rangle\langle v|C$ is $C^\dagger|v\rangle\langle u|B^\dagger A^\dagger$; the ket $AB|u\rangle\langle v|C|w\rangle$ has as its conjugate bra $\langle w|C^\dagger|v\rangle\langle u|B^\dagger A^\dagger$; the complex conjugate of $\langle t|AB|u\rangle\langle v|C|w\rangle$ is $\langle w|C^\dagger|v\rangle\langle u|B^\dagger A^\dagger|t\rangle$, and so forth.

8. Hermitean (or Self-Adjoint) Operators, Positive Definite Hermitean Operators, Unitary Operators

By definition a linear operator H is *Hermitean* if it is its own adjoint:

$$H = H^\dagger.$$

An operator I is *anti-Hermitean* if it has the opposite sign of its adjoint:

$$I = -I^\dagger.$$

From these definitions, one easily derives the following properties.

Any linear operator can be put uniquely in the form of a sum of two operators, one Hermitean, the other anti-Hermitean

$$A = H_A + I_A \tag{VII.26}$$

and one has

$$H_A = \frac{A + A^\dagger}{2}, \qquad I_A = \frac{A - A^\dagger}{2}. \tag{VII.27}$$

Any linear combination with *real* coefficients of Hermitean operators is Hermitean. *The product HK of two Hermitean operators H and K is not necessarily Hermitean* since, according to equation (VII.24)

$$(HK)^\dagger = KH \tag{VII.28}$$

HK is Hermitean if and only if H and K commute. Moreover, the commutator $[H, K]$ is anti-Hermitean and the decomposition (VII.26) of the product HK is written

$$HK = \frac{HK + KH}{2} + \frac{1}{2}[H, K]. \tag{VII.29}$$

The operator $|a\rangle\langle a|$ is a particular Hermitean operator. With two distinct kets $|a\rangle$ and $|b\rangle$, one can form the two Hermitean operators

$|a\rangle\langle a|$ and $|b\rangle\langle b|$, but the product of these two operators $|a\rangle\langle a|b\rangle\langle b|$ is proportional to $|a\rangle\langle b|$ which is not Hermitean; the product is therefore not Hermitean (except if $|a\rangle$ and $|b\rangle$ are orthogonal, in which case this product vanishes).

By definition the *Hermitean* operator H is *positive definite* if

$$\langle u|H|u\rangle \geqslant 0 \quad \text{for any} \quad |u\rangle.$$

The operator $|a\rangle\langle a|$ is a particular positive definite Hermitean operator.

The operators of this type have remarkable properties (cf. Problems VII.7 and VII.8). In particular, if H is positive definite Hermitean, one has the generalized Schwarz inequality:

$$|\langle u|H|v\rangle|^2 \leqslant \langle u|H|u\rangle\langle v|H|v\rangle,$$

which holds for any $|u\rangle$ and $|v\rangle$; the equality obtains if and only if $H|u\rangle$ and $H|v\rangle$ are proportional. Furthermore, the equality

$$\langle u|H|u\rangle = 0$$

necessarily implies: $H|u\rangle = 0$.

An operator U is *unitary* if it is the inverse of its own adjoint:

$$UU^\dagger = U^\dagger U = 1.$$

The product $W = UV$ of two unitary operators U, V is a unitary operator. In fact [properties (VII.15) and (VII.24)],

$$W^{-1} = V^{-1}U^{-1} = V^\dagger U^\dagger = W^\dagger.$$

9. Eigenvalue Problem and Observables

Let A be a linear operator. By definition the complex number a is an eigenvalue of A and the ket $|u\rangle$ an eigenket associated with a if

$$A|u\rangle = a|u\rangle.$$

Similarly $\langle u'|$ is an eigenbra associated with the eigenvalue a' if

$$\langle u'|A = a'\langle u'|.$$

If $|u\rangle$ is an eigenket of A, any multiple $c|u\rangle$ of this ket vector is an

eigenket of A belonging to the same eigenvalue; if there exist several linearly independent eigenkets belonging to the same eigenvalue a, any linear combination of these kets is an eigenket of A belonging to this eigenvalue. In other words, the ensemble of eigenkets of A belonging to a given eigenvalue a forms a vector space; we shall call it the *subspace of the eigenvalue a*. If this subspace has but one dimension, the eigenvalue is said to be single or non-degenerate. If not, we have *degeneracy*, the order of the degeneracy being by definition the number of dimensions of this subspace; the degeneracy may be of infinite order.

Similar remarks apply to the eigenbras of A. If A is arbitrary, there is no simple relation between the eigenvalue problem for the kets, and the eigenvalue problem for the bras. However, these two problems are closely related if A is Hermitean, the only case of practical interest.

If A is Hermitean, one has in fact the following properties:

(*i*) the two eigenvalue spectra are identical;

(*ii*) all the eigenvalues are *real*;

(*iii*) any bra conjugate to an eigenket of A is an eigenbra belonging to the same eigenvalue, and *vice versa*; in other words, the subspace of the eigenbras of a given eigenvalue is the dual space of the subspace of the eigenkets of the same eigenvalue.

The proof of property (*ii*) is the same as the one given in Ch. V, § 5, except for notation. If $A = A^\dagger$ and if $A|u\rangle = a|u\rangle$,

$$\langle u|A|u\rangle = a\langle u|u\rangle,$$

and since

$$\langle u|A|u\rangle^* = \langle u|A^\dagger|u\rangle = \langle u|A|u\rangle,$$

$\langle u|A|u\rangle$ is real, as well as $\langle u|u\rangle$; hence a is real. An analogous proof can be given for an eigenvalue belonging to a bra.

Moreover, since any eigenvalue is real, $A|u\rangle = a|u\rangle$ necessarily implies $\langle u|A = a\langle u|$, and conversely; from this one easily derives the properties (*i*) and (*iii*).

Another important property of the eigenvectors of a Hermitean operator is the *orthogonality* property of the eigenvectors belonging to distinct eigenvalues. The proof here is also patterned after the

one of Ch. V, § 5. If $|u\rangle$ and $|v\rangle$ are eigenkets belonging to the eigenvalues a and b, respectively,

$$A|u\rangle = a|u\rangle, \qquad \langle v|A = b\langle v|.$$

Multiplying the first equation from the left by $\langle v|$ (scalar product), the second equation by $|u\rangle$ from the right, and subtracting term by term, one obtains

$$0 = (a-b)\,\langle v|u\rangle.$$

Consequently, if $a \neq b$,

$$\langle v|u\rangle = 0.$$

In all these arguments it is implicitly assumed that the eigenvectors belong to Hilbert space. Stated in this way, the eigenvalue problem is too restrictive to satisfy all the requirements of the Theory. One must also consider as acceptable eigensolution, the vectors of infinite norm subject to the conditions stated at the end of Ch. VII, § 4. These vectors are associated with the eigenvalues of the *continuous spectrum.*

The difficulties of the continuous spectrum were discussed at length in Chapter V, and we shall not return to them. The principal results of this discussion are easily translated into our new language. The properties (*i*), (*ii*), and (*iii*) remain true in the case of the continuous spectrum. As to the orthogonality property, it is conveniently expressed by means of the Dirac δ-"function".

By way of an example, let us return to the Hermitean operator of Ch. V, § 9 whose eigenvalue spectrum is made up of a series of discrete eigenvalues a_n and a continuous part $a(\nu)$. The eigenfunctions $\varphi_n^{(r)}$ belonging to the eigenvalue a_n represent orthonormal kets which we denote by the symbols $|nr\rangle$, respectively. Likewise, the eigenfunctions $\varphi^{(r)}(\nu, \varrho)$ represent the kets $|\nu\varrho r\rangle$. The orthonormality relations between these various kets are written [eq. (V.38)]:

$$\langle nr|n'r'\rangle = \delta_{nn'}\,\delta_{rr'} \tag{VII.30a}$$

$$\langle nr|\nu'\varrho'r'\rangle = 0 \tag{VII.30b}$$

$$\langle \nu\varrho r|\nu'\varrho'r'\rangle = \delta(\nu-\nu')\,\delta(\varrho-\varrho')\,\delta_{rr'}. \tag{VII.30c}$$

If the ensemble of these vectors spans the entire space, in other words if any vector of finite norm can be expanded in a series (or integral) of these vectors, they are said to form a *complete* set and

the Hermitean operator is called an *observable*. Observables are the only Hermitean operators of $\mathscr{E}$-space capable of having a physical interpretation.

The important problem of knowing if a given Hermitean operator is an observable is in most cases a difficult mathematical problem. However, for a very important class of operators this problem is easily solved; namely the projection operators, or projectors for short.

10. Projectors (Projection Operators)

Let $\mathscr{S}$ be a subspace of Hilbert space $\mathscr{E}$, and $\mathscr{S}^\times$ its complementary subspace. Any ket $|u\rangle$ possesses a projection in $\mathscr{S}$ and a projection in $\mathscr{S}^\times$; these two vectors, $|u_S\rangle$ and $|u_S^\times\rangle$, are uniquely defined and one has

$$|u\rangle = |u_S\rangle + |u_S^\times\rangle. \tag{VII.31}$$

To each ket $|u\rangle$ there thus corresponds one and only one ket $|u_S\rangle$. This correspondence is easily seen to be linear. The linear operator P_S thus defined is called the *projector* (or *projection operator*) *on the subspace* $\mathscr{S}$

$$P_S|u\rangle = |u_S\rangle.$$

It is a Hermitean operator. Indeed, we have for any $|v\rangle$

$$\langle u|P_S|v\rangle = \langle u|v_S\rangle = \langle u_S|v_S\rangle = \langle u_S|v\rangle$$

hence

$$\langle u|P_S = \langle u_S|.$$

P_S is evidently an observable possessing a total of two eigenvalues 0 and 1, whose subspaces are respectively $\mathscr{S}^\times$ and $\mathscr{S}$.

Moreover, since, for any $|u\rangle$

$$P_S^2|u\rangle = P_S(P_S|u\rangle) = P_S|u_S\rangle = |u_S\rangle = P_S|u\rangle,$$

P_S satisfies the operator equation

$$P_S^2 = P_S.$$

Conversely, *any Hermitean operator* P *satisfying the equation*

$$P^2 = P \tag{VII.32}$$

is a projector. The subspace $\mathscr{S}$ *on which it projects is the subspace of its eigenvalue* 1.

Indeed, if p is an eigenvalue of this operator, and $|p\rangle$ one of the corresponding eigenvectors,

$$P|p\rangle = p|p\rangle.$$

According to eq. (VII.32)

$$0 = (P^2 - P)|p\rangle = (p^2 - p)|p\rangle$$

and since the ket $|p\rangle$ does not vanish, $p^2 - p = 0$. In other words, the only possible eigenvalues of P are equal to 0 or 1.

Furthermore, P is an observable since any vector $|u\rangle$ can be put in the form of a sum of eigenvectors of P. Indeed one has

$$|u\rangle = P|u\rangle + (1 - P)|u\rangle. \tag{VII.33}$$

The vector $P|u\rangle$ is an eigenket of P belonging to the eigenvalue 1, since by virtue of eq. (VII.32)

$$P^2|u\rangle \equiv P(P|u\rangle) = P|u\rangle.$$

The vector $(1-P)|u\rangle$ is an eigenket of P belonging to the eigenvalue 0, since likewise

$$P(1 - P)|u\rangle = (P - P^2)|u\rangle = 0.$$

One can easily verify that the vectors $P|u\rangle$ and $(1-P)|u\rangle$ are orthogonal and, therefore, that the sum of their norms is equal to the norm of $|u\rangle$. These two vectors thus certainly have a finite norm: they belong to Hilbert space.

Let $\mathscr{S}$ be the subspace of the eigenvectors of P belonging to the eigenvalue 1. The subspace $\mathscr{S}^\times$, which is complementary to $\mathscr{S}$, i.e. the subspace of the vectors which are orthogonal to the vectors of the subspace $\mathscr{S}$, is composed of the sum total of the eigenvectors of P belonging to the eigenvalue 0. Following the decomposition (VII.33), the action of P on an arbitrary vector $|u\rangle$ yields the projection of this vector on $\mathscr{S}$. P is therefore just the projector on $\mathscr{S}$. It is clear that $(1-P)$ is the projector on $\mathscr{S}^\times$.

The property relating to the norm of $P|u\rangle$ stated earlier can be rewritten

$$0 \leqslant \langle u|P|u\rangle \leqslant \langle u|u\rangle. \tag{VII.34}$$

If $\langle u|P|u\rangle = 0$, $|u\rangle$ lies entirely in $\mathscr{S}^\times$.

If $\langle u|P|u\rangle = \langle u|u\rangle$, $|u\rangle$ lies entirely in $\mathscr{S}$.

Two extreme cases deserve mention. When the projection space $\mathscr{S}$

is the space $\mathscr{E}$ itself, any ket $|u\rangle$ is its own projection: one has $\langle u|P|u\rangle = \langle u|u\rangle$ for any $|u\rangle$; the complementary space $\mathscr{S}^\times$ is zero. This is the case $P=1$.

The other extreme case is the one where the space $\mathscr{S}$ vanishes (the complementary space $\mathscr{S}^\times$ then being the space $\mathscr{E}$ itself): $\langle u|P|u\rangle = 0$ for any $|u\rangle$. This is the case $P=0$.

As an illustration, we give some typical examples of projectors.

Let $|a\rangle$ be a ket normalized to unity. It spans a one-dimensional subspace. Let us denote by $|u_a\rangle$ the projection of an arbitrary vector $|u\rangle$ in this subspace:

$$|u\rangle = |u_a\rangle + |u_a^\times\rangle. \tag{VII.35}$$

By hypothesis

$$\langle a|u_a^\times\rangle = 0 \qquad |u_a\rangle = c|a\rangle.$$

Multiplying both sides of eq. (VII.35) from the left by $\langle a|$ (scalar product), we have $c=\langle a|u\rangle$. Hence

$$|u_a\rangle = |a\rangle\langle a|u\rangle.$$

Consequently, the projector on $|a\rangle$ is the operator

$$P_a \equiv |a\rangle\langle a| \qquad (\langle a|a\rangle = 1). \tag{VII.36}$$

Projection operators of this type will be called *elementary*.

Consider now a set of orthonormal vectors $|1\rangle, |2\rangle, \ldots, |N\rangle$

$$\langle m|n\rangle = \delta_{mn} \qquad (m, n = 1, 2, \ldots, N).$$

These vectors span a certain subspace $\mathscr{E}_1$ (with N dimensions) of the ket vector space to which they belong. One easily verifies that the operator

$$P_1 \equiv \sum_{m=1}^{N} |m\rangle\langle m| \tag{VII.37}$$

is the projector on $\mathscr{E}_1$.

All the vectors considered thus far have finite norm. But one might also contemplate a set of ket vectors $|\xi\rangle$ depending upon the continuous index ξ varying in a certain domain (ξ_1, ξ_2). We assume that the eigen-

differentials formed with these kets have finite norm and belong to the Hilbert space under study. Therefore, any linear combination of these vectors also belongs to Hilbert space. The ensemble of these linear combinations forms a certain subspace $\mathscr{E}_2$ of the total Hilbert space: $\mathscr{E}_2$ is the subspace spanned by the vectors $|\xi\rangle$. Let us further suppose that these vectors satisfy the "orthonormality" condition

$$\langle\xi'|\xi\rangle = \delta(\xi' - \xi). \tag{VII.38}$$

Clearly

$$P_2 \equiv \int_{\xi_1}^{\xi_2} |\xi\rangle \, \mathrm{d}\xi \, \langle\xi| \tag{VII.39}$$

is the projector upon $\mathscr{E}_2$. Indeed, the vector

$$P_2|u\rangle \equiv \int_{\xi_1}^{\xi_2} |\xi\rangle \, \mathrm{d}\xi \, \langle\xi|u\rangle$$

obtained by letting P_2 act on an arbitrarily chosen vector $|u\rangle$ is certainly located in $\mathscr{E}_2$ since it is a linear combination of vectors $|\xi\rangle$; the difference $(1 - P_2)|u\rangle$ is orthogonal to all vectors of the set $|\xi\rangle$:

$$\begin{aligned}\langle\xi'|(1 - P_2)|u\rangle &= \langle\xi'|u\rangle - \int_{\xi_1}^{\xi_2} \langle\xi'|\xi\rangle \, \mathrm{d}\xi \, \langle\xi|u\rangle \\ &= \langle\xi'|u\rangle - \int_{\xi_1}^{\xi_2} \delta(\xi' - \xi) \, \mathrm{d}\xi \, \langle\xi|u\rangle = 0.\end{aligned}$$

and hence orthogonal to $\mathscr{E}_2$.

11. Projector Algebra

The projectors of Hilbert space are of great interest in view of their very simple geometrical interpretation [1]). We shall give here the main properties of the algebra of these operators. Since the proofs are for the most part elementary, we shall merely state the principle of these proofs, leaving the task of completing them to the reader.

---

[1]) The treatment of the continuous spectrum by the method of von Neumann is based on the systematic study of the properties of the projectors of Hilbert space; in this way one succeeds in surmounting all the difficulties of the continuous spectrum without ever leaving Hilbert space; cf. von Neumann, *loc. cit.*, p. 157, Ch. V.

Let P_i, P_j be the projectors upon the subspaces $\mathscr{E}_i, \mathscr{E}_j$ of the Hilbert space $\mathscr{E}$. In order that the *product*

$$P_{[ij]} \equiv P_i P_j$$

also be a projector, it is necessary and sufficient that P_i and P_j commute.

The condition is necessary, otherwise $P_{[ij]}$ would not be Hermitean. It is sufficient since, in that case, $P_{[ij]}$ is Hermitean and

$$P_{[ij]}^2 = P_i P_j P_i P_j = P_i^2 P_j^2 = P_i P_j = P_{[ij]}.$$

The subspace $\mathscr{E}_{[ij]}$ which corresponds to this is the *intersection* of the subspaces $\mathscr{E}_i$ and $\mathscr{E}_j$, that is to say the subspace of the vectors common to $\mathscr{E}_i$ and $\mathscr{E}_j$. Two extreme cases can arise, the one where $\mathscr{E}_{[ij]}$ is identical with one of the two subspaces from which we started, $\mathscr{E}_j$, say, and the one where $\mathscr{E}_{[ij]}$ vanishes. In the first case, $\mathscr{E}_j$ is a subspace of $\mathscr{E}_i$; in the second case $\mathscr{E}_i$ and $\mathscr{E}_j$ are orthogonal.

One easily derives the two following properties.

In order that $\mathscr{E}_j$ be a subspace of $\mathscr{E}_i$ (i.e. in order that any vector of the subspace $\mathscr{E}_j$ be a vector of the subspace $\mathscr{E}_i$), the necessary and sufficient condition is that

$$P_i P_j = P_j.$$

For $\mathscr{E}_i$ and $\mathscr{E}_j$ to be orthogonal, it is necessary and sufficient that

$$P_i P_j = 0. \tag{VII.40}$$

In that case, by an extension of the meaning of the term, we say that the *projectors are orthogonal.*

Concerning the *sum of projectors*, one has the important theorem:

Let $P_i, P_j, P_k, \ldots$ be the projectors upon the subspaces $\mathscr{E}_i, \mathscr{E}_j, \mathscr{E}_k, \ldots$ respectively. In order that their sum $P_i + P_j + P_k + \ldots$ likewise be a projector, it is necessary and sufficient that these operators be mutually orthogonal. The subspace upon which the projection is carried out is the *direct sum* of the subspaces $\mathscr{E}_i, \mathscr{E}_j, \mathscr{E}_k, \ldots$ (i.e. the ensemble of vectors formed by linear superposition of vectors belonging to any one of these partial subspaces).

The condition of orthogonality is obviously sufficient.

To show that it is necessary, it suffices to show that, if the sum $S = P_i + P_j + P_k + \ldots$ is a projector, the action of $P_j P_i$ on every ket vector $|u\rangle$ of ξ_i gives 0. To this effect, we apply property (VII.34) to the projectors S and P_i, P_j, P_k, ... The upper limit of the sum

$$\langle u|S|u\rangle = \langle u|P_i|u\rangle + \langle u|P_j|u\rangle + \langle u|P_k|u\rangle + \ldots$$

is $\langle u|u\rangle$ and all its terms are positive. If $P_i|u\rangle = |u\rangle$, the first term is equal to this upper limit and all the other terms vanish. Further, since $\langle u|P_j|u\rangle = 0$ implies $P_j|u\rangle = 0$, we have:

$$P_j P_i|u\rangle = P_j|u\rangle = 0.$$

The operator P_1 defined by eq. (VII.37) is a sum of orthogonal projectors. Specifically, the projectors $|m\rangle\langle m|$ which occur in this sum are elementary projectors. It is quite clear that the space $\mathscr{E}_1$ upon which the projection is carried out is the direct sum of the spaces upon which the projections relating to each term of the sum are carried out. Being the direct sum of N one-dimensional spaces, $\mathscr{E}_1$ is a space with N dimensions, while the operator P_1 is the sum of N elementary orthogonal projectors.

If $N \neq 1$, P_1 can be put into this form in infinitely many ways. Indeed, designate by $\{n\}$ a set $|1\rangle, |2\rangle, \ldots, |N\rangle$ of N orthonormal vectors of $\mathscr{E}_1$. The set $\{n\}$ forms a basis in $\mathscr{E}_1$, in the sense that any vector of $\mathscr{E}_1$ can be expressed as a linear function of these N vectors. We adopt the convention of not considering two bases distinct whose vectors differ only by a phase factor or by the relative order in which they occur. Now it is clear that

$$P_1 = \sum_{n=1}^{N} |n\rangle\langle n|$$

and that there are as many expressions of P_1 of this type as there are distinct bases.

These considerations can be extended with only minor changes to the case where the subspace $\mathscr{S}$ upon which the projection is performed possesses an infinite number of dimensions. One shows in the theory of Hilbert space that it is always possible to choose in $\mathscr{S}$, in an infinite number of ways, a basis $\{n\} \equiv \{|1\rangle, |2\rangle, \ldots, |n\rangle, \ldots\}$ containing a denumerably infinite set of orthonormal vectors. The projector P on $\mathscr{S}$ can then be put *in the form of a series of elementary, orthogonal projectors*:

$$P = \sum_{n=1}^{\infty} |n\rangle\langle n|.$$

However, one may also form a basis in $\mathscr{S}$ containing kets depending upon a continuous index. Let us suppose, for instance, that there exists a (non-denumerably infinite) set of vectors $|\xi\rangle$ of infinite norm, depending on the continuous index ξ and satisfying the "orthonormality" relations (VII.38); let us further suppose that the subspace $\mathscr{S}$ is made up of the sum total of the vectors of finite norm formed by linear superposition of the kets $|\xi\rangle$ of a certain domain (ξ_1, ξ_2). In that case, one can also write P in the form of the type (VII.39),

$$P = \int_{\xi_1}^{\xi_2} |\xi\rangle \, \mathrm{d}\xi \, \langle\xi|.$$

In that form P can still be regarded as a sum of orthogonal projectors. Let us subdivide the domain of integration (ξ_1, ξ_2) into a number of partial domains. P is the sum of the projectors obtained by integrating $|\xi\rangle \, \mathrm{d}\xi \, \langle\xi|$ over each of these partial domains. The latter can actually be arbitrarily small. Denote by δP the operator obtained by integration over the infinitesimal domain $(\xi, \xi + \delta\xi)$:

$$\delta P = \int_{\xi}^{\xi+\delta\xi} |\xi'\rangle \, \mathrm{d}\xi' \, \langle\xi'|.$$

P is the sum of an infinitely large number of operators of the type δP. We shall call the operators of the type δP *differential projectors*; the projection space corresponding to this type of operator has an infinity of dimensions.

12. Observables Possessing an Entirely Discrete Spectrum

Let A be a Hermitean operator. In this section, we shall study the eigenvalue problem, limiting ourselves to the eigenvectors located in Hilbert space. The eigenvalues then form a discrete set $a_1, a_2, \ldots, a_n, \ldots$. We designate by $\mathscr{E}_n$ the subspace belonging to the eigenvalue a_n, by P_n the projector upon this subspace. If the eigenvalue a_n is not degenerate, $\mathscr{E}_n$ has only one dimension and P_n is an elementary projector. If it is degenerate, one can always choose in an infinity of ways a basis in $\mathscr{E}_n$, $|n1\rangle, |n2\rangle, \ldots, |nr\rangle, \ldots$, and one has

$$P_n = \sum_r |nr\rangle \langle nr|. \qquad \text{(VII.41)}$$

The subspaces $\mathscr{E}_n, \mathscr{E}_{n'}$, belonging to two distinct eigenvalues $a_n, a_{n'}$, are orthogonal, consequently

$$P_n P_{n'} = 0 \qquad (n \neq n'). \qquad \text{(VII.42)}$$

Carrying out the summation of the projectors P_n corresponding to all the eigenvalues of the discrete spectrum, one obtains the projector

$$P_A \equiv \sum_n P_n \tag{VII.43}$$

whose subspace of projection $\mathscr{E}_A$ is formed of the direct sum of all the $\mathscr{E}_n$; $\mathscr{E}_A$ is the vector space formed by linear superposition of the eigenkets of A located in Hilbert space.

If A is an observable and if its spectrum is entirely discrete, $\mathscr{E}_A$ by definition coincides with the total space $\mathscr{E}$, in other words

$$P_A \equiv \sum_n P_n = 1. \tag{VII.44}$$

One sometimes calls the left-hand side of (VII.44) the *decomposition of unity* with respect to the eigenvalues of A. It is clear that this decomposition is unique, hence that any ket $|u\rangle$ can be written in a unique way as a sum of eigenkets of A each of which belongs to a different eigenvalue. To write down this sum, one merely operates on $|u\rangle$ with each side of eq. (VII.44) and equates the two kets thus obtained, namely:

$$|u\rangle = \sum_n P_n|u\rangle. \tag{VII.45}$$

According to the definition of P_n, $P_n|u\rangle$ is either zero, or an eigenket of A belonging to the eigenvalue a_n, for any $|u\rangle$. Therefore

$$(A - a_n)P_n = 0. \tag{VII.46}$$

Multiplying eq. (VII.44) term by term by A, one obtains, taking into account eqs. (VII.46)

$$A = \sum_n a_n P_n. \tag{VII.47}$$

According to this equation the observable A is completely defined once its eigenvalues and their respective subspaces are given. Furthermore, from expression (VII.47) it is obvious that the operator A commutes with all the projectors P_n.

The relations (VII.44), (VII.45), and (VII.47) are characteristic of observables possessing an entirely discrete spectrum, whether the number of their eigenvalues be finite or infinite. We shall not elaborate here on the question of convergence of the series; in practice that convergence always obtains.

Especially convenient expressions result by substituting everywhere for every P_n its expression (VII.41). Thus the left-hand side of eq. (VII.44) is written in the form of a sum of elementary projectors, and one obtains the *closure relation*

$$P_A \equiv \sum_{n,\,r} |nr\rangle \langle nr| = 1. \tag{VII.48}$$

Together with the orthogonality relations

$$\langle nr|n'r'\rangle = \delta_{nn'}\,\delta_{rr'} \tag{VII.49}$$

this expresses the fact that the assembly of vectors $|nr\rangle$ forms a complete orthonormal set.

Upon applying the closure relation to any vector $|u\rangle$, one obtains the expansion of $|u\rangle$

$$|u\rangle = \sum_{n,\,r} |nr\rangle \langle nr|u\rangle \tag{VII.50}$$

in a series of eigenvectors $|nr\rangle$. The coefficients of the expansion are equal to the scalar products $\langle nr|u\rangle$ [cf. eqs. (V.14) and (V.15)]. Moreover

$$\langle u|u\rangle = \langle u|P_A|u\rangle = \sum_{n,\,r} \langle u|nr\rangle \langle nr|u\rangle = \sum_{n,\,r} |\langle nr|u\rangle|^2. \tag{VII.51}$$

The norm of $|u\rangle$ is equal to the sum of the squares of the absolute values of the expansion coefficients: this is the Parseval relation [cf. eq. (V.16)].

The observable A can be written in the form of a series of elementary, orthogonal projectors. By a procedure analogous to the one which yielded eq. (VII.47), one actually obtains:

$$A = AP_A = \sum_{n,\,r} |nr\rangle a_n \langle nr|. \tag{VII.52}$$

13. Observables in the General Case. Generalized Closure Relation

The Hermitean operator A is an observable if the space $\mathscr{E}_A$ of all the vectors of finite norm, which can be obtained by superposing the eigenvectors of A, coincides with the total Hilbert space $\mathscr{E}$, that is if the operator P_A of projection upon $\mathscr{E}_A$ is equal to unity.

When the spectrum is entirely discrete, P_A can be written in the form of an expansion in a series of elementary, orthogonal projectors formed with the eigenvectors of A. The condition that A be an observable is conveniently expressed by the closure relation (VII.48).

The extension of this relation to the general case necessitates the introduction of differential projectors; they play, in the treatment of the continuous portion of the spectrum, the role which the elementary projectors play in the treatment of the discrete spectrum.

Let us deal first with the case where the spectrum of A possesses no degeneracy. We assume that this spectrum contains a continuous portion labelled by the continuous index ν and a discrete portion labelled by the integral index n. Thus a_n is an eigenvalue of the discrete spectrum, $a(\nu)$ an eigenvalue of the continuous spectrum; $a(\nu)$ is a monotonic function of ν taking on all values of some interval $[a(\nu_1), a(\nu_2)]$. We denote by $|n\rangle$ and $|\nu\rangle$ the eigenkets belonging to the eigenvalues a_n and $a(\nu)$, respectively. These kets are orthonormal; in particular

$$\langle \nu | \nu' \rangle = \delta(\nu - \nu').$$

The operator

$$\delta P = \int_{\nu}^{\nu+\delta\nu} |\nu'\rangle \, \mathrm{d}\nu' \, \langle \nu' |$$

is the projector upon the subspace spanned by the kets $|\nu'\rangle$ of the interval $(\nu, \nu+\delta\nu)$. By adding projectors of this type, one forms the projector

$$P_{\mathrm{c}} = \int_{\nu_1}^{\nu_2} |\nu\rangle \, \mathrm{d}\nu \, \langle \nu |$$

on the subspace $\mathscr{E}_{\mathrm{c}}$ spanned by the eigenkets of the continuous spectrum. This subspace is orthogonal to the subspace $\mathscr{E}_{\mathrm{d}}$ spanned by the eigenkets of the discrete spectrum, a subspace whose projector is

$$P_{\mathrm{d}} = \sum_n |n\rangle \, \langle n|.$$

The condition that A be an observable can be written

$$P_A \equiv P_{\mathrm{c}} + P_{\mathrm{d}} = 1,$$

or else

$$P_A \equiv \sum_n |n\rangle \, \langle n| + \int_{\nu_1}^{\nu_2} |\nu\rangle \, \mathrm{d}\nu \, \langle \nu | = 1. \qquad \text{(VII.53)}$$

The closure relation (VII.53) is the necessary and sufficient condition that the ensemble of the orthonormal vectors $|n\rangle$, $|\nu\rangle$ form a complete set.

The extension to the most general case where all or part of the

spectrum of A is degenerate is quite straightforward. As an illustration let us take up once again the example mentioned at the end of Ch. VII, § 9. The eigenkets of A, $|nr\rangle$, $|\nu\varrho r\rangle$ satisfy the orthonormality relations (VII.30). Furthermore, if A is an observable, in other words, if these kets form a complete set, they satisfy the *closure relation*

$$\boxed{P_A \equiv \sum_{n,r} |nr\rangle \langle nr| + \sum_r \iint |\nu\varrho r\rangle \, d\nu \, d\varrho \, \langle \nu\varrho r| = 1} \qquad \text{(VII.54)}$$

As in the case where the spectrum is entirely discrete, it is very convenient to use the closure relation to form the expansion of any vector $|u\rangle$ of Hilbert space in a series of the basic kets of the observable A. For ease of writing, we suppose that the spectrum of A is non-degenerate [relation (VII.53)]. One obtains

$$|u\rangle = P_A|u\rangle = \sum_n |n\rangle \langle n|u\rangle + \int_{\nu_1}^{\nu_2} |\nu\rangle \, d\nu \, \langle \nu|u\rangle. \qquad \text{(VII.55)}$$

Similarly one finds the generalized Parseval relation

$$\langle u|u\rangle = \langle u|P_A|u\rangle = \sum_n |\langle n|u\rangle|^2 + \int_{\nu_1}^{\nu_2} |\langle \nu|u\rangle|^2 \, d\nu \qquad \text{(VII.56)}$$

and the expansion of A in a series of projectors

$$A = AP_A = \sum_n |n\rangle \, a_n \, \langle n| + \int_{\nu_1}^{\nu_2} |\nu\rangle \, a(\nu) \, d\nu \, \langle \nu|. \qquad \text{(VII.57)}$$

In conclusion, note that it is sometimes convenient to replace the normalization condition of the eigenvectors of the continuous spectrum by the more general condition

$$\langle \nu|\nu'\rangle = f(\nu) \, \delta(\nu - \nu'),$$

where $f(\nu)$ is a real, positive function of ν. This amounts to multiplying each vector $|\nu\rangle$ by a constant of modulus $\sqrt{f}$. In that case, all the above remains valid provided that one replaces in all formulae the differential element

$$|\nu\rangle \, d\nu \, \langle \nu| \quad \text{by} \quad |\nu\rangle \frac{d\nu}{f(\nu)} \langle \nu|.$$

Likewise, if one replaces the normalization condition (VII.30*c*) by

$$\langle \nu\varrho r|\nu'\varrho' r'\rangle = F_r(\nu, \varrho) \, \delta(\nu - \nu') \, \delta(\varrho - \varrho') \, \delta_{rr'},$$

the expression for P_A in the closure relation (VII.54) remains valid provided that one divides the integrand there by the function $F_\tau(\nu, \varrho)$.

14. Functions of an Observable

A linear operator is completely defined by its action upon the vectors of a complete set of orthonormal eigenvectors. Its action on any arbitrary linear superposition of these vectors can be derived immediately therefrom, provided that it is convergent, if one is dealing with an infinite series (the conditions of convergence have already been set forth in Chapter V). In particular, any function $f(a)$ of the eigenvalues of an observable A enables us to define a linear operator $f(A)$, a function of this observable. The action of $f(A)$ on an eigenvector $|a\rangle$ of A belonging to the eigenvalue a is by definition

$$f(A)|a\rangle = f(a)|a\rangle. \tag{VII.58}$$

When the function f is a polynomial, this definition coincides with the one obtained by simply applying the rules of operator algebra; but it remains meaningful in more general cases.

From its very definition it follows that any eigenvector of A is an eigenvector of $f(A)$.

Conversely, if every eigenvector of an observable A is an eigenvector of a linear operator F, the latter is a function of A.

This is rather obvious if none of the eigenvalues of A are degenerate. Let us therefore consider a degenerate eigenvalue, and let $|a1\rangle$, $|a2\rangle$ be two linearly independent eigenvectors belonging to this eigenvalue. By hypothesis, they are eigenvectors of F

$$F|a1\rangle = f_a^{(1)}|a1\rangle \qquad F|a2\rangle = f_a^{(2)}|a2\rangle.$$

Moreover, any linear combination of these two vectors is an eigenvector of F

$$F(\lambda_1|a1\rangle + \lambda_2|a2\rangle) = f_a^{(\lambda)}(\lambda_1|a1\rangle + \lambda_2|a2\rangle);$$

consequently

$$\lambda_1(f_a^{(\lambda)} - f_a^{(1)})|a1\rangle + \lambda_2(f_a^{(\lambda)} - f_a^{(2)})|a2\rangle = 0,$$

and since $|a1\rangle$ and $|a2\rangle$ are linearly independent,

$$f_a^{(1)} = f_a^{(\lambda)} = f_a^{(2)}.$$

Hence all the eigenfunctions of A belonging to the same eigenvalue a are eigenfunctions of F belonging to the same eigenvalue f; the latter is a certain function of a, namely $f(a)$, and one actually has $F=f(A)$.

Any function $f(A)$ of the observable A may be expressed like A itself, in the form of a linear combination of elementary or differential projectors. Suppose, to be definite, that A satisfies equation (VII.57). One likewise has

$$f(A) = f(A)P_A = \sum_n |n\rangle\, f(a_n)\, \langle n| + \int_{\nu_1}^{\nu_2} |\nu\rangle\, f[a(\nu)]\, \mathrm{d}\nu\, \langle \nu|. \qquad \text{(VII.59)}$$

As examples of functions of the observable A, let us mention the projector upon the subspace of a particular eigenvalue, the projector upon the space spanned by the eigenvectors whose eigenvalues are located in a given region. Let us also mention the exponential $\exp(\mathrm{i}\xi A)$ (ξ=given constant), and the inverse A^{-1}. The function $\exp(\mathrm{i}\xi A)$ is always defined; the inverse A^{-1} is well defined only if none of the eigenvalues of A vanish.

15. Operators which Commute with an Observable. Commuting Observables

The functions of the observable A belong to a more general class of operators, *the operators which commute with* A. Such operators have a particularly simple effect upon the eigenvectors of A. Indeed, if $|a\rangle$ is an eigenvector,

$$A|a\rangle = a|a\rangle,$$

and if

$$[A, X] = 0,$$

it follows that

$$0 = (AX - XA)|a\rangle = (A - a)\, X|a\rangle.$$

$X|a\rangle$ is an eigenvector of A belonging to the same eigenvalue (unless it vanishes).

Conversely, in order that an operator X commute with the observable A, it suffices that its action on each of the vectors of a complete orthonormal set of eigenvectors of A yield an eigenvector of A belonging to the same eigenvalue.

Indeed, according to our earlier hypothesis, the action of the

commutator $[A, X]$ on each vector of this complete set gives zero. Therefore, we have $[A, X] = 0$.

All this applies especially to *commuting observables*. Moreover, all the considerations developed in Ch. V, §§ 14 and 15 on the subject of commuting observables may be taken up here once again, except for some details of terminology. We shall merely state the results.

We call *basis* of a given observable any complete orthonormal set of eigenvectors of this observable, having adopted the convention to consider two bases to be identical if their eigenvectors differ only by phase factors, by the order in which the vectors of the discrete spectrum are arranged, or by the choice of continuous indices relating to those of the continuous spectrum.

We have the following important theorem:

If two observables A and B commute, they have at least one basis in common, and, conversely, if two observables A and B possess a common basis, they commute.

Any function $f(a, b)$ of the eigenvalues of two commuting observables, A and B, makes it possible to define a linear operator $f(A, B)$, function of these two observables, by an obvious generalization of the concept of function of an observable. One easily shows that if every eigenvector common to the observables A and B is an eigenvector of a linear operator F, the latter is some function of A and B.

All this can be easily generalized to an arbitrary number of pairwise commuting observables.

Finally, one says that a set $A, B, C, \ldots$ of observables form a *complete set of commuting observables*, if these observables all commute in pairs, and if their common basis is uniquely defined. To each set of eigenvalues $a, b, c, \ldots$ corresponds one and only one common eigenvector (defined to within a constant). This vector may be regarded as a function of the eigenvalues $a, b, c, \ldots$. One commonly denotes it by the symbol $|abc\ldots\rangle$.

III. REPRESENTATION THEORY

16. General Remarks on Finite Matrices

By definition a matrix A of the type $M \times N$ is a set of MN quantities A_{mn} $(m = 1, 2, \ldots, M;\ n = 1, 2, \ldots, N)$ which one usually

arranges in a rectangular array with M rows and N columns

$$(A) \equiv \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1N} \\ A_{21} & \cdots & & \vdots \\ \vdots & & & \\ A_{M1} & \cdots & & A_{MN} \end{pmatrix}$$

A_{mn} is the matrix element of the mth row and the nth column.

If $M=N$, it is called a *square matrix*; the number of its rows and of its columns is the number of dimensions or *order* of that matrix. If one of the two integers M or N is equal to 1, the matrix elements can be likened to the components of a vector. We shall call *column vector* a matrix with one column ($M=$number of dimensions of the vector; $N=1$), and *row vector* a matrix with one row ($M=1$; $N=$ number of dimensions of the vector). A *scalar* is a special matrix for which $M=N=1$.

From a matrix A of the type $M \times N$ one can derive new matrices by carrying out certain *conjugation operations*, namely:

(*i*) the *complex conjugate* matrix A^*, matrix of the type $M \times N$ whose elements are the complex conjugates of the elements of A: $(A^*)_{kl}=A^*_{kl}$;

(*ii*) the *transposed* matrix $\tilde{A}$, matrix of the type $N \times M$, derived from A by interchanging rows and columns:

$$(\tilde{A})_{kl} = A_{lk};$$

(*iii*) the *Hermitean conjugate* matrix $A^\dagger$, matrix of the type $N \times M$ obtained by performing upon A both of the foregoing operations: $(A^\dagger)_{kl}=A^*_{lk}$.

The complex conjugate of a column vector is a column vector. The transposed and Hermitean conjugate matrix of a column vector are row vectors, and *vice versa*. The complex conjugate, transpose, and Hermitean conjugate of a square matrix of order N are square matrices of order N.

One defines the following *algebraic operations* with matrices:

(*a*) *multiplication* of a matrix A *by a constant* c; the product cA is a matrix of the same type as A:

$$(cA)_{mn}=cA_{mn};$$

(*b*) *sum* $S = A + B$ of two matrices of the same type; S is a matrix of the same type as A and B:

$$S_{mn} = A_{mn} + B_{mn};$$

(*c*) *product* (from the left) $P = AB$ of a matrix B of type $M_B \times N_B$ by a matrix A of type $M_A \times N_A$ whose number of columns is equal to the number of rows of B: $N_A = M_B = K$. It is a matrix of type $M_A \times N_B$ whose elements are:

$$P_{mn} = \sum_{k=1}^{K} A_{mk} B_{kn}.$$

The product from the right of B by A (if it exists) is equal to the product from the left of A by B and is written BA.

Clearly

$$(A+B)^* = A^* + B^* \qquad \widetilde{(A+B)} = \tilde{A} + \tilde{B} \qquad (A+B)^\dagger = A^\dagger + B^\dagger$$
$$(AB)^* = A^* B^* \qquad \widetilde{(AB)} = \tilde{B}\tilde{A} \qquad (AB)^\dagger = B^\dagger A^\dagger.$$

Note the reversal of the order of factors in the last two equations.

The product from the left of an N-dimensional column vector by an N-dimensional row vector is a scalar. The product from the left of an N-dimensional row vector by an N-dimensional column vector is a square matrix of order N.

Another important operation is the *tensor product* of two matrices. With a matrix $A^{(1)}$ of type $M_1 \times N_1$ and a matrix $A^{(2)}$ of type $M_2 \times N_2$ one can form, by tensor product, a matrix $A^{(12)} \equiv A^{(1)} \otimes A^{(2)}$ of the type $M_1 M_2 \times N_1 N_2$. The $M_1 M_2$ rows of this matrix are labelled by two indices m_1 and m_2 ($m_1 = 1, 2, \ldots, M_1$; $m_2 = 1, 2, \ldots, M_2$), its $N_1 N_2$ columns by two indices n_1 and n_2 ($n_1 = 1, 2, \ldots, N_1$; $n_2 = 1, 2, \ldots, N_2$):

$$A^{(12)}_{m_1 m_2 ; n_1 n_2} = A^{(1)}_{m_1 n_1} A^{(2)}_{m_2 n_2}.$$

17. Square Matrices

In this paragraph, we shall state a number of definitions and properties peculiar to square matrices.

In a square matrix A of order N, one distinguishes the diagonal

elements A_{nn} ($n = 1, 2, \ldots, N$) from the off-diagonal elements A_{kl} ($k \neq l$). The *trace* of A is the sum of its diagonal elements:

$$\mathrm{Tr}\, A \equiv \sum_n A_{nn}. \tag{VII.60}$$

The *determinant* of A, $\det A$, is the determinant of the array of its elements.

The *unit matrix* I is a matrix whose diagonal elements are all equal to 1 and whose off-diagonal elements all vanish.

$$I_{mn} = \delta_{mn}.$$

The product of the unit matrix by a constant is, by definition, a *constant matrix*. A *diagonal matrix* is a matrix whose off-diagonal elements all vanish.

A square matrix is real, symmetrical or Hermitean according to whether it is equal to its complex conjugate, its transpose or its Hermitean conjugate.

The sum, and the product of two matrices of order N are always defined; they are matrices of order N. The sum is associative and commutative. The product is associative, distributive with respect to the sum, but it is not necessarily commutative. The algebra of matrices of order N is a *non-commutative algebra*.

In order that a matrix of order N commute with all matrices of order N, it is necessary and sufficient that it be constant (Problem VII.4). In particular, the unit matrix I is such that for any A

$$IA = AI = A. \tag{VII.61}$$

Two diagonal matrices necessarily commute. In order that a matrix of order N commute with all diagonal matrices of order N, it is necessary and sufficient that it be diagonal (Problem VII.4).

The *trace* of a product of matrices is *invariant under cyclic permutation* of these matrices

$$\mathrm{Tr}\, ABC = \mathrm{Tr}\, CAB. \tag{VII.62}$$

The determinant of a matrix product is equal to the product of their determinants:

$$\det ABC = \det A \cdot \det B \cdot \det C. \tag{VII.63}$$

A matrix B is by definition the *inverse* of a matrix A if one has

$$AB = 1 \quad \text{and} \quad BA = 1. \tag{VII.64}$$

Indeed, if one of these inequalities is realised, so is the other. One writes:

$$B = A^{-1}.$$

A necessary and sufficient condition for matrix A to possess an inverse is that its determinant be different from zero: $\det A \neq 0$ If the determinant vanishes, the matrix is said to be *singular*.

One easily verifies that

$$(\tilde{A})^{-1} = (\widetilde{A^{-1}}) \qquad (A^*)^{-1} = (A^{-1})^* \qquad (A^\dagger)^{-1} = (A^{-1})^\dagger$$

and that

$$(PQ)^{-1} = Q^{-1}P^{-1}.$$

A matrix O is *orthogonal* if its transpose $\tilde{O}$ is equal to its inverse:

$$O\tilde{O} = \tilde{O}O = I.$$

A matrix U is *unitary* if its Hermitean conjugate equals its inverse:

$$UU^\dagger = U^\dagger U = I.$$

The product from the left of an N-dimensional column vector by a matrix of order N is an N-dimensional column vector. The product from the right of an N-dimensional row vector by a matrix of order N is an N-dimensional row vector.

The action of a diagonal matrix upon such vectors is particularly simple. Let

$$D_{mn} = d_m \, \delta_{mn}$$

be the elements of such a matrix, and u_n the components of a column vector u

$$(Du)_n = d_n \, u_n.$$

Likewise, if v_n are the components of a row vector v

$$(vD)_n = v_n \, d_n.$$

If a matrix is singular, there exists at least one column vector u such that one has $Au = 0$, and conversely.

From this result we obtain the important theorem:

A and B being two matrices of order N, in order that there exist a column vector u (with N dimensions) such that

$$Au = \lambda Bu$$

it is necessary and sufficient that the constant λ be a solution of the equation

$$\det (A - \lambda B) = 0.$$

In particular:

A being a matrix of order N, in order that there exist a column vector u such that

$$Au = \lambda u$$

it is necessary and sufficient that the constant λ be a solution of the equation

$$\det (A - \lambda I) = 0.$$

This algebraic equation, of the Nth degree at most, is called the *secular equation* of A.

Analogous properties hold for the row vectors.

The tensor product of two matrices of order N_1 and N_2, respectively, is a matrix of order $N_1 N_2$. In particular, the tensor product of the unit matrices $I^{(1)}$, $I^{(2)}$, is a unit matrix $I^{(12)}$ of order $N_1 N_2$.

By way of an example, let us mention the fourth order matrices obtained by the tensor product of second-order matrices by second-order matrices. It is common practice to introduce the following two-dimensional matrices (Pauli matrices):

$$1^{(\sigma)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad \text{(VII.65)}$$

All matrices in two dimensions can be put in the form of a linear combination of these four Hermitean matrices. On the other hand, let us consider the matrices of another two-dimensional space

$$1^{(\varrho)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \varrho_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \varrho_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \varrho_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad \text{(VII.66)}$$

Forming the tensor product of a matrix of the type (σ) by a matrix of the type (ϱ) one obtains a matrix with four rows and four

columns. Let us give some explicit examples of matrices of the type $(\varrho\sigma)$ thus formed:

$$\varrho_1 \otimes \sigma_1 = \begin{pmatrix} 0 & \sigma_1 \\ \sigma_1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

$$\varrho_1 \otimes \sigma_2 = \begin{pmatrix} 0 & \sigma_2 \\ \sigma_2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & -\mathrm{i} \\ 0 & 0 & \mathrm{i} & 0 \\ 0 & -\mathrm{i} & 0 & 0 \\ \mathrm{i} & 0 & 0 & 0 \end{pmatrix}$$

$$\varrho_3 \otimes 1^{(\sigma)} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & +1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

One can regard these tensor product matrices as matrices of one of the spaces, the (ϱ)-space for instance, each matrix element of which is a matrix of the other space. This was done in the center of each of the equations above. The right-hand sides give the explicit expressions of the matrices; if one agrees to label the rows (and the columns) by two indices $m_\varrho m_\sigma$, the first of which refers to the components of (ϱ)-space, and the second to those of (σ)-space, the rows (and columns) will be arranged in the order: 11, 12, 21, 22.

By linear combination of tensor product matrices, one forms square matrices with two indices

$$A_{m_1 m_2; n_1 n_2} \qquad (m_1, n_1 = 1, 2, \ldots, N_1;\ m_2, n_2 = 1, 2, \ldots, N_2)$$

and with $N_1 N_2$ dimensions. As is shown by the foregoing example, they may be regarded as matrices of type (1) whose elements are matrices of type (2). Forming the sum of the diagonal elements of such a matrix, one obtains a matrix of type (2) in the ordinary sense of the term. It is by definition the partial trace in space (1) of the matrix from which we started out:

$$(\mathrm{Tr}_1 A)_{m_2 n_2} \equiv \sum_{n=1}^{N_1} A_{n\, m_2;\, n\, n_2}. \tag{VII.67}$$

Likewise one defines the partial trace in space (2). It is obvious that

$$\mathrm{Tr}\, A = \mathrm{Tr}_2 (\mathrm{Tr}_1 A) = \mathrm{Tr}_1 (\mathrm{Tr}_2 A) \tag{VII.68}$$

and that, if A is equal to the tensor product $A^{(1)} \otimes A^{(2)}$,

$$\mathrm{Tr}\, (A^{(1)} \otimes A^{(2)}) = (\mathrm{Tr}_1 A^{(1)})\, (\mathrm{Tr}_2 A^{(2)}). \tag{VII.69}$$

18. Extension to Infinite Matrices

Most of the remarks on finite matrices may be extended to infinite matrices, whose rows and columns are labelled by one or several discrete indices or, still more generally, by a number of indices which can take on a finite or denumerably infinite number of discrete values, and a number of continuous indices which can assume all values located in a given interval. An infinite matrix is square if its rows and columns are labelled by the same system of indices. It is a column vector if it has but one column, and a row vector it if has but one row.

The operations of complex conjugation, transposition, and Hermitean conjugation extend to infinite matrices without change. The same is true for multiplication by a constant, and for the sum. As far as the product (from the left) of B by A is concerned, it is understood that the rows of B and the columns of A must be labelled by the same system of indices. Moreover, if some indices are continuous, the summation must be replaced by an integration. Let us suppose, to be definite, that B and A are square matrices depending upon a continuous index q which can take any value in the interval (q_1, q_2). The matrix element $P(q; q')$ of the product $P = AB$ is

$$P(q; q') = \int_{q_1}^{q_2} A(q; q'')\, B(q''; q')\, \mathrm{d}q''.$$

The product is well defined only if the summations or integrals which occur in the definition of its elements converge.

Aside from any questions of convergence, the considerations of § 17 on square matrices may be extended to matrices of infinite order, with the exception of the notion of determinant. The only points to be made concern the definition of diagonal matrices in the case where the indices are continuous, and the conditions concerning the existence of an inverse matrix.

By definition, a continuous matrix $D(q; q')$ is *diagonal* if it is of the form

$$D(q; q') = d(q)\, \delta(q - q') \qquad \text{(VII.70)}$$

where $d(q)$ is an arbitrary function of the index q. In this way the two characteristic properties of diagonal matrices are preserved, namely that two diagonal matrices commute, and that the action of a diagonal

matrix upon a vector consists in multiplying each of its components by the corresponding diagonal element. Thus, the action of the diagonal matrix defined by expression (VII.70) upon the column vector g with components $g(q)$ yields the vector $h = Dg$ with components

$$h(q) = \int D(q, q')\, g(q')\, \mathrm{d}q' = d(q)\, g(q).$$

Note that the continuous matrix $\delta'(q-q')$ *is not diagonal.*

As far as *the existence of the inverse* of a given matrix is concerned, in contrast to the case of finite matrices, the fact that

$$AB - I \tag{VII.71a}$$

does not necessarily imply that

$$BA = I. \tag{VII.71b}$$

The two equations (VII.71a) and (VII.71b) must be satisfied simultaneously in order that one may assert that A and B are the inverse of each other.

Actually, it is not necessary that A be a square matrix for it to possess an inverse; for instance, it may happen that a matrix A whose rows are labelled by a discrete index and whose columns are labelled by a continuous index, possesses an inverse. In that case, the inverse A^{-1} has a discrete column index and a continuous row index. In particular, in a *unitary matrix* U, a matrix satisfying by definition the two equations

$$UU^{\dagger} = 1, \qquad U^{\dagger}U = 1 \tag{VII.72}$$

the row indices and the column indices are not necessarily of the same kind. However, the unit matrices of the right-hand sides in the two defining equations (VII.72) are necessarily square matrices. If the matrix U is not a square matrix, the systems of indices of each of these unit matrices are different.

19. Representation of Vectors and Operators by Matrices

Consider a vector space $\mathscr{E}$ and choose a basis in this space. The latter may possibly be the eigenvectors of a complete set of commuting observables. To simplify the writing, we shall make our arguments on a basis whose vectors are labelled by a discrete index n. We assume,

for instance, that they are the eigenvectors of some observable Q

$$Q|n\rangle = q_n|n\rangle.$$

We shall say that they are the basis vectors of the representation $\{Q\}$. These vectors form a complete orthonormal set

$$\langle m|n\rangle = \delta_{mn} \tag{VII.73}$$

$$P_Q \equiv \sum_n |n\rangle\langle n| = 1. \tag{VII.74}$$

Equations (VII.73) and (VII.74) are the fundamental equations of representation $\{Q\}$.

For any ket vector $|u\rangle$,

$$|u\rangle = P_Q|u\rangle = \sum_n |n\rangle\langle n|u\rangle.$$

The quantities $u_n = \langle n|u\rangle$ can be regarded as the elements of a matrix with one column whose row index is n. This column vector completely defines $|u\rangle$: it is the matrix representing $|u\rangle$ in representation $\{Q\}$.

For any bra vector $\langle v|$,

$$\langle v| = \langle v|P_Q = \sum_n \langle v|n\rangle\langle n|.$$

The quantities $\langle v|n\rangle$ are the complex conjugates of the components v_n of the column vector representing the ket $|v\rangle$ in the representation $\{Q\}$. They may be regarded as the elements of a row vector; this row vector defines $\langle v|$ completely; it is the vector representing $\langle v|$ in representation $\{Q\}$. With this convention, the bra conjugate to a given ket is represented by the Hermitean conjugate of the vector representing that ket.

Any linear operator A can be expanded in a unique manner in a double series of operators of basis $|m\rangle\langle n|$

$$A = P_Q A P_Q = \sum_{mn} |m\rangle\langle m|A|n\rangle\langle n|.$$

The coefficients of the expansion $A_{mn} = \langle m|A|n\rangle$ completely define A and can be regarded as the elements of a square matrix whose row index is m and whose column index is n: it is the matrix representing A in representation $\{Q\}$.

Having established a one-to-one correspondence between vectors and operators on the one hand, and matrices on the other hand, we

shall examine how each operation concerning the operators and vectors of the space $\mathscr{E}$ is translated into the language of the matrices that represent them.

To the conjugation relations between vectors or between operators correspond relations of Hermitean conjugation between matrices. We have already noted this fact in connection with the conjugation between bras and kets. Likewise, the matrices representing two Hermitean-conjugate operators A and $A^\dagger$ are Hermitean conjugate; their elements actually satisfy the relations characteristic of Hermitean conjugation:

$$A^\dagger_{mn} \equiv \langle m|A^\dagger|n\rangle = \langle n|A|m\rangle^* = A^*_{nm}.$$

As for the different algebraic operations between vectors and operators, they are translated into matrix language *by the same algebraic operations upon their representative matrices.* One can convince oneself by merely examining each of the elementary operations defined in the two preceding sections.

This is evident for the multiplication by a constant and the sum; thus to any linear combination $\lambda_1 A_1 + \lambda_2 A_2$ of two operators corresponds the same linear combination of their representative matrices

$$\langle m|\ (\lambda_1 A_1 + \lambda_2 A_2)\ |n\rangle = \lambda_1 \langle m|A_1|n\rangle + \lambda_2 \langle m|A_2|n\rangle.$$

Likewise the different products defined for vectors and operators are represented by the corresponding matrix products, namely

(*i*) scalar product of $|u\rangle$ by $|v\rangle$:

$$\langle v|u\rangle = \langle v|P_Q|u\rangle = \sum_n \langle v|n\rangle\langle n|u\rangle = \sum_n v_n^* u_n;$$

$\langle v|u\rangle$ is equal to the product (from the left) of the matrix (column vector) representing $|u\rangle$ by the Hermitean conjugate of the matrix representing $|v\rangle$;

(*ii*) action of an operator A on a ket $|u\rangle$ or on a bra $\langle v|$

$$\langle n|A|u\rangle = \langle n|AP_Q|u\rangle = \sum_k \langle n|A|k\rangle\ \langle k|u\rangle$$

$$\langle v|A|n\rangle = \langle v|P_Q A|n\rangle = \sum_l \langle v|l\rangle\ \langle l|A|n\rangle.$$

The matrix (column vector) representing $A|u\rangle$ is the product (from the left) of the matrix representing $|u\rangle$ by the matrix representing A.

The matrix (row vector) representing $\langle v|A|$ is the product (from the right) of the matrix representing $\langle v|$ by the matrix representing A;

(*iii*) the product AB of the operator B by the operator A:

$$\langle m|AB|n\rangle = \langle m|AP_Q B|n\rangle = \sum_k \langle m|A|k\rangle \langle k|B|n\rangle$$

the matrix representing AB is the product (from the left) of the matrix representing B by the matrix representing A;

(*iv*) operator $|u\rangle\langle v|$; the element (m, n) of its representative matrix is $\langle m|u\rangle \langle v|n\rangle$; the latter therefore results from the product (from the left) of the matrix (row vector) representing $\langle v|$ by the matrix (column vector) representing $|u\rangle$ (which actually yields a square matrix).

To sum up, we have succeeded in defining a representation of vectors and operators of $\mathscr{E}$ space by matrices, in such a way that there exist very simple correspondence rules between the various operations on the vectors and operators, and the operations on matrices. Any geometrical problem in $\mathscr{E}$ space may be treated either by the methods of pure geometry by directly manipulating the vectors and operators which occur there, or by manipulations of algebra or analysis on the matrices representing them in a suitable representation.

In this last case, a more or less judicious choice of representation will lead to a more or less simple solution of the problem, in the same way as a more or less happy choice of coordinates makes the solution of a problem of analytical geometry more or less simple. In practice one will choose the representation in which the vectors and operators under study are represented by the simplest possible matrices.

Note in this connection that the observable Q is represented by a particularly simple matrix in representation $\{Q\}$: it is a diagonal matrix. More generally, any function $f(Q)$ is represented by a diagonal matrix,

$$\langle m|f(Q)|n\rangle = f(q_n)\,\delta_{mn}.$$

Operators which commute with Q are also represented by very simple matrices. In fact, if $[X, Q] = 0$,

$$(q_n - q_m)\,\langle m|X|n\rangle = 0$$

and consequently $\langle m|X|n\rangle = 0$ for any pair (m, n) such that $q_m \neq q_n$. In other words, all the matrix elements whose row and column indices refer to distinct eigenvalues of Q, are necessarily zero (cf. Ch. VII, § 15).

All the foregoing considerations can be extended without difficulty to the entire space $\mathscr{E}_1 \otimes \mathscr{E}_2$ resulting from the tensor product of the two spaces $\mathscr{E}_1$ and $\mathscr{E}_2$. The vectors and operators formed by tensor product can be represented by the tensor product matrices of the matrices representing the vectors and operators of the spaces $\mathscr{E}_1$ and $\mathscr{E}_2$.

20. Matrix Transformations

Consider once again matrices of finite order. In what follows, we shall designate by a capital letter square matrices of order N, and by a small letter N-dimensional column or row vectors. Let T be a non-singular matrix (T^{-1} exists). This matrix enables us to define a transformation in which the matrix A', the transform of A, is defined by

$$A' = TAT^{-1}. \tag{VII.75}$$

The correspondence between A and A' is one-to-one, A being deduced from A' by the inverse transformation

$$A = T^{-1}A'T. \tag{VII.76}$$

Such a transformation conserves the trace and the determinant

$$\operatorname{Tr} A = \operatorname{Tr} A', \qquad \det A = \det A' \tag{VII.77}$$

(properties of the trace and the determinant of a matrix product). It is likewise clear that it conserves any algebraic equation between matrices. If, for instance,

$$A = \lambda BC + \mu DEF,$$

then, multiplying term by term by T from the left and by T^{-1} from the right, and inserting into each monomial the expression $T^{-1}T$ as many times as necessary, we have

$$TAT^{-1} = \lambda TBT^{-1}\, TCT^{-1} + \mu TDT^{-1}\, TET^{-1}\, TFT^{-1},$$

that is

$$A' = \lambda B'C' + \mu D'E'F'.$$

One likewise defines the transform u' of a column vector u:

$$u' = Tu, \qquad u = T^{-1}u' \tag{VII.78}$$

and the transform v' of a row vector v:

$$v' = vT^{-1}, \qquad v = v'T. \tag{VII.79}$$

As may be easily verified, *the transformation conserves in all generality the algebraic equations* involving square matrices and vectors of one or the other type. On the other hand, if c is an arbitrary constant, the transforms of square matrices do not change upon replacing T by cT; however, the column vectors are multiplied by c, and the row vectors by $1/c$.

A transformation of the kind just described does not in general conserve the various conjugation relations between matrices (Problem VII.5). In particular, let us look for the condition which T must satisfy in order that the transformation conserve the Hermitean conjugation. In order that $A' = TAT^{-1}$ imply

$$A'^{\dagger} = TA^{\dagger}T^{-1}$$

for any A, one must have

$$TAT^{-1} = (TA^{\dagger}T^{-1})^{\dagger} = (T^{-1})^{\dagger}AT^{\dagger},$$

or else, multiplying term by term from the left by $T^{\dagger}$ and from the right by T,

$$T^{\dagger}TA = AT^{\dagger}T.$$

It is therefore necessary that $T^{\dagger}T$ commute with all matrices A, and thus that $T^{\dagger}T$ be a multiple of unity:

$$T^{\dagger}T = cI.$$

Moreover, in order that $u' = Tu$ imply $u'^{\dagger} = u^{\dagger}T^{-1}$ for any u, one must have $u = T^{\dagger}Tu$ for any u, hence that $c = 1$. In conclusion, the matrix T must be *unitary*. It is clear that this condition which is necessary for the Hermitean conjugation relations to be conserved, is also sufficient.

One calls *unitary transformation* a transformation whose matrix U

is unitary. Since, in that case, $U^{-1} = U^\dagger$, the definitions of the transforms of a matrix A, of a column vector u, and of a row vector v are, respectively,

$$\begin{aligned} A' &= UAU^\dagger & A &= U^\dagger A' U \\ u' &= Uu & u &= U^\dagger u' \\ v' &= vU^\dagger & v &= v'U. \end{aligned} \tag{VII.80}$$

As all transformations, *a unitary transformation conserves the trace and the determinant of the matrices* and *the algebraic equations between matrices and vectors. In addition, it conserves the Hermitean conjugation relations.*

Furthermore, one has the two following fundamental theorems which we state here without proof.

A) Any Hermitean matrix H may be diagonalized by a unitary transformation

$$H' = UHU^\dagger, \qquad H' \text{ diagonal.}$$

The diagonal elements of H' are the "eigenvalues" of H. They are all real (H' is Hermitean) and are solutions of the secular equation

$$\det (H - xI) = 0.$$

B) In order that two Hermitean matrices H, K might be diagonalized by one and the same unitary transformation, it is necessary and sufficient that they commute.

All definitions and properties relating to matrices of finite order are extended without difficulty to infinite matrices. Any infinite matrix T possessing an inverse defines a transformation of square matrices or (column or row) vectors, provided that the sums or integrals which enter in the definitions converge. In contrast to the transformations of finite matrices it is not necessary that T be a square matrix. Of course the rows and columns of the (square) matrices which one transforms are labelled by the same system of indices as the columns of T; the same holds for the components of the column and row vectors. As for the rows and columns of the transformed matrices and the components of the transformed vectors, their system of indices is that of the rows of T.

The properties of conservation of the trace (with the proviso that it converges), of the algebraic equations and, in the case of unitary

transformations, of Hermitean conjugation persist in the transformations of infinite matrices. However, the two fundamental theorems concerning the diagonalization of Hermitean matrices by a unitary transformation do not hold for all Hermitean matrices; we shall assume that they apply to all those we shall encounter.

21. Change of Representation

Let us return to the problem of the representation of vectors and operators of a vector space $\mathscr{E}$ by matrices. To each basis in this space corresponds a given representation. It is convenient to know how to relate the matrices representing one and the same operator or vector in each of the representations one can thus form. We shall see that one passes from one to the other by *unitary transformation.*

Consider two bases, one formed with the eigenvectors $|n\rangle$ ($n = 1, 2, \ldots, \infty$) of the observable Q of § 19, the other with the eigenvectors $|\xi\rangle$ of another observable Ξ whose spectrum we suppose to be continuous. These two bases define the representations $\{Q\}$ and $\{\Xi\}$. The fundamental equations of the representation $\{Q\}$ have already been written down [eqs. (VII.73) and (VII.74)]. Those of the representation $\{\Xi\}$ are

$$\langle\xi|\xi'\rangle = \delta(\xi - \xi') \tag{VII.81}$$

$$P_{\Xi} \equiv \int |\xi\rangle \, \mathrm{d}\xi \, \langle\xi| = 1. \tag{VII.82}$$

The basis vectors of each of the representations can be expanded in a series of basis vectors of the other:

$$|n\rangle = \int |\xi\rangle \, \mathrm{d}\xi \, \langle\xi|n\rangle, \qquad |\xi\rangle = \sum_n |n\rangle \, \langle n|\xi\rangle. \tag{VII.83}$$

The scalar product $\langle\xi|n\rangle$ occurring as coefficient in the expansion of $|n\rangle$ may be regarded as the element $S(\xi; n)$ of a matrix S whose row index is ξ and column index is n. The scalar product $\langle n|\xi\rangle$ occurring as coefficient in the expansion of $|\xi\rangle$ may be regarded as the element $T(n; \xi)$ of a matrix T whose row index is n and column index is ξ. Moreover, since $\langle\xi|n\rangle = \langle n|\xi\rangle^*$,

$$T = S^\dagger.$$

Furthermore,

$$\sum_n \langle\xi|n\rangle \, \langle n|\xi'\rangle = \langle\xi|\xi'\rangle = \delta(\xi - \xi'),$$

$$\int \langle n|\xi\rangle \, \mathrm{d}\xi \, \langle\xi|n'\rangle = \langle n|n'\rangle = \delta_{nn'}.$$

In other words

$$SS^{\dagger} = I \tag{VII.84a}$$

$$TT^{\dagger} \equiv S^{\dagger}S = I. \tag{VII.84b}$$

The matrix S is unitary.

Now designate by $(u)_Q$ the column vector with components $\langle 1|u\rangle$, $\langle 2|u\rangle$, ... representing the ket $|u\rangle$ in representation $\{Q\}$, and by $(u)_\Xi$ the column vector with components $\langle \xi|u\rangle$ representing this same ket in representation $\{\Xi\}$. By application of relation (VII.74),

$$\langle \xi|u\rangle = \sum_k \langle \xi|k\rangle \langle k|u\rangle,$$

in other words,

$$(u)_\Xi = S(u)_Q. \tag{VII.85}$$

Likewise designate by $(A)_Q$ and by $(A)_\Xi$ the matrices representing a given operator A in the representations $\{Q\}$ and $\{\Xi\}$, respectively. One has

$$\langle \xi|A|\xi'\rangle = \sum_{kl} \langle \xi|k\rangle \langle k|A|l\rangle \langle l|\xi'\rangle$$

or else

$$(A)_\Xi = S(A)_Q S^{\dagger}. \tag{VII.86}$$

One would likewise obtain between the row vectors $(v)_Q$ and $(v)_\Xi$ representing the same bra $\langle v|$:

$$(v)_\Xi = (v)_Q S^{\dagger}. \tag{VII.87}$$

Eqs. (VII.85), (VII.86), and (VII.87) are the characteristic transformation equations of the unitary transformation S [eq. (VII.80)].

The elements of this matrix have the following remarkable properties:

– considered as functions of the column index n, the elements $\langle \xi|n\rangle$ of the ξth row are the components of the row vector $(\xi)_Q$ representing the eigenbra $\langle \xi|$ of Ξ in representation $\{Q\}$;

– considered as functions of the row index ξ, the elements $\langle \xi|n\rangle$ of the nth column are the components of the column vector $(n)_\Xi$ representing the eigenket $|n\rangle$ of Q in representation $\{\Xi\}$.

In particular, the solution in representation $\{Q\}$ of the eigenvalue problem of the operator Ξ is a problem mathematically equivalent to the determination of the transformation S which diagonalizes the

matrix $(\Xi)_Q$. Likewise, the solution in the representation $\{\Xi\}$ of the eigenvalue problem of the operator Q is equivalent to the problem of determining the transformation S which diagonalizes the matrix $(Q)_{\Xi}$.

It is important to recognize the quantities and relations which can be defined independently of any representation. All quantities and all relations defined directly by means of vectors and operators obviously possess this property. Thus the *scalar product* of two vectors is invariant under a change of representation. The *Hermitean conjugation relations*, and the *algebraic equations* between vectors and between operators likewise possess this invariance property.

Let us also mention the *conservation of the trace*: the trace (if it converges) of the matrix representing an operator retains the same value no matter what representation is used; it is a quantity characteristic of the operator itself. In particular it is easy to show that (Problem VII.6)

$$\mathrm{Tr}\ |u\rangle\langle u| = \langle u|u\rangle \tag{VII.88}$$

$$\mathrm{Tr}\ |u\rangle\langle v| = \langle v|u\rangle. \tag{VII.89}$$

22. Unitary Transformations of Operators and Vectors

The matrix S of the preceding paragraph does not represent an operator. The matrix representative of an operator is defined in a given representation whereas the transformation matrix straddles, so to speak, two representations. In the particular example examined earlier this is revealed especially since the matrix S is not a square matrix.

However, and here we have an important special case, it may happen that a one-to-one correspondence exists between the basis vectors of the first representation and those of the second. In that case the vectors of the two bases are labelled by the same set of indices. Let us consider, to be definite, a representation $\{Q\}$ whose basis vectors $|n\rangle$ are labelled by a discrete index n, and a representation $\{\bar{Q}\}$ whose basis vectors $|\bar{n}\rangle$ are labelled by the same index. Two kets $|n\rangle$, $|\bar{n}\rangle$ labelled by the same index correspond to each other. Let U be the linear operator defined by this correspondence:

$$|n\rangle = U|\bar{n}\rangle.$$

One has

$$U = U(\sum_n |\bar{n}\rangle\langle\bar{n}|) = \sum_n |n\rangle\langle\bar{n}| \tag{VII.90}$$

and

$$U^\dagger = \sum_n |\bar{n}\rangle \langle n|$$

taking into account the orthonormality relations of the $|n\rangle$ and the $|\bar{n}\rangle$,

$$UU^\dagger = U^\dagger U = 1. \tag{VII.91}$$

Hence, U is a *unitary operator*. In fact, the unitary matrix $\langle \bar{m}|n\rangle$ defining the change of representation from $\{Q\}$ to $\{\bar{Q}\}$ is the matrix representing U in representation $\{\bar{Q}\}$.

In the case where one can form a unitary operator U, one can define a manipulation which is to some extent complementary to the change of representation. Instead of transforming the basis $\{Q\}$ into a new basis $\{\bar{Q}\}$ whose vectors are given by the equation

$$|\bar{n}\rangle = U^\dagger |n\rangle \tag{VII.92}$$

one can carry out the transformation on the vectors and the operators of the space $\mathscr{E}$ themselves, and associate with each vector $|u\rangle$, the vector $|\hat{u}\rangle = U|u\rangle$, and with each operator A, the operator $\hat{A} = UAU^\dagger$.

Taking into account the fact that U is unitary, it is clear that the transformation U conserves the conjugation relations and the equations between vectors and operators. In particular,

(*i*) the scalar product is conserved: $\langle \hat{u}|\hat{A}|\hat{v}\rangle = \langle u|A|v\rangle$;

(*ii*) the hermiticity is conserved.

In fact, if A is an observable, its transform $\hat{A}$ is an observable possessing the same eigenvalue spectrum, since the eigenvalue equation of A

$$A|a\rangle = a|a\rangle$$

transforms into the equation

$$\hat{A}|\hat{a}\rangle = a|\hat{a}\rangle. \tag{VII.93}$$

The eigenkets of $\hat{A}$ corresponding to a given eigenvalue a are the transforms of the eigenkets of A corresponding to the same eigenvalue. Note that the matrix representing $\hat{A}$ in representation $\{Q\}$ is the same as that which represents A in representation $\{\bar{Q}\}$. Likewise the vector $|\hat{a}\rangle$ has the same components in $\{Q\}$ as $|a\rangle$ has in $\{\bar{Q}\}$.

To carry out successively two transformations defined, respectively, by the operators U and V is equivalent to the transformation defined by the operator $W = VU$. Since W is unitary, the resulting transformation is unitary. In other words, *the product of two unitary transformations is a unitary transformation.*

If the operator U defining a unitary transformation is infinitely close to 1, the transformation is said to be *infinitesimal*. U takes the form

$$U \equiv 1 + i\varepsilon F \tag{VII.94}$$

where ε is a real, infinitesimal quantity. The unitarity condition (VII.91) reads

$$(1 - i\varepsilon F^\dagger)(1 + i\varepsilon F) = (1 + i\varepsilon F)(1 - i\varepsilon F^\dagger) = 1,$$

or else, retaining only terms of the first order in ε,

$$F = F^\dagger.$$

The operator F is *Hermitean*.

In such an infinitesimal transformation, the transformed vectors and operators are given by the expressions

$$|\hat{u}\rangle \equiv |u\rangle + \delta|u\rangle = (1 + i\varepsilon F)\,|u\rangle$$

$$\hat{A} \equiv A + \delta A = (1 + i\varepsilon F)\,A(1 - i\varepsilon F) = A + i\varepsilon[F, A]$$

or

$$\delta|u\rangle = i\varepsilon F|u\rangle \tag{VII.95}$$

$$\delta A = i\varepsilon[F, A]. \tag{VII.96}$$

EXERCISES AND PROBLEMS

1. By definition, a projector P_i is less than or equal to another projector P_j, if $P_iP_j = P_i$; one then uses the notation $P_i \leqslant P_j$. Show that if $P_i \leqslant P_j$, one necessarily has $\langle u|P_i|u\rangle \leqslant \langle u|P_j|u\rangle$ for any $|u\rangle$, and conversely. Show either directly or by making use of this last property, that the inequality thus defined actually satisfies the characteristic axioms of an inequality, namely that (*i*) $P_i \leqslant P_j$ and $P_j \leqslant P_i$ imply $P_i = P_j$; (*ii*) $P_i \leqslant P_j$ and $P_j \leqslant P_k$ imply $P_i \leqslant P_k$.

2. $P_1, P_2, \ldots, P_K$ being projectors, show that their sum is likewise a projector if, and only if

$$\sum_{i=1}^{K} \langle u|P_i|u\rangle \leqslant \langle u|u\rangle$$

for any vector $|u\rangle$ of Hilbert space.

3. (*i*) An observable A possesses a finite number N of eigenvalues. One denotes them by $a_1, a_2, \ldots, a_N$ and sets

$$f(A) \equiv (A-a_1)(A-a_2)\ldots(A-a_N)$$
$$\equiv (A-a_n)\, g_n(A).$$

Show that $f(A) = 0$ and that the projector P_n upon the subspace of the nth eigenvalue is given by the expression

$$P_n = g_n(A)/g_n(a_n).$$

(*ii*) Derive the converse property, namely:

If A is a Hermitean operator satisfying the algebraic equation of the Nth degree

$$f(A) \equiv (A-a_1)(A-a_2)\ldots(A-a_N) = 0$$

and if it satisfies no other algebraic equation of degree less than N, it is an observable possessing N eigenvalues, and these are the N necessarily real and distinct roots of the equation $f(x) = 0$.

4. Show that a matrix of order N

(*i*) is necessarily a constant if it commutes with all matrices of order N;

(*ii*) is necessarily diagonal if it commutes with all diagonal matrices of order N.

5. Show that: (*a*) in order that a transformation conserve complex conjugation between matrices, it is necessary and sufficient that the transformation matrix be real; (*b*) in order that a transformation conserve the transposition relation between matrices, it is necessary and sufficient that the transformation matrix be orthogonal.

6. Let $|u\rangle$ and $|v\rangle$ be two vectors of finite norm. Show that

$$\mathrm{Tr}\ |u\rangle\langle u| = \langle u|u\rangle$$
$$\mathrm{Tr}\ |u\rangle\langle v| = \langle v|u\rangle.$$

7. Let H be a positive definite, Hermitean operator. Show that for any $|u\rangle$ and $|v\rangle$,

$$|\langle u|H|v\rangle|^2 \leqslant \langle u|H|u\rangle\langle v|H|v\rangle$$

and that the equality $\langle u|H|u\rangle = 0$ necessarily implies $H|u\rangle = 0$. Show also that $\mathrm{Tr}\, H \geqslant 0$ and that the equality implies $H = 0$.

8. Show that if H and K are two positive definite observables, $\mathrm{Tr}\, HK > 0$ and that the equality implies $HK = 0$.

9. A being a linear operator, show that $A^\dagger A$ is a positive definite Hermitean operator and that its trace is equal to the sum of the squares of the moduli of the matrix elements representing A in an arbitrarily chosen representation. Deduce that $\mathrm{Tr}\, A^\dagger A > 0$, and that the equality $\mathrm{Tr}\, A^\dagger A = 0$ implies $A = 0$.

GENERAL FORMALISM

(B) DESCRIPTION OF PHYSICAL PHENOMENA

1. Introduction

In Classical Physics, the dynamical state of a given system is defined at every instant, once the values assumed by the assembly of quantities or dynamical variables associated with the system are known at that instant. The latter may, in principle, all be determined simultaneously with infinite precision. The object of Classical Theory is to enumerate these dynamical variables, and then to discover and to study their equations of motion.

In the Quantum Theory, the relationship between dynamical states and dynamical variables is much less direct. In the measurement process of a given dynamical variable, the dynamical state of the system on which the measurement is performed, is in general modified by the intervention of the measuring device. This modification, which is usually neglected in Classical Physics, ceases to be negligible on the microscopic scale; it appears as an unpredictable and uncontrollable disturbance of the system and sets a limit to the precision with which the dynamical variables can all be measured simultaneously. One therefore abandons the fundamental postulate of Classical Physics, according to which all the various quantities belonging to a system take on well-defined values at each instant of time. One can only determine for each of these variables a *statistical distribution* of values, which is the probability law of the results of measurement in the eventuality that such a measurement is performed.

According to the usual terminology (Ch. VI, § 17) all the dynamical variables of a quantum system are not compatible with each other. One assumes, however, that one can always add to every dynamical variable of the system a certain number of others and thus form a *complete set of compatible variables*; by definition, all the variables of such a set are compatible with each other and there exist no other variables compatible with each of them, aside from functions of these variables themselves. The precise determination of the variables of a complete set constitutes the largest possible amount of information one can obtain on the dynamical state of a quantum system. Conse-

quently, the dynamical state of a quantum system is no longer defined, as in Classical Theory, by the precise specification of all the dynamical variables associated with the system, but rather by the specification of those which occur in one of the various complete sets of compatible variables one can construct.

One establishes the principle that the dynamical states of a quantum system are *linearly superposable*. In accordance with this principle (Ch. VII, § 1), one can associate with such a system a certain vector space $\mathscr{E}$ such that each dynamical state of the system is represented by a vector of this space. One further supposes that $\mathscr{E}$ is a Hilbert space. Hereafter, we shall use the notation and properties of Hilbert spaces as they were stated in Chapter VII. Hence, *to each dynamical state corresponds a certain ket* $|u\rangle$ *of space* $\mathscr{E}$ [1]). Likewise, *to each dynamical variable is attached an observable of space* $\mathscr{E}$. According to whether wo variables are or are not compatible, the observables corresponding to them do or do not commute.

The general formalism of thc Quantum Theory is based upon this correspondence between dynamical states and physical quantities on the one hand, and between vectors and operators on the other. In Section I we shall define this correspondence in precise fashion, and we shall indicate a practical way to construct the Hilbert space $\mathscr{E}$, and the physical significance to be attached to its vectors and operators. In Section II, this general theoretical scheme is complemented by giving the *equations of motion*. In Section III, we show that there exist as many particular formulations of the theory as there are particular matrix representations of vectors and operators in $\mathscr{E}$ space. Wave Mechanics is one of these particular formulations. When the dynamical state of a quantum system is incompletely known, one can represent it by a statistical mixture of vectors, following the usual methods of statistics. An equivalent procedure consists in representing this state by an operator of a special type, the *density operator*; the formalism of the density operator is outlined in the fourth and last section of this chapter.

---

[1]) The space $\mathscr{E}$ plays a similar role in Quantum Theory to that of phase space in Classical Theory. Each point of phase space represents a classical dynamical state; likewise each vector of $\mathscr{E}$ space represents a dynamical quantum state. In this latter case, however, it is not a one-to-one correspondence since two vectors of $\mathscr{E}$ space which are multiples of each other represent the same state; cf. § 2 below.

I. DYNAMICAL STATES AND PHYSICAL QUANTITIES

2. Definition of Probabilities. Postulates Concerning Measurement

In a given dynamical state, a definite statistical distribution of values is associated with each dynamical variable of the system. In order to define these probabilities, one starts out from the following fundamental postulate:

The mean value of any function $F(A)$ of a given physical quantity A is

$$\boxed{\langle F(A)\rangle = \langle u|F(A)|u\rangle,} \qquad \text{(VIII.1)}$$

an expression in which the ket $|u\rangle$ represents the dynamical state of the system, and the observable A represents the physical quantity in question.

In particular, the characteristic function $f(\xi)$ of the statistical distribution of A is the mean value of the function $\exp(i\xi A)$:

$$f(\xi) = \langle u|\, e^{i\xi A}\, |u\rangle. \qquad \text{(VIII.2)}$$

Since a statistical distribution is completely determined upon specifying its characteristic function, the fundamental postulate we have just stated completely defines the statistical distributions of all the dynamical variables of the system.

Let us see how this postulate affects the correspondence between dynamical states and ket vectors. Whatever the operator $F(A)$, expression (VIII.1) remains unchanged when multiplying vector $|u\rangle$ by an arbitrary phase factor $\exp(i\alpha)$ (α real and arbitrary). Hence the statistical distributions calculated for two vectors which differ only by a phase factor, are strictly identical: two such vectors represent the same dynamical state. In other words, *to each dynamical state corresponds a vector defined to within a phase factor*. On the other hand, one necessarily has $f(0) = 1$ (the average value of 1 is equal to 1); the vector $|u\rangle$ must therefore be *normalized to unity*

$$\langle u|u\rangle = 1. \qquad \text{(VIII.3)}$$

It is sometimes convenient to relax this last condition. To this end, one replaces the definition (VIII.1) of the mean values by the more general definition

$$\langle F(A)\rangle = \frac{\langle u|F(A)|u\rangle}{\langle u|u\rangle}. \qquad \text{(VIII.4)}$$

With this definition, two vectors which are multiples of each other represent the same dynamical state (it being understood that the vectors with which we are dealing have finite norm).

To obtain the statistical distribution of A explicitly, one calculates the expression (VIII.2) of the characteristic function $f(\xi)$ [or the expression (VIII.4) if $|u\rangle$ is not normalized to unity] in a representation where A is diagonal. Except for slight differences in terminology, this method has already been outlined in Chapter V. It will not be restated here. We shall merely state the results:

1) *The only precise values which the quantity A may assume are those of the eigenvalue spectrum of the observable associated with A.*

2) *Let $\mathscr{E}_D$ be the subspace spanned by the eigenvectors of A corresponding to the eigenvalues located in a certain domain D of the spectrum of A; denote by $|u_D\rangle \equiv P_D|u\rangle$ the projection of ket $|u\rangle$ on $\mathscr{E}_D$. The probability w_D that the result of a measurement of A belongs to the domain D is* [1])

$$\boxed{w_D = \langle P_D \rangle = \frac{\langle u_D | u_D \rangle}{\langle u | u \rangle}} \qquad \text{(VIII.5)}$$

Relation (VIII.5) summarizes the results obtained in each particular case studied in Chapter V (Problem VIII.1). Indeed, D may simply be an eigenvalue of the discrete spectrum, in which case relation (VIII.5) is identical to relation (V.21). But D can equally well be made up of an assembly of several discrete eigenvalues, or of a portion of the continuous spectrum, or even of a combination of both. In particular, when D is an infinitesimal portion $[a(v), a(v+\mathrm{d}v)]$ of the continuous spectrum, as in the example at the end of Ch. V, § 10, $w_D = \omega(v)\,\mathrm{d}v$ and the probability density $\omega(v)$, calculated by means of relation (VIII.5) is precisely the one given by equation (V.44).

The next point to consider is how to define the dynamical state of the system, once the measurement is completed. The answer to this question depends upon the particular conditions under which the measurement was carried out; it is simple only in the case of an

[1]) w_D is the average value of the projector P_D, that is to say of the function of A equal to 1 for all eigenvectors of A located in $\mathscr{E}_D$, and equal to 0 for all eigenvectors of A orthogonal to $\mathscr{E}_D$.

ideal measurement (cf. Ch. V, § 13). If, with the hypothesis of an ideal measurement, the observation carried out on the system indicates that it is in an eigenstate of A belonging to the domain D defined above, *its dynamical state after measurement is represented by the projection of vector* $|u\rangle$ *upon space* $\mathscr{E}_D$. In other words, the (non-causal) evolution of the state vector in the course of the measurement corresponds to the scheme

$$|u\rangle \rightarrow \text{ideal measurement yielding result } D \rightarrow P_D|u\rangle$$

This postulate of the filtering of the wave packet may be regarded as a genuine definition of the ideal measurement.

With the convention that dynamical states are always represented by vectors normalized to unity – $|u\rangle$ is then assumed normalized to unity – the state vector of the system after measurement is the product $P_D|u\rangle$ by a normalization factor defined to within a phase, the square of whose modulus is equal to $1/w_D$, that is $1/\langle u|P_D|u\rangle$.

3. Observables of a Quantized System and Their Commutation Relations

The first step in the study of a quantized system consists in enumerating the dynamical variables of the system and in defining the algebra of the observables associated with the variables. In fact, the various observables can be expressed as function of some set of "fundamental observables"; one completes the definition of their algebra by giving the commutation relations of these fundamental observables.

When the considered quantum system possesses a classical analogue, as is the case for all those we have mentioned so far, one follows a general procedure based upon the correspondence principle.

In a classical system with N dimensions, the most general dynamical variable is a function of $2N$ independent variables, the N coordinates $q_1, q_2, \ldots, q_N$ and the N momenta $p_1, p_2, \ldots, p_N$. One assigns the same dynamical variables to the corresponding quantum system. One thus introduces N position variables and N momentum variables. To these variables correspond observables which we shall denote by the same symbols $q_1, q_2, \ldots, q_N, p_1, p_2, \ldots, p_N$ as the dynamical variables themselves. One postulates further that the only observables which do not commute are the N pairs formed by associating each coordinate

with its conjugate momentum; for such pairs one has $[q_r, p_r] = i\hbar$. In other words,

$$[q_r, q_s] = 0, \qquad [p_r, p_s] = 0 \tag{VIII.6}$$

$$[q_r, p_s] = i\hbar\delta_{rs} \tag{VIII.7}$$

$$(r, s = 1, 2, \ldots, N).$$

Since the most general observable is a function of the q's and the p's, the commutator of any two observables is defined unambiguously by specifying the fundamental commutator relations (VIII.6) and (VIII.7); one can actually calculate it explicitly by making use of the rules of commutator algebra (Chapter V, § 17). This correspondence between observables of a quantum system and quantities of the analogous classical system has already been discussed several times (Ch. II, § 15 and Ch. V, § 3). In order to remove all ambiguities, one always starts from the cartesian coordinates in configuration space and conforms to the empirical rules of Ch. II, § 15. In particular, the "symmetrization" rule given in that paragraph insures that with every real quantity connected with the system, there is associated a Hermitean operator.

All quantum systems cannot be treated by this correspondence method. It often happens that the dynamical variables introduced through correspondence with an appropriate classical analogue are not sufficient to exhaust all physical properties of the quantum system one wishes to study. It is then necessary to introduce additional variables. The choice of these new variables and of the commutation relations associated with them is purely a matter of intuition.

Among the physical quantities attached to a quantum system, special mention must be made of its energy. The observable H representing it is called the Hamiltonian of the system. When the system has a classical analogue, H is derived by correspondence from the Hamiltonian function of Classical Mechanics.

4. Heisenberg's Uncertainty Relations

The position-momentum uncertainty relations of Heisenberg result directly from the commutation relations (VIII.7).

Indeed, let us show in all generality that if two observables A and B satisfy the equation

$$[A, B] = i\hbar \tag{VIII.8}$$

the product of their root-mean-square deviations will always remain greater or equal to $\hbar/2$:

$$\Delta A \cdot \Delta B \geqslant \tfrac{1}{2}\hbar. \tag{VIII.9}$$

The proof we shall give is essentially the same as that of Ch. IV, § 8.

By definition,

$$\Delta A = (\langle A^2 \rangle - \langle A \rangle^2)^{\frac{1}{2}}, \qquad \Delta B = (\langle B^2 \rangle - \langle B \rangle^2)^{\frac{1}{2}}.$$

Let us introduce the observables

$$\hat{A} = A - \langle A \rangle, \qquad \hat{B} = B - \langle B \rangle.$$

Clearly

$$[\hat{A}, \hat{B}] = \mathrm{i}\hbar$$

and

$$\Delta A = \Delta \hat{A} = \langle \hat{A}^2 \rangle^{\frac{1}{2}}, \qquad \Delta B = \Delta \hat{B} = \langle B^2 \rangle^{\frac{1}{2}}.$$

Assume that the dynamical state of the system is represented by the ket $|u\rangle$ normalized to unity, and apply the Schwarz inequality to the vectors $\hat{A}|u\rangle$ and $\hat{B}|u\rangle$:

$$(\Delta A)^2 (\Delta B)^2 \equiv \langle u|\hat{A}^2|u\rangle \langle u|\hat{B}^2|u\rangle \geqslant |\langle u|\hat{A}\hat{B}|u\rangle|^2.$$

Separating in $\hat{A}\hat{B}$ the Hermitean from the anti-Hermitean part [cf. eq. (VII.29)]:

$$\hat{A}\hat{B} = \frac{\hat{A}\hat{B}+\hat{B}\hat{A}}{2} + \frac{\hat{A}\hat{B}-\hat{B}\hat{A}}{2} = \frac{\hat{A}\hat{B}+\hat{B}\hat{A}}{2} + \frac{\mathrm{i}\hbar}{2},$$

one can separate in $\langle u|\hat{A}\hat{B}|u\rangle$ the real from the imaginary part:

$$\langle u|\hat{A}\hat{B}|u\rangle = \left\langle \frac{\hat{A}\hat{B}+\hat{B}\hat{A}}{2} \right\rangle + \frac{\mathrm{i}\hbar}{2},$$

and rewrite the Schwarz inequality

$$(\Delta A)^2 \cdot (\Delta B)^2 \geqslant \left\langle \frac{\hat{A}\hat{B}+\hat{B}\hat{A}}{2} \right\rangle^2 + \frac{\hbar^2}{4}.$$

A fortiori

$$\Delta A \cdot \Delta B \geqslant \tfrac{1}{2}\hbar. \qquad \text{Q.E.D.}$$

In order that the product $\Delta A \cdot \Delta B$ be equal to its minimum value $\frac{1}{2}\hbar$, it is necessary on the one hand that the Schwarz inequality reduce

to an equality, hence that $\hat{A}|u\rangle = c\hat{B}|u\rangle$ (c an arbitrary constant), and on the other hand that the mean value of $\hat{A}\hat{B} + \hat{B}\hat{A}$ be zero, that is

$$\langle u|\hat{A}\hat{B}|u\rangle + \langle u|\hat{B}\hat{A}|u\rangle = (c^* + c)\,\langle u|\hat{B}^2|u\rangle = 0,$$

or Re $c = 0$. To sum up, for inequality (VIII.9) to reduce to an equality, it is necessary and sufficient that $|u\rangle$ satisfy the equation

$$(A - \alpha)|u\rangle = \mathrm{i}\gamma(B - \beta)|u\rangle \qquad \text{(VIII.10)}$$

in which α, β, and γ are arbitrary real constants.

The application of this general result to the position-momentum pairs (q_r, p_r) of the foregoing paragraph yields the uncertainty relations

$$\Delta q_r \cdot \Delta p_r \geqslant \tfrac{1}{2}\hbar \qquad (r = 1, 2, \ldots, N), \qquad \text{(VIII.11)}$$

the equality being actually realized if $|u\rangle$ is a solution of the equation

$$(p_r - \mathrm{i}\gamma q_r)|u\rangle = (\alpha - \mathrm{i}\gamma\beta)|u\rangle$$

(α, β, γ are arbitrary real constants).

5. Definition of the Dynamical States and Construction of Space $\mathscr{E}$

Once the observables of our quantum system have been enumerated and their commutation relations established, one must precisely define the various possible quantum states. One must construct the Hilbert space $\mathscr{E}$ in which these observables act. For this it is sufficient to select a basis of this space, and then to define each observable by its action upon the vectors of this basis. This action must be defined in such a way that all operators representing physical quantities actually turn out to be observables, and that the rules of algebra of these observables are actually obeyed.

To define the basis, one selects from the entire collection of observables a complete set of commuting observables $A, B, C, \ldots$. The simultaneous measurement of the dynamical variables which they represent constitutes the maximum of information one can obtain on the state of the system; it therefore completely defines a particular dynamical state of the system. Consequently, each set of eigenvalues a, b, c, ... of these observables defines a vector of $\mathscr{E}$ to within a constant: fixing this constant arbitrarily, one obtains a certain representative vector $|a\,b\,c\ldots\rangle$. The ensemble of vectors $|a\,b\,c\ldots\rangle$

obtained by letting each of the eigenvalues a, b, c, ... vary over the entire range of the respective spectra of A, B, C, ... forms a complete orthogonal set in $\mathscr{E}$-space. In fact, if one fixes the normalizations of the vectors $|a\,b\,c\,\ldots\rangle$ in an appropriate manner – normalization to unity for all vectors of finite norm, normalization by means of the Dirac δ-function for vectors of infinite norm – it is a complete *orthonormal* set in $\mathscr{E}$. One thus obtains a basis of $\mathscr{E}$ by specifying the respective spectra of the observables A, B, C, ...

The action of the basic observables A, B, C, ... on each of these vectors is obvious. The action of the various other observables susceptible of representing physical quantities is yet to be determined.

The mere knowledge of the commutation relation is generally sufficient –

(*i*) to show that the set A, B, C, ... forms a complete set of commuting observables;

(*ii*) to deduce their respective spectra;

(*iii*) to deduce the action of the other observables upon the vectors of their basis.

In other words, the knowledge of the algebra of the observables of the system in general suffices to define unambiguously the space $\mathscr{E}$ in which they act [1]).

It remains then to check the internal consistency of the construction thus defined; in other words, one must verify that the operators representing physical quantities are actually observables.

Note that at this stage the theory already lends itself to experimental check. The physical quantities are defined in principle by well-defined operations of measurement, and their spectrum is directly accessible to experiment. It is necessary that the theoretical spectrum, that is to say the eigenvalue spectrum of the observable associated with each physical quantity, coincide with this experimental spectrum.

6. One-Dimensional Quantum System Having a Classical Analogue

Let us apply the method of construction of § 5 to a one-dimensional quantum system having a classical analogue. Its observables are

[1]) This is correct only if $\mathscr{E}$ is irreducible with respect to the said observables; this condition, which we merely point out here for the record, is implicitly postulated in all the arguments below. We shall discuss more fully the notion of irreducibility and its physical significance in Chapter XV (§ 6).

functions of two of them, q and p, obeying the commutation relation

$$[q, p] = i\hbar. \tag{VIII.12}$$

The position q constitutes a complete set of commuting observables by itself. Indeed, the commutator of q and any given function $A(q, p)$ of p and q is [eq. (V.68)]

$$[q, A] = i\hbar \frac{\partial A}{\partial p}; \tag{VIII.13}$$

q commutes with A if, and only if A is independent of p; in other words, the only observables which commute with q are functions of q.

Simple considerations of internal consistency impose very restrictive conditions upon the eigenfunctions and the eigenvalue spectrum of q. Indeed, let $|q_0\rangle$ be an eigenket of q

$$q|q_0\rangle = q_0|q_0\rangle.$$

Let us write that the identical operators $[q, p]$ and $i\hbar$ have the same diagonal element corresponding to $|q_0\rangle$:

$$i\hbar \langle q_0|q_0\rangle = \langle q_0|qp|q_0\rangle - \langle q_0|pq|q_0\rangle.$$

$|q_0\rangle$ certainly does not have a finite norm; otherwise the right-hand side would vanish identically, while the left-hand side would be finite and non-zero.

On the other hand, consider the operator

$$S(\xi) = e^{-ip\xi/\hbar}. \tag{VIII.14}$$

It is a function of the observable p depending upon the parameter ξ. Clearly, it is a unitary operator:

$$S^\dagger S = SS^\dagger = 1,$$

since its Hermitean conjugate is

$$S^\dagger(\xi) = S(-\xi) = e^{-ip\xi/\hbar}.$$

Applying equation (VIII.12) yields

$$[q, S] = i\hbar \frac{\partial S}{\partial p} = \xi S.$$

In other words

$$qS = S(q+\xi) \tag{VIII.15}$$

and consequently,

$$qS|q_0\rangle = S(q+\xi)|q_0\rangle = (q_0+\xi)S|q_0\rangle. \tag{VIII.16}$$

Hence $S|q_0\rangle$ is an eigenvector of q belonging to the eigenvalue $(q_0+\xi)$. This vector is certainly not zero (otherwise S would not have an nverse); in fact its (infinite) norm is the same as that of $|q_0\rangle$ since S is unitary:

$$\langle q_0|S^\dagger S|q_0\rangle = \langle q_0|q_0\rangle.$$

This operation can be performed no matter what the value taken by ξ in the entire interval $(-\infty, +\infty)$. Thus, by a suitable unitary transformation on $|q_0\rangle$, one can form an eigenket of q belonging to any given eigenvalue in the interval $(-\infty, +\infty)$.

In conclusion, the spectrum of q is necessarily continuous, non-degenerate, and extends from $-\infty$ to $+\infty$; its eigenvectors necessarily have infinite norm.

Denote by $|q'\rangle$ one of the eigenkets of q belonging to the eigenvalue q':

$$q|q'\rangle = q'|q'\rangle;$$

$|q'\rangle$ is defined to within a constant whose modulus we fix by the normalization condition

$$\langle q'|q''\rangle = \delta(q'-q''). \tag{VIII.17}$$

The space $\mathscr{E}$ is by definition the space formed by linear superposition of the vectors $|q'\rangle$.

q is obviously an observable of that space. Actually, the vectors $|q'\rangle$ form the basis of a certain representation of vectors and operators of $\mathscr{E}$, the representation $\{q\}$, in which q is diagonal:

$$\langle q'|q|q''\rangle = q'\,\delta(q'-q''). \tag{VIII.18}$$

Let us show that p is a well-defined Hermitean operator in space $\mathscr{E}$; to this effect one merely has to determine its matrix in the representation $\{q\}$.

Consider first the unitary operator $S(\xi)$ defined by equation (VIII.14).

Since this operator satisfies eq. (VIII.15), $S(\xi)|q'\rangle$ is an eigenvector of q belonging to the eigenvalue $(q'+\xi)$:

$$S(\xi)|q'\rangle = c|q'+\xi\rangle,$$

c is a phase factor [1]) which may depend on ξ and on q'. We choose the phases of the vectors of the basis in such a way that

$$|q'\rangle = S(q')\,|0\rangle.$$

In this way the phase factor c is equal to 1, no matter what ξ and q'. Indeed

$$\begin{aligned} S(\xi)|q'\rangle &= S(\xi)\,S(q')|0\rangle = \mathrm{e}^{-\mathrm{i}p\xi/\hbar}\,\mathrm{e}^{-\mathrm{i}pq'/\hbar}\,|0\rangle \\ &= \mathrm{e}^{-\mathrm{i}p(\xi+q')/\hbar}|0\rangle = S(q'+\xi)|0\rangle \\ &= |q'+\xi\rangle, \end{aligned} \tag{VIII.19}$$

or else

$$\langle q'|S(\xi)|q''\rangle = \langle q'|q''+\xi\rangle = \delta(q'-q''-\xi).$$

From the matrix elements of $S(\xi)$ calculated above, we deduce the representative matrix of p by noting that in the limit where ξ is equal to an infinitesimal quantity ε,

$$S(\varepsilon) \sim 1 - \frac{\mathrm{i}}{\hbar}\,p\varepsilon$$

and, consequently

$$\delta(q'-q''-\varepsilon) = \langle q'|S(\varepsilon)|q''\rangle \sim \delta(q'-q'') - \frac{\mathrm{i}}{\hbar}\,\varepsilon\,\langle q'|p|q''\rangle,$$

whence

$$\langle q'|p|q''\rangle = \frac{\hbar}{\mathrm{i}} \lim_{\varepsilon\to 0} \frac{\delta(q'-q'') - \delta(q'-q''-\varepsilon)}{\varepsilon} = \frac{\hbar}{\mathrm{i}}\,\delta'(q'-q''). \tag{VIII.20}$$

As the "function" $\delta'(x)$ is odd, it is clear that $\langle q''|p|q'\rangle = \langle q'|p|q''\rangle^*$, hence that the operator p is Hermitean.

---

[1]) Since S is unitary, one has

$$\langle q''|S^\dagger(\xi)\,S(\xi)|q'\rangle = c^*(\xi, q'')\,c(\xi, q')\,\delta(q''-q') = \delta(q''-q'),$$

from which $|c(\xi, q')| = 1$.

Let us likewise verify that q and p actually satisfy the commutation relation (VIII.12)

$$\begin{aligned}\langle q'|(qp - pq)|q''\rangle &= (q' - q'')\langle q'|p|q''\rangle \\ &= \frac{\hbar}{\mathrm{i}}(q' - q'')\,\delta'(q' - q'') \\ &= \mathrm{i}\hbar\,\delta(q' - q'')\end{aligned}$$

[we have used the identity (A.30) of Appendix A].

It remains to be shown that p is an observable. To see this, we solve the eigenvalue problem of p in the representation $\{q\}$. Let $|p'\rangle$ be the eigenket belonging to the eigenvalue p'. The equation

$$p|p'\rangle = p'|p'\rangle$$

is written in the representation $\{q\}$, taking into account eq. (VIII.20),

$$\begin{aligned}p'\langle q'|p'\rangle = \langle q'|p|p'\rangle &= \int \langle q'|p|q''\rangle\,\mathrm{d}q''\,\langle q''|p'\rangle \\ &= \frac{\hbar}{\mathrm{i}}\int \delta'(q' - q'')\,\langle q''|p'\rangle\,\mathrm{d}q'' \\ &= \frac{\hbar}{\mathrm{i}}\frac{\mathrm{d}}{\mathrm{d}q'}(\langle q'|p'\rangle).\end{aligned}$$

This is a differential equation of the function $\langle q'|p\rangle$ of the variable q', whose general solution is

$$\langle q'|p'\rangle = a\,\mathrm{e}^{\mathrm{i}p'q'/\hbar},$$

where a is an arbitrary constant. We thus verify that p has a continuous spectrum of eigenvalues p' extending from $-\infty$ to $+\infty$. The eigenvectors have infinite norm. They satisfy the conditions of orthonormality

$$\langle p'|p''\rangle = \delta(p' - p'')$$

if one takes for the value of the constant: $a = (2\pi\hbar)^{-\frac{1}{2}}$. Now p is obviously an observable, since the vectors $|p'\rangle$ satisfy the closure relation. Indeed, the projection operator

$$P_p \equiv \int_{-\infty}^{+\infty} |p'\rangle\,\mathrm{d}p'\,\langle p'|$$

has for its matrix elements, in the representation $\{q'\}$,

$$\langle q'|P_p|q''\rangle = \int_{-\infty}^{+\infty} \langle q'|p'\rangle \, dp' \, \langle p'|q''\rangle$$

$$= \frac{1}{2\pi\hbar} \int_{-\infty}^{+\infty} e^{ip'(q'-q'')/\hbar} \, dp' = \delta(q' - q'');$$

whence the closure relation:

$$P_p = 1.$$

With the fundamental observables p, q one can build each of the operators $F(p, q)$ representing the various dynamical variables of the system. One can easily ensure that these operators are Hermitean. To be complete, one must show further that they actually are observables. It is traditional in Quantum Theory to pass over these points of rigor and to admit without proof that all Hermitean operators representing physical quantities are observables.

7. Construction of the $\mathscr{E}$ Space of a System by Tensor Product of Simpler Spaces

Knowing how to build up the $\mathscr{E}$ space for a system having a one-dimensional classical analogue, it is simple to solve the same problem for a system having a classical analogue in N dimensions.

In the latter case, the dynamical variables are functions of the $2N$ fundamental variables of position and momentum. The observables representing the latter obey the commutation relations (VIII.6) and (VIII.7). They may be grouped in N pairs (q_1, p_1), (q_2, p_2), ..., (q_N, p_N), each made up of one of the position components and its conjugate momentum. Each observable of a pair commutes with all the observables of the other pairs.

The observables of a given pair, (q_i, p_i) for instance, may be regarded as the fundamental observables of a one-dimensional system of the type studied in the preceding paragraph. One knows how to build up the ket space $\mathscr{E}_i$ of such a system. According to the results of § 6, $\mathscr{E}_i$ is spanned by the orthonormal kets $|q_i'\rangle$ whose index q_i' is continuous and may take on all values in the interval $(-\infty, +\infty)$.

The space $\mathscr{E}$ of the dynamical states of the N-dimensional system is the tensor product (cf. Ch. VII, § 6) of the spaces $\mathscr{E}_1, \mathscr{E}_2, \ldots, \mathscr{E}_N$

thus constructed

$$\mathscr{E} = \mathscr{E}_1 \otimes \mathscr{E}_2 \otimes \ldots \otimes \mathscr{E}_N,$$

that is to say the space spanned by the kets:

$$|q_1' q_2' \ldots q_N'\rangle \equiv |q_1'\rangle\, q_2'\rangle \ldots |q_N'\rangle. \tag{VIII.21}$$

To each operator q_i, p_i of the space $\mathscr{E}_i$ corresponds a well-defined operator q_i, p_i of the product space $\mathscr{E}$. To represent the $2N$ fundamental variables, we thus obtain $2N$ well-defined operators in $\mathscr{E}$. In fact, in accordance with the rules of tensor multiplication, to every observable of a partial space corresponds an observable of the product space; two observables coming from different partial spaces commute; two observables coming from the same partial space $\mathscr{E}_i$ obey the same commutation relations in $\mathscr{E}$ as in $\mathscr{E}_i$. Therefore, the $2N$ operators $q_1, \ldots, q_N, p_1, \ldots, p_N$ which we established in $\mathscr{E}$ are observables and actually obey the commutation relations (VIII.6) and (VIII.7).

The ensemble of vectors $|q_1', \ldots, q_N'\rangle$ obtained by varying each of the eigenvalues $q_1', \ldots, q_N'$ from $-\infty$ to $+\infty$ forms a basis in $\mathscr{E}$-space and defines a certain representation, the representation $\{q\}$. It is instructive to write down explicitly the matrix elements representing the q's and the p's in this representation. To this effect, we make use of the shorthand notation

$$|q'\rangle \equiv |q_1'\, q_2' \ldots q_N'\rangle \tag{VIII.22}$$

$$\left.\begin{aligned} \delta(q' - q'') &\equiv \prod_{i=1}^{N} \delta(q_i' - q_i'') \\ &\equiv \delta(q_1' - q_1'')\delta(q_2' - q_2'') \ldots \delta(q_N' - q_N'') \end{aligned}\right\} \tag{VIII.23}$$

$$\frac{\partial}{\partial q_n'}\,[\delta(q' - q'')] \equiv \delta'(q_n' - q_n'') \prod_{i \neq n} \delta(q_i' - q_i''). \tag{VIII.24}$$

In the last expression $\prod_{i \neq n}$ designates a product of the $(N-1)$ terms obtained by letting the index i assume all integral values from 1 to N with the exception of the value n.

Upon applying the relations (VIII.17), (VIII.18), and (VIII.20) of § 6, we successively obtain the orthonormality relations,

$$\langle q'|q''\rangle = \prod_{i=1}^{N} \langle q_i'|q_i''\rangle = \delta(q' - q''), \tag{VIII.25}$$

the elements of the (diagonal) matrices representing the coordinates

$$\begin{aligned}\langle q'|q_n|q''\rangle &= \langle q_n'|q_n|q_n''\rangle \prod_{i\neq n}\langle q_i'|q_i\rangle \\ &= q_n'\,\delta(q'-q''),\end{aligned} \tag{VIII.26}$$

and the elements of the (non-diagonal) matrices representing the momenta

$$\begin{aligned}\langle q'|p_n|q''\rangle &= \langle q_n'|p_n|q_n''\rangle \prod_{i\neq n}\langle q_i'|q_i''\rangle \\ &= \frac{\hbar}{i}\,\delta'(q_n'-q_n'')\prod_{i\neq n}\delta(q_i'-q_i'') \\ &= \frac{\hbar}{i}\,\frac{\partial}{\partial q_n'}\,[\delta(q'-q'')].\end{aligned} \tag{VIII.27}$$

One easily verifies the commutation relations (VIII.6) and (VIII.7) using these explicit expressions for the matrices representing the p's and the q's.

Any dynamical variable of the system is a function of the p's and the q's; therefore a well-defined operator of $\mathscr{E}$-space corresponds to it. One must make sure that this operator is an observable. In keeping with the tradition mentioned above, this point is in most cases admitted without proof.

The construction of the ket space of a system from the tensor product of simpler spaces is a very general procedure. In practice, one can always express the dynamical variables of a system as a function of a certain number of "basis" variables; now, the latter can usually be classified into a number of sets such that any variable belonging to a given set is compatible with all the variables of the other sets. To be specific, suppose that these "basis" variables have been grouped into two sets (A_1, B_1, ...) and (A_2, B_2, ...) and that each variable of the type (1) is compatible with every variable of type (2). Each set taken separately defines a partial system, whose ket space one knows how to construct. Let $\mathscr{E}_1$, $\mathscr{E}_2$ be the spaces of the partial systems (1) and (2), respectively. Clearly the ket space $\mathscr{E}$ of the total system is the tensor product of the two partial spaces

$$\mathscr{E} = \mathscr{E}_1 \otimes \mathscr{E}_2.$$

II. THE EQUATIONS OF MOTION

8. Evolution Operator and the Schrödinger Equation

From the fact that on the scale of precision of quantum phenomena, there exists no net separation between the system itself and the observing instrument, the evolution of a quantum system ceases to be strictly causal as soon as it is subjected to an observation. However, a quantum system isolated from any external influence evolves in an exactly predictable manner. Let $|\psi(t_0)\rangle$ be the ket vector representing its dynamical state at time t_0; the ket vector $|\psi(t)\rangle$ representing its state at the later time t is exactly determined by specifying $|\psi(t_0)\rangle$ if, as we shall assume henceforth, the system was not subjected to any observation during the time interval (t_0, t). This fundamental law of evolution will be discussed in this paragraph.

In the first place, we postulate that *the linear superposition of states is preserved in the course of time*. Consequently, the correspondence between $|\psi(t_0)\rangle$ and $|\psi(t)\rangle$ is linear and defines a certain linear operator $U(t, t_0)$ which is called the *evolution operator*:

$$|\psi(t)\rangle = U(t, t_0)\,|\psi(t_0)\rangle. \tag{VIII.28}$$

If the system is conservative, that is to say if its energy, as represented by the Hamiltonian H, does not depend explicitly upon the time, $U(t, t_0)$ can be deduced from the requirement that the motion of a system of energy E be periodic, and that its (angular) frequency ω be given by Einstein's law

$$E = \hbar\omega. \tag{VIII.29}$$

Indeed, since the eigenvectors of H span the space $\mathscr{E}$, it is sufficient, for a determination of U, to know its action upon each of these vectors. Let $|u_E(t_0)\rangle$ be an eigenvector of H corresponding to the energy E

$$H|u_E(t_0)\rangle = E|u_E(t_0)\rangle. \tag{VIII.30}$$

According to Einstein's law, we postulate that this vector evolves in the course of time according to the law

$$|u_E(t)\rangle = \mathrm{e}^{-\mathrm{i}\omega(t-t_0)}\,|u_E(t_0)\rangle = \mathrm{e}^{-\mathrm{i}E(t-t_0)/\hbar}\,|u_E(t_0)\rangle$$

or else, taking into account eq. (VIII.30),

$$|u_E(t)\rangle = \mathrm{e}^{-\mathrm{i}H(t-t_0)/\hbar}\,|u_E(t_0)\rangle.$$

Consequently

$$U(t, t_0) = e^{-iH(t-t_0)/\hbar}. \tag{VIII.31}$$

Upon differentiating [1]) the two sides of this equation with respect to t, we obtain the differential equation

$$\boxed{i\hbar \frac{d}{dt} U(t, t_0) = HU(t, t_0)} \tag{VIII.32}$$

$U(t, t_0)$ is the solution of this equation satisfying the initial condition

$$\boxed{U(t_0, t_0) = 1.} \tag{VIII.33}$$

By a very natural extension, we postulate that the operator $U(t_0, t_0)$ is the solution of the differential equation (VIII.32) satisfying the initial condition (VIII.33), *even when* the quantum system is not conservative. In the latter case, H depends upon the time explicitly, relation (VIII.29) loses all meaning, and the operator U is no longer given by eq. (VIII.31).

Note that U is likewise defined by the integral equation

$$\boxed{U(t, t_0) = 1 - \frac{i}{\hbar} \int_{t_0}^{t} HU(t', t_0)\, dt'} \tag{VIII.34}$$

The equations (VIII.32) and (VIII.33), or the integral equation (VIII.34) express the fundamental law of evolution of the quantum system. An equivalent expression of this law is the *Schrödinger equation*, that is, the differential equation of motion of the dynamical states of the system. One obtains this equation by differentiating equation (VIII.28) term by term,

$$\frac{d}{dt} |\psi(t)\rangle = \left(\frac{d}{dt} U(t, t_0)\right) |\psi(t_0)\rangle$$

---

[1]) The derivative with respect to t of an operator $X(t)$ depending upon a continuous parameter t is defined just as the derivative of a function:

$$\frac{dX}{dt} = \lim_{\varepsilon \to 0} \frac{X(t+\varepsilon) - X(t)}{\varepsilon}.$$

(Cf. Problem VIII.3.)

and substituting for $\mathrm{d}/\mathrm{d}t\ U(t, t_0)$ its expression (VIII.32). We obtain

$$\boxed{i\hbar \frac{\mathrm{d}}{\mathrm{d}t} |\psi(t)\rangle = H|\psi(t)\rangle.} \qquad \text{(VIII.35)}$$

In order that the norm of the vector $|\psi(t)\rangle$ remain constant in the course of time, it is necessary and sufficient that H be Hermitean; this is easily shown starting from the Schrödinger equation. The Hermitean property of the Hamiltonian will always be assumed to hold.

In fact, if H is Hermitean, $U(t, t_0)$ *is a unitary operator*. When H does not depend upon time, this can immediately be verified in the explicit expression for U, (VIII.31). However, even if H depends upon the time, we have according to the Schrödinger equation

$$|\psi(t + \mathrm{d}t)\rangle = \left(1 - \frac{\mathrm{i}}{\hbar} H\, \mathrm{d}t\right) |\psi(t)\rangle.$$

Since H is Hermitean, the operator

$$U(t + \mathrm{d}t, t) \equiv 1 - \frac{\mathrm{i}}{\hbar} H\, \mathrm{d}t$$

is an infinitesimal unitary operator (cf. Ch. VII, § 22): one passes from the ket vector at the time t, to the ket vector at the time $t + \mathrm{d}t$ by an infinitesimal unitary transformation. The transformation $U(t, t_0)$ which enables us to pass from $|\psi; t_0\rangle$ to $|\psi; t\rangle$ is thus a succession of infinitesimal unitary transformations; $U(t, t_0)$, a product of infinitesimal unitary operators, is unitary.

9. Schrödinger "Representation"

The Schrödinger equation completes the general scheme of description of quantum phenomena which we intended to present in this chapter. If we limit ourselves to the essentials, this scheme may be summarized in the following manner.

1) DEFINITION OF DYNAMICAL STATES.

The dynamical state of a quantum system is defined by a collection of precisely defined quantities, namely the particular values taken by the dynamical variables of a complete set of compatible variables. By carrying out the simultaneous measurement of the variables of

such a set, one defines unambiguously the state of the system at time t when the measurement was performed.

2) DEFINITION OF THE KET SPACE OF THE SYSTEM.

Each state can be represented (principle of superposition) by a ket vector $|\chi\rangle$ (normalized to unity and defined to within a phase factor) of a certain vector space $\mathscr{E}$. Each dynamical variable is represented by an observable of this space; the only states for which this variable has a well-defined value are the states represented by the eigenvectors of this observable, the value of the variable then being the eigenvalue belonging to the eigenvector in question. The observables satisfy certain rules of algebra which one can specify completely by giving the commutation relations. Compatible variables are represented by commuting observables.

3) DEFINITION OF PROBABILITIES.

If one performs on the quantum system a simultaneous measurement of a complete set of compatible dynamical variables, the probability of finding the system in the state $|\chi\rangle$ (i.e. of finding the particular values of these variables defining the dynamical state represented by $|\chi\rangle$) is equal to the square of the modulus of the scalar product of the vector $|\psi\rangle$ (normalized to unity) representing the dynamical state of the system at the instant the measurement is carried out, by $|\chi\rangle$, namely

$$|\langle\chi|\psi\rangle|^2.$$

More generally, the probability of finding the system in the subspace $\mathscr{E}_D$ (i.e. of finding the system in any one of the states of that subspace) is equal to the average of the projector P_D on that subspace, that is

$$\langle P_D\rangle = \langle\psi|P_D|\psi\rangle.$$

4) EQUATION OF MOTION.

In the absence of any external interference, the dynamical state of a system evolves in a strictly causal manner in the course of time. The vector $|\psi(t)\rangle$ which represents it in the space $\mathscr{E}$ moves continuously according to the Schrödinger equation (VIII.35). In other words, one goes over from $|\psi(t_0)\rangle$ to $|\psi(t)\rangle$ by the unitary transformation (VIII.28) in which $U(t, t_0)$ is a unitary operator defined by eqs. (VIII.32) and (VIII.33).

Knowing the dynamical state $|\psi\rangle$ of the system at a given initial

time t_0, one is thus able to predict the statistical distribution of the results of any measurement performed upon the system at a given later time t_1. Indeed, the dynamical state of the system at the instant when the measurement begins is

$$|\psi(t_1)\rangle = U(t_1, t_0)|\psi\rangle,$$

and consequently, the probability of finding the system in a given state $|\chi\rangle$ is

$$|\langle\chi|\psi(t_1)\rangle|^2 = |\langle\chi|U(t_1, t_0)|\psi\rangle|^2. \qquad \text{(VIII.36)}$$

In the above description of phenomena, the state of the system is represented by a moving ket vector $|\psi(t)\rangle$. On the other hand, the physical quantities, or at least those which do not depend upon the time explicitly, are represented by stationary observables of $\mathscr{E}$ space; likewise, the eigenvectors of these observables are stationary vectors of $\mathscr{E}$ space, as is the case for the vectors $|\chi\rangle$ and $|\psi\rangle$ of expression (VIII.36). This mode of description of quantum phenomena is called the *Schrödinger "representation"* [1]).

10. Heisenberg "Representation"

One obtains a mode of description of phenomena that is strictly equivalent to the foregoing one by performing a unitary transformation on the kets and observables of the Schrödinger "representation", and assigning to the transformed quantities the same physical significance as to those from which they originated. In such a transformation the observables transform into observables possessing the same eigenvalue spectrum; the eigenvectors transform into eigenvectors; the algebraic relations, the conjugation relations, and the scalar products are conserved. Since the only measurable quantities are moduli of scalar products [cf. eq. (VIII.36)], it is clear that the

[1]) One must not confuse this concept of "representation" with the notion of the representation of vectors and operators of vector space by matrices. The "representation" with which we are dealing here is that of the motion of the quantum system. In order to avoid confusion, it would be preferable here to speak of the Schrödinger "mode of description". The term "representation" is unfortunately entrenched by usage. In order to distinguish the two concepts, we agree to place the word "representation" between quotation marks whenever it is used in its present sense. The distinction to be made here is analogous to the distinction between unitary transformation of matrices and unitary transformation of vectors and operators (cf. Ch. VII, Sec. III).

predictions made by means of the new quantities are identical to the predictions made with the old quantities.

In particular, one defines the *Heisenberg "representation"* by performing upon the kets and observables of the Schrödinger "representation" the unitary, time-dependent transformation $U^{\dagger}(t, t_0)$. Let us attach the subscript S to the old quantities, and the subscript H to the new quantities. The ket

$$|\psi_S(t)\rangle = U(t, t_0)\,|\psi_S(t_0)\rangle$$

which represents the dynamical state of the system at time t is transformed into a stationary ket

$$|\psi_H\rangle = U^{\dagger}(t, t_0)|\psi_S(t)\rangle = |\psi_S(t_0)\rangle. \tag{VIII.37}$$

Conversely, an observable A_S of the Schrödinger "representation" transforms into

$$A_H(t) = U^{\dagger}(t, t_0)\, A_S U(t, t_0). \tag{VIII.38}$$

In general, A_H is not stationary, even when A_S does not depend upon the time explicitly. Indeed, if one takes into account differential equation (VIII.32) and its Hermitean conjugate equation, one obtains, on differentiating the last equation term by term

$$\begin{aligned} i\hbar \frac{dA_H}{dt} &= - U^{\dagger} H A_S U + i\hbar U^{\dagger} \frac{\partial A_S}{\partial t} U + U^{\dagger} A_S H U \\ &= U^{\dagger}[A_S, H] U + i\hbar U^{\dagger} \frac{\partial A_S}{\partial t} U. \end{aligned} \tag{VIII.39}$$

In this equation, H is the Hamiltonian of the Schrödinger "representation". Introducing the Hamiltonian of the Heisenberg "representation":

$$H_H = U^{\dagger} H U,$$

one has

$$U^{\dagger}[A_S, H]\, U = [A_H, H_H].$$

On the other hand, A_S, a function of the fundamental observables of the Schrödinger "representation", may depend upon the time explicitly; the right-hand side of (VIII.39) takes account of this fact. $\partial A_S/\partial t$ is some function of the observables of the Schrödinger "represetantion". If $\partial A_H/\partial t$ is the function obtained by replacing the latter

observables by the corresponding observables of the Heisenberg "representation", it is clear that

$$\frac{\partial A_{\mathrm{H}}}{\partial t} = U^{\dagger} \frac{\partial A_{\mathrm{S}}}{\partial t} U.$$

Equation (VIII.39) is therefore written

$$\boxed{\mathrm{i}\hbar \frac{\mathrm{d}A_{\mathrm{H}}}{\mathrm{d}t} = [A_{\mathrm{H}}, H_{\mathrm{H}}] + \mathrm{i}\hbar \frac{\partial A_{\mathrm{H}}}{\partial t}.} \qquad \text{(VIII.40)}$$

This equation is known as the *Heisenberg equation.*

In conclusion, the Heisenberg "representation" is obtained by imposing upon the vector space of the Schrödinger "representation" an overall motion chosen in such a way that the dynamical state of the quantum system is represented by a stationary ket $|\psi_{\mathrm{H}}\rangle$. In other words, *any stationary ket of the Heisenberg "representation" represents a possible motion of the quantum system.* Conversely, the various physical quantities are represented by observables evolving in the course of time according to the law (VIII.38) or, what amounts to the same thing, according to the Heisenberg equation (VIII.40) with the initial condition $A_{\mathrm{H}}(t_0) = A_{\mathrm{S}}(t_0)$.

Equations (VIII.38) and (VIII.40) apply equally well to any function of the observables of the Heisenberg representation, and especially to the expression $\exp(\mathrm{i}\xi A_{\mathrm{H}})$ or to the projector $P_D^{(\mathrm{H})}$ upon the subspace of the eigenvectors belonging to the eigenvalues of a given domain D of the spectrum of A_{H}.

Likewise, a ket $|\chi_{\mathrm{H}}\rangle$ representing an ensemble of compatible variables generally depends upon the time and is deduced from its homologue $|\chi_{\mathrm{S}}\rangle$ of the Schrödinger "representation" by the formula

$$|\chi_{\mathrm{H}}(t)\rangle = U^{\dagger}(t, t_0)|\chi_{\mathrm{S}}\rangle. \qquad \text{(VIII.41)}$$

Let us suppose that the motion of the quantum system is represented after time t_0 by the (stationary) ket $|\psi_{\mathrm{H}}\rangle$. The probability of finding it in the state $|\chi_{\mathrm{H}}\rangle$ upon performing a measurement at the later time t_1 is

$$|\langle\chi_{\mathrm{H}}(t_1)|\psi_{\mathrm{H}}\rangle|^2,$$

a quantity which is obviously equal to the one obtained with the corre-

sponding kets of the Schrödinger "representation" [eq. (VIII.36)], since the scalar product is invariant under the unitary transformation $U^{\dagger}(t_1, t_0)$.

11. Heisenberg "Representation" and Correspondence Principle

As shown above, the "representations" of Schrödinger and of Heisenberg are strictly equivalent. In practice, the Schrödinger "representation" is more often used, because it lends itself better to calculations. In fact, the Schrödinger equation, an equation between vectors, is *a priori* easier to solve than the Heisenberg equation which is an equation between operators. However, certain general properties of quantum systems are more immediately apparent in the Heisenberg "representation".

The formal analogy between the Classical Theory and the Quantum Theory is particularly striking in the Heisenberg "representation". The motion of a quantum system in fact, just like the motion of a classical system, appears in the Heisenberg "representation" as a motion of the dynamical variables which are associated with it.

Let us consider a quantum system possessing a classical analogue, and let us compare the motions of the two systems. To every physical quantity of the classical system corresponds a physical quantity of the quantum system. The only difference lies in the fact that the physical quantities of the classical system are quantities obeying the rules of ordinary algebra, whereas their quantum analogues are operators obeying the rules of a non-commutative algebra. But, to the extent that one can identify the expressions of a non-commutative algebra with expressions of ordinary algebra, the equations of motion of the quantized quantities are identical to those of their classical analogues. Indeed, the Heisenberg equations for the variables $q_1, \ldots, q_N$ and $p_1, \ldots, p_N$ are written

$$\begin{aligned} \frac{\mathrm{d}q_i}{\mathrm{d}t} &= \frac{1}{\mathrm{i}\hbar}[q_i, H] = \frac{\partial H}{\partial p_i} \qquad (i = 1, 2, \ldots, N) \\ \frac{\mathrm{d}p_i}{\mathrm{d}t} &= \frac{1}{\mathrm{i}\hbar}[p_i, H] = -\frac{\partial H}{\partial q_i} \qquad (i = 1, 2, \ldots, N). \end{aligned} \tag{I}$$

To obtain these expressions, the fundamental commutation relations between the q's and the p's as well as the properties (V.67) and (V.68) to which they give rise, have been taken into account. The

system of equations (I) is formally identical to Hamilton's canonical equations of Classical Mechanics.

More generally, a classical dynamical variable $A_{\text{cl.}} = A(q_1, \ldots, q_N; p_1, \ldots, p_N; t)$ obeys the equation of motion

$$\frac{\mathrm{d}A_{\text{cl.}}}{\mathrm{d}t} = \{A_{\text{cl.}}, H_{\text{cl.}}\} + \frac{\partial A_{\text{cl.}}}{\partial t}, \tag{VIII.42}$$

where $\{A_{\text{cl.}}, H_{\text{cl.}}\}$ designates the Poisson bracket of $A_{\text{cl.}}$ and $H_{\text{cl.}}$ according to the definition

$$\{A, H\} \equiv \sum_i \left(\frac{\partial A}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial A}{\partial p_i} \frac{\partial H}{\partial q_i} \right).$$

We see that the classical equation (VIII.42) is identical to the corresponding Heisenberg equation to the extent that one can identify the Poisson bracket $\{A, H\}$ with the commutator $[A_{\text{H}}, H_{\text{H}}]/i\hbar$. Making use of the fundamental commutation relations and of the similarity between the rules of commutator algebra and the rules of Poisson bracket algebra, one can actually prove the identity of these two expressions provided one makes a suitable choice of the order of the q's and the p's in the explicit expression of the Poisson bracket.

12. Constants of the Motion

The notion of constant of the motion is particularly simple to grasp in the Heisenberg "representation". A dynamical variable *which does not depend upon the time explicitly* is a constant of the motion if the observable C_{H} representing it in the Heisenberg "representation" remains constant in time. Consequently, its system of eigenvectors remains stationary and the statistical distribution of the results of a possible measurement of this quantity is always independent of the time at which this measurement is undertaken.

According to the above definition of the constant of the motion

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t} C_{\text{H}} = [C_{\text{H}}, H_{\text{H}}] = 0.$$

The constants of the motion are thus represented by *observables which commute with the Hamiltonian*. This result holds true equally well in the Schrödinger "representation" and in the Heisenberg "representation"

since the commutation relations are conserved in passing from one to the other.

Moreover, since C_H is time-independent, it is equal to its initial value C_S

$$C_H(t) = C_H(t_0) = C_S = C.$$

If, in particular, the dynamical state of the system is represented by an eigenvector of C in the Heisenberg "representation",

$$C|\psi_H\rangle = c|\psi_H\rangle,$$

the variable C keeps the same well-defined value c in the course of time; the eigenvalue c is then said to be a *good quantum number*. As can easily be shown, C commutes with the evolution operator $U(t, t_0)$; hence, the ket $|\psi_S(t)\rangle$ of the Schrödinger "representation" remains forever in the subspace of the eigenvalue c,

$$C|\psi_S(t)\rangle = c|\psi_S(t)\rangle.$$

13. Equations of Motion for the Mean Values. Time-Energy Uncertainty Relation

Starting from the Heisenberg "representation", it is particularly simple to write down a differential equation for the mean value of a given observable A_H. Indeed, since $|\psi_H\rangle$ is time-independent,

$$\frac{d\langle A\rangle}{dt} = \frac{d}{dt}\langle\psi_H|A_H|\psi_H\rangle = \langle\psi_H|\frac{dA_H}{dt}|\psi_H\rangle.$$

Using the Heisenberg equation, we arrive at eq. (V.72) once again:

$$\frac{d}{dt}\langle A\rangle = \frac{1}{i\hbar}\langle[A, H]\rangle + \left\langle\frac{\partial A}{\partial t}\right\rangle. \qquad \text{(VIII.43)}$$

In particular, one obtains the Ehrenfest equations (Ch. VI, § 2) by carrying out this manipulation on system (I).

As an application of eq. (VIII.43), we shall give a precise statement of the time-energy uncertainty relation (cf. Ch. IV, § 10). Consider a system whose Hamiltonian H does not explicitly depend upon the time, and let A be another observable of this system which does not depend upon the time explicitly. We consider the dynamical state of the system at a given time t. Let $|\psi\rangle$ be the vector representing

that state. Call ΔA, ΔE the root-mean-square deviations of A and of H, respectively. Applying the Schwarz inequality to the vectors $(A-\langle A\rangle)|\psi\rangle$ and $(H-\langle H\rangle)|\psi\rangle$ and carrying out the same manipulations as in § 4, we find after some calculations

$$\Delta A \cdot \Delta E \geqslant \tfrac{1}{2}|\langle [A, H]\rangle|, \tag{VIII.44}$$

the equality being realized when $|\psi\rangle$ satisfies the equation

$$(A-\alpha)|\psi\rangle = \mathrm{i}\gamma(H-\varepsilon)|\psi\rangle$$

where α, γ, and ε are arbitrarily real constants [cf. eq. (VIII.10)]. However, according to eq. (VIII.43), $\langle [A, H]\rangle = \mathrm{i}\hbar\, \mathrm{d}\langle A\rangle/\mathrm{d}t$; the inequality (VIII.44) may equally well be written

$$\frac{\Delta A}{|\mathrm{d}\langle A\rangle/\mathrm{d}t|} \cdot \Delta E \geqslant \tfrac{1}{2}\hbar,$$

or else

$$\tau_A \cdot \Delta E \geqslant \tfrac{1}{2}\hbar \tag{VIII.45}$$

if one puts

$$\tau_A = \frac{\Delta A}{|\mathrm{d}\langle A\rangle/\mathrm{d}t|}\,; \tag{VIII.46}$$

τ_A appears as a time characteristic of the evolution of the statistical distribution of A. It is the time required for the center $\langle A\rangle$ of this distribution to be displaced by an amount equal to its width ΔA; in other words, it is the time necessary for this statistical distribution to be appreciably modified. In this manner we can define a characteristic evolution time for each dynamical variable of the system.

Let τ be the shortest of the times thus defined. τ may be considered as a characteristic time of evolution of the system itself: whatever the measurement carried out on the system at an instant of time t', the statistical distribution of the results is essentially the same as would be obtained at the instant t, as long as the difference $|t-t'|$ is less than τ.

According to the inequality (VIII.45), this time τ and the energy spread ΔE satisfy the time-energy uncertainty relation

$$\tau \cdot \Delta E \geqslant \tfrac{1}{2}\hbar. \tag{VIII.47}$$

If, in particular, the system is in a stationary state, $\mathrm{d}\langle A\rangle/\mathrm{d}t = 0$ no matter what A, and consequently τ is infinite; however, $\Delta E = 0$, in conformity with relation (VIII.47).

14. Intermediate "Representations"

The Schrödinger and Heisenberg "representations" are not the only possible ones. Any unitary transformation of the vectors and the observables of the Schrödinger (or Heisenberg) "representations" defines a new "representation". All these "representations" furnish strictly equivalent descriptions of quantum phenomena. In practice, one therefore adopts the "representation" which lends itself best to the solution of each particular problem.

Any problem of Quantum Mechanics essentially consists of a more or less complete and more or less precise determination of the properties of the unitary operator $U(t, t_0)$; indeed, all the predictions of the theory are given by matrix elements of $U(t, t_0)$ such as the one occurring in eq. (VIII.36). The solution of equation (VIII.32) is thus the central problem of the theory. When one knows an approximate solution $U^{(0)}(t, t_0)$ of this equation, it is often convenient to set

$$U = U^{(0)}U'. \tag{VIII.48}$$

Substituting this expression in eq. (VIII.32) and multiplying both sides from the left by the unitary operator $U^{(0)\dagger}$, we obtain the differential equation

$$i\hbar \frac{d}{dt} U' = U^{(0)\dagger}\left(HU^{(0)} - i\hbar \frac{dU^{(0)}}{dt}\right) U'; \tag{VIII.49}$$

U' is the solution of this equation satisfying the initial condition

$$U'(t_0, t_0) = 1.$$

If the approximation is justified, U' is an operator changing slowly as a function of time; this is actually quite evident from eq. (VIII.49), since in that case, the operator $HU^{(0)} - i\hbar(dU^{(0)}/dt)$ almost vanishes. Equation (VIII.49), therefore, lends itself better than equation (VIII.32) to an approximate solution [1]).

Since $U^{(0)}$ is unitary, the operator

$$H^{(0)}(t) \equiv i\hbar\left[\frac{d}{dt} U^{(0)}(t, t_0)\right] U^{(0)\dagger}(t, t_0)$$

[1]) The procedure discussed here is the generalization to the differential equations between operators, of the method of variation of constants of the elementary theory of differential equations.

is Hermitean (Problem VIII.6). $U^{(0)}(t, t_0)$ is therefore the rigorous solution of the Schrödinger equation:

$$i\hbar \frac{d}{dt} U^{(0)} = H^{(0)} U^{(0)}, \qquad U^{(0)}(t_0, t_0) = 1. \tag{VIII.50}$$

The Hamiltonian H is thus the sum of two Hermitean operators

$$H = H^{(0)} + H'$$

of which one, H', may be considered as a small perturbation, under the hypothesis considered here, while the other, $H^{(0)}$, is the Hamiltonian of a Schrödinger equation whose solution is known. With this notation, eq. (VIII.49) can then simply be written

$$i\hbar \frac{d}{dt} U' = H_I' U', \tag{VIII.51}$$

an expression in which H_I' is deduced from H' by the time-dependent unitary transformation

$$H_I' = U^{(0)\dagger} H' U^{(0)}. \tag{VIII.52}$$

It is then convenient to adopt, throughout the discussion, a "representation" intermediate between that of Schrödinger and of Heisenberg; namely the one obtained by applying to the vectors and observables of the Schrödinger "representation" the unitary transformation $U^{(0)\dagger}(t, t_0)$. Let us label the vectors and observables of this new representation by a subscript I:

$$|\psi_I(t)\rangle = U^{(0)\dagger}|\psi_S(t)\rangle. \tag{VIII.53}$$

$$A_I(t) = U^{(0)\dagger} A_S U^{(0)}. \tag{VIII.54}$$

In this intermediate "representation", the vector $|\psi_I(t)\rangle$ representing a possible motion of the quantum system is equal to $U'|\psi_S(t_0)\rangle$. According to eq. (VIII.51), this vector moves (slowly) in time satisfying a Schrödinger equation whose Hamiltonian is the perturbation energy H_I':

$$i\hbar \frac{d}{dt} |\psi_I(t)\rangle = H_I'|\psi_I(t)\rangle. \tag{VIII.55}$$

On the other hand, the physical quantities are represented by moving observables; these observables are in fact subject to the Heisenberg

equations of motion written with the "unperturbed" Hamiltonian $H_I^{(0)}$

$$i\hbar \frac{d}{dt} A_I = [A_I, H_I^{(0)}] + i\hbar \frac{\partial A_I}{\partial t}, \qquad \text{(VIII.56)}$$

as one can easily show by carrying out on eq. (VIII.54) a procedure analogous to the one applied to eq. (VIII.38) in order to establish the Heisenberg equation.

III. VARIOUS REPRESENTATIONS OF THE THEORY

15. Definition of a Representation

According to the theory developed in the first two sections, all the elements for the description of a quantum system are present when one has defined its fundamental dynamical variables, the commutation relations obeyed by the representative observables, and the explicit expression, as a function of these fundamental observables, of the Hamiltonian which governs the motion of the system. One can then build up the $\mathscr{E}$ space of the vectors representing the various possible dynamical states of the system; define the physical meaning of the vectors of $\mathscr{E}$ by solving the eigenvalue problems for the various observables associated with the system, establish and solve the fundamental equations of motion, and finally, perform the calculation of the statistical distribution of the results of measurement which the theory is in a position to predict exactly.

In order to solve all these problems of analysis or of algebra in $\mathscr{E}$ space, one can always, and in an infinite number of ways at that, choose in that space a complete, orthonormal set of vectors and represent the operators and vectors of $\mathscr{E}$ by their matrices in the representation having this set of vectors as a basis.

In this way, any dynamical variable of the system is represented by a square Hermitean matrix, and any dynamical state by a column vector (or equally well by the Hermitean conjugate row vector) defined to within a constant.

There are as many possible representations of the Theory as there are distinct bases. One passes from one to the other by unitary transformation. These unitary transformations of matrices must not be confused with the unitary transformations of operators and vectors which permit, according to the discussion of Sec. II, to change the "representation" of the motion of the quantum system itself.

Most often a representation is defined by giving a complete set of commuting observables; their common eigenvectors are the basis of the representation. The basis observables of the representation and all functions of these observables are represented by diagonal matrices.

16. Wave Mechanics

Wave Mechanics is obtained by formulating the Quantum Theory in the Schrödinger "representation" and in a representation where the position variables are diagonal.

Let us return to the quantum system having a classical analogue in N dimensions, studied in § 7. The position coordinates $q_1, q_2, \ldots, q_N$ form a complete set of commuting observables and define a certain representation, the $\{q\}$ representation. This representation has already been used to construct the $\mathscr{E}$ space itself. With a suitable choice of phases and normalization for the basis vectors, we obtained very simple expressions for the matrix elements of the q's and the p's [eqs. (VIII.26) and (VIII.27)].

The fundamental equations of the $\{q\}$ representation are the orthonormality relations (VIII.25) and the closure relation which, following the abbreviated notation of § 7 ($\mathrm{d}q \equiv \mathrm{d}q_1\, \mathrm{d}q_2 \ldots \mathrm{d}q_N$), is written

$$P_q \equiv \int |q'\rangle\, \mathrm{d}q'\, \langle q'| = 1. \qquad \text{(VIII.57)}$$

Any ket $|\psi\rangle$ is represented by the single-column matrix with components $\langle q'|\psi\rangle$. This function of the coordinates $q_1', q_2', \ldots, q_N'$ of configuration space, which can equally well be written $\psi(q_1', q_2', \ldots, q_N')$, is the wave function representing the dynamical state of the system in the language of Wave Mechanics:

$$\langle q'|\psi\rangle \equiv \langle q_1'\, q_2' \ldots q_N'|\psi\rangle \equiv \psi(q_1', q_2', \ldots, q_N'). \qquad \text{(VIII.58)}$$

The scalar product of $|\psi\rangle$ by $|\varphi\rangle$ is equal to the scalar product of the corresponding wave functions as defined in Wave Mechanics:

$$\langle\varphi|\psi\rangle = \langle\varphi|P_q|\psi\rangle = \int \langle\varphi|q'\rangle\, \mathrm{d}q'\, \langle q'|\psi\rangle = \int \varphi^*(q')\, \psi(q')\, \mathrm{d}q'. \qquad \text{(VIII.59)}$$

Let us verify the identity between the operators of Wave Mechanics and the matrices representing the observables in the $\{q\}$ representation.

Using expression (VIII.26) of the matrix representing the observable q_n, we verify that $q_n|\psi\rangle$ is represented by the wave function

$$\langle q'|q_n|\psi\rangle = q_n'\langle q'|\psi\rangle = q_n'\psi(q').$$

More generally, the action of an arbitrary function $V(q) \equiv V(q_1, q_2, \ldots, q_N)$ of the coordinates of configuration space, on a ket $|\psi\rangle$ is represented by the function obtained by multiplying $\psi(q')$ by $V(q')$,

$$\langle q'| V(q) |\psi\rangle = V(q')\, \psi(q'). \tag{VIII.60}$$

Using the explicit expression (VIII.27) of the matrix representing the observable p_n, we verify that $p_n|\psi\rangle$ is represented by the wave function

$$\begin{aligned} \langle q'|p_n|\psi\rangle &= \int \langle q'|p_n|q''\rangle \, \mathrm{d}q'' \, \langle q''|\psi\rangle \\ &= \frac{\hbar}{\mathrm{i}} \int \frac{\partial}{\partial q_n'} [\delta(q' - q'')]\, \psi(q'')\, \mathrm{d}q'' \\ &= \frac{\hbar}{\mathrm{i}} \frac{\partial}{\partial q_n'} \psi(q'). \end{aligned} \tag{VIII.61}$$

p_n is therefore represented by the operation of partial differentiation $-\mathrm{i}\hbar\partial/\partial q_n$ applied to the wave function to the right.

The identification which we intended to make is therefore actually verified for the functions of position coordinates [eq. (VIII.60)] and for the components of the momentum [eq. (VIII.61)]. Now, since any observable is an algebraic function of the p's, and of functions of the q's, this identification holds true for any observable; hence, any physical quantity $A(q; p)$ is represented in Wave Mechanics by he operator

$$A\left(q; \frac{\hbar}{\mathrm{i}} \frac{\partial}{\partial q}\right).$$

As an illustration, let us consider the energy H of the system, assuming a purely static potential energy. H reads

$$H(q; p) = \sum_i \frac{p^2}{2m_i} + V(q_1, \ldots, q_N).$$

The matrix which represents the energy in the $\{q\}$ representation is of the form

$$\begin{aligned} \langle q'|H|q''\rangle &= H\left(q'; \frac{\hbar}{\mathrm{i}} \frac{\partial}{\partial q'}\right) \delta(q' - q'') \\ &= \left[\sum_i \left(-\frac{\hbar^2}{2m_i}\right) \frac{\partial^2}{\partial q_i'^2} + V(q')\right] \delta(q' - q''). \end{aligned}$$

In the last line, the expression in brackets must be regarded as an operator acting on $\delta(q'-q'')$ considered as a function of the q'. Therefore, the vector $H|\psi\rangle$ is represented by the wave function

$$\langle q'|H|\psi\rangle \equiv H\psi(q') = \left[\sum_i -\frac{\hbar^2}{2m_i}\frac{\partial}{\partial {q_i'}^2} + V(q')\right]\psi(q').$$

Finally, if we write, in the framework of the Schrödinger "representation", the fundamental equation of motion (VIII.35) in the $\{q\}$ representation, we again arrive at the Schrödinger equation in its usual form:

$$i\hbar\frac{\partial}{\partial t}\psi(q;t) = H\psi(q;t).$$

This completes the proof of the identity between Wave Mechanics and the formulation of Quantum Theory in the $\{q\}$ representation and in the Schrödinger "representation".

17. Momentum Representation. ($\{p\}$ Representation)

As another example of representation, we mention the $\{p\}$ representation, in which the components of the momentum are diagonal. Let us denote by $|p'\rangle \equiv |p_1'\rangle\,|p_2'\rangle \ldots |p_N'\rangle$ the basis vectors of this representation. They are the eigenvectors common to $p_1, p_2, \ldots, p_N$ belonging to the eigenvalues $p_1', p_2', \ldots, p_N'$. These vectors are assumed to be orthonormal:

$$\langle p'|p''\rangle = \delta(p'-p'').$$

They satisfy the closure relation

$$P_p \equiv \int |p'\rangle\,\mathrm{d}p'\,\langle p'| = 1,$$

(we make use in this whole section of abbreviated notations analogous to those introduced in §§ 7 and 16).

In accordance with the results of § 6, the wave function of the vector $|p'\rangle$ in the $\{q\}$ representation is

$$\langle q'|p'\rangle \equiv \prod_{i=1}^{N}\langle q_i'|p_i'\rangle = (2\pi\hbar)^{-\frac{1}{2}N}\,\mathrm{e}^{\mathrm{i}(p_1'q_1'+\ldots+p_N'q_N')/\hbar}.$$

Considered as a function of the q' and the p', $\langle p'|q'\rangle$ $(=\langle q'|p'\rangle^*)$ is the unitary matrix S which transforms the matrices of the $\{q\}$

representation into the matrices of the $\{p\}$ representation. A ket vector $|\psi\rangle$ is represented in the latter by its "wave function in momentum space"

$$\Phi(p') \equiv \langle p'|\psi\rangle.$$

Clearly $\Phi(p')$ is the Fourier transform (suitably normalized) of the wave function $\Psi(q') \equiv \langle q'|\psi\rangle$ of configuration space:

$$\begin{aligned}\Phi(p') = \langle p'|\psi\rangle &= \int \langle p'|q'\rangle \, dq' \, \langle q'|\psi\rangle \\ &= (2\pi\hbar)^{-\frac{1}{2}N} \int \Psi(q') \, e^{-i(p_1' q_1' + \ldots + p_N' q_N')/\hbar} \, dq'.\end{aligned}$$

It is easy to show (one could in fact prove it directly) that the action of the operator p_n on the function $\Phi(p')$ reduces to a multiplication by p_n' and that the action of q_n is represented by the operation of partial differentiation $i\hbar\partial/\partial p_n'$.

As an example, let us write the Schrödinger equation of a particle of mass m in a static potential $V(\boldsymbol{r})$ in the $\{p\}$ representation. The energy of the system is represented by the observable

$$H(\boldsymbol{r}, \boldsymbol{p}) \equiv \frac{p^2}{2m} + V(\boldsymbol{r}).$$

The basis vectors $|\boldsymbol{p}'\rangle$ depend upon the three momentum coordinates p_x', p_y', p_z' and satisfy the orthonormality and closure relations:

$$\delta(\boldsymbol{p}' - \boldsymbol{p}'') \equiv \delta(p_x' - p_x'') \, \delta(p_y' - p_y'') \, \delta(p_z' - p_z'') = \langle \boldsymbol{p}'|\boldsymbol{p}''\rangle$$

$$P_p \equiv \int |\boldsymbol{p}'\rangle \, d\boldsymbol{p}' \, \langle \boldsymbol{p}'| = 1.$$

The unitary transformation matrix S which transforms the matrices of the $\{\boldsymbol{r}\}$ representation into matrices of the $\{\boldsymbol{p}\}$ representation is

$$\langle \boldsymbol{p}'|\boldsymbol{r}'\rangle = (2\pi\hbar)^{-3/2} \, e^{-i(\boldsymbol{p}'\cdot\boldsymbol{r}')/\hbar}.$$

Therefore, the matrix elements of $V(\boldsymbol{r})$ in the $\{\boldsymbol{p}\}$ representation read explicitly

$$\begin{aligned}\langle \boldsymbol{p}'|V|\boldsymbol{p}''\rangle &= \iint \langle \boldsymbol{p}'|\boldsymbol{r}'\rangle \, d\boldsymbol{r}' \, \langle \boldsymbol{r}'|V(\boldsymbol{r})|\boldsymbol{r}''\rangle \, d\boldsymbol{r}'' \, \langle \boldsymbol{r}''|\boldsymbol{p}''\rangle \\ &= (2\pi\hbar)^{-3} \iint e^{-i(\boldsymbol{p}'\cdot\boldsymbol{r}')/\hbar} \, d\boldsymbol{r}' \; V(\boldsymbol{r}') \, \delta(\boldsymbol{r}' - \boldsymbol{r}'') \, d\boldsymbol{r}'' \, e^{i(\boldsymbol{p}''\cdot\boldsymbol{r}'')/\hbar} \\ &= (2\pi\hbar)^{-3} \int V(\boldsymbol{r}') \, e^{-i(\boldsymbol{p}' - \boldsymbol{p}'')\cdot\boldsymbol{r}'/\hbar} \, d\boldsymbol{r}'.\end{aligned}$$

Let us put

$$\mathscr{V}(\boldsymbol{k}) = (2\pi\hbar)^{-3} \int V(\boldsymbol{r}) \, e^{-i(\boldsymbol{k}\cdot\boldsymbol{r})/\hbar} \, d\boldsymbol{r}.$$

We then have

$$\langle \mathbf{p}' | V | \mathbf{p}'' \rangle = \mathscr{V}(\mathbf{p}' - \mathbf{p}'')$$

and the matrix elements in the $\{\mathbf{p}\}$ representation of the operator H are:

$$\langle \mathbf{p}' | H | \mathbf{p}'' \rangle = \frac{p'^2}{2m} \delta(\mathbf{p}' - \mathbf{p}'') + \mathscr{V}(\mathbf{p}' - \mathbf{p}'').$$

Let us denote by $\Phi(\mathbf{p}')$ the wave function in momentum space of a dynamical state $|\psi\rangle$ whose wave function in the usual sense of the term is $\Psi(\mathbf{r}')$:

$$\begin{aligned} \Phi(\mathbf{p}') &= \langle \mathbf{p}' | \psi \rangle = \int \langle \mathbf{p}' | \mathbf{r}' \rangle \, \mathrm{d}\mathbf{r}' \, \langle \mathbf{r}' | \psi \rangle \\ &= (2\pi\hbar)^{-3/2} \int \mathrm{e}^{-\mathrm{i}(\mathbf{p}' \cdot \mathbf{r}')/\hbar} \, \Psi(\mathbf{r}') \, \mathrm{d}\mathbf{r}'. \end{aligned}$$

The Schrödinger equation which is written

$$\mathrm{i}\hbar \frac{\partial}{\partial t} \Psi(\mathbf{r}; t) = \left[-\frac{\hbar^2}{2m} \triangle + V(\mathbf{r}) \right] \Psi(\mathbf{r}; t)$$

in Wave Mechanics, takes the form of an integro-differential equation in the $\{\mathbf{p}\}$ representation.

$$\mathrm{i}\hbar \frac{\partial}{\partial t} \Phi(\mathbf{p}; t) = \frac{p^2}{2m} \Phi(\mathbf{p}; t) + \int \mathscr{V}(\mathbf{p} - \mathbf{p}') \, \Phi(\mathbf{p}'; t) \, \mathrm{d}\mathbf{p}'.$$

18. An Example: Motion of a Free Wave Packet

As an application of the foregoing considerations, consider the motion of the free wave packet ($V = 0$).

Denote by $|\psi\rangle$ the state vector at the time $t = 0$, by $\psi(\mathbf{r})$ and $\varphi(\mathbf{p})$ the wave functions representing this vector in the $\{\mathbf{r}\}$ and $\{\mathbf{p}\}$ representations, respectively. At some later time t, the dynamical state of the system is given by the vector

$$|\Psi\rangle = \mathrm{e}^{-\mathrm{i}Ht/\hbar} |\psi\rangle,$$

where $H \equiv \mathbf{p}^2/2m$ is the Hamiltonian of the free particle. The momentum being a constant of the motion, its average value remains constant in time; the same holds true for the group velocity

$$\mathbf{v} = \frac{\langle \mathbf{p} \rangle}{m}.$$

One knows (Ch. VI, § 3) that the spreading of the wave packet remains negligible if t is sufficiently small. We intend to make this statement more precise and to show that when these conditions of negligible spreading are fulfilled, the wave packet propagates practically without distortion, and that it may, to a very good approximation, be written

$$\psi(\boldsymbol{r} - \boldsymbol{v}t).$$

This approximate wave function represents the vector

$$|\overline{\Psi}\rangle \equiv \mathrm{e}^{-\mathrm{i}(\boldsymbol{p}\cdot\boldsymbol{v})t/\hbar}\,|\psi\rangle$$

as can be easily shown either by making use of and generalizing the property (VIII.16), or by mere inspection of the corresponding wave function in the $\{\boldsymbol{p}\}$ representation. The approximation is better the closer to unity the probability for the system to be in the state $|\overline{\Psi}\rangle$; in other words, one must have

$$1 - |\langle\Psi|\overline{\Psi}\rangle|^2 \ll 1.$$

Replacing $|\Psi\rangle$ and $|\overline{\Psi}\rangle$ by the expressions given above, one finds

$$|\langle\Psi|\overline{\Psi}\rangle| = \left|\langle\psi|\,\exp\left(\frac{\mathrm{i}}{\hbar}\,\frac{(\boldsymbol{p}-m\boldsymbol{v})^2}{2m}\,t\right)|\psi\rangle\right|.$$

The matrix element of the right-hand side is particularly simple to evaluate in the $\{\boldsymbol{p}\}$ representation; we have

$$\langle\psi|\,\exp\left(\frac{\mathrm{i}}{\hbar}\,\frac{(\boldsymbol{p}-m\boldsymbol{v})^2}{2m}\,t\right)|\psi\rangle = \int |\varphi(\boldsymbol{p})|^2\,\mathrm{e}^{\mathrm{i}(\boldsymbol{p}-m\boldsymbol{v})^2 t/2m\hbar}\,\mathrm{d}\boldsymbol{p}.$$

If one assumes that the wave packet $\varphi(\boldsymbol{p})$ is of the type described in Fig. IV.1, it is a function exhibiting a very pronounced peak of linear dimensions Δp around the mean value $\boldsymbol{p} = m\boldsymbol{v}$, Δp being the magnitude of the vector $\Delta\boldsymbol{p}$ representing the root-mean-square deviation of the momentum of the particle. Under this hypothesis, the exponential of the right-hand side stays very close to unity over the entire region of this peak as long as

$$(\Delta p)^2 t/2m\hbar \ll 1$$

or else

$$\frac{\Delta p}{m}\,t \ll \frac{2\hbar}{\Delta p}.$$

The approximation is therefore quite justified when this condition is fulfilled. Now this condition of validity is just the condition

$$\text{spreading} \ll \text{width}$$

which we have determined in the study of the spreading of the wave packet of Chapter VI, § 3 [condition (VI.15)].

19. Other Representations. Representations in which the Energy is Diagonal

The preceding examples show us that the various equations of the Theory take on very different forms according to which representation is adopted. Hence the procedures of calculation can vary greatly from one representation to another.

Among the various representations of the Quantum Theory, some turn out to be particularly useful in the treatment of conservative systems, by reason of the very simple form taken by the Schrödinger equation: they are the representations in which the energy H is diagonal [1]). The basis vectors $|E\alpha\rangle$ of such a representation can be labelled by the eigenvalue E of the energy and by the assembly α of the eigenvalues of other constants of the motion forming with H the basis of observables. The vector $|\psi(t)\rangle$ representing the dynamical state of the system in the Schrödinger "representation" is represented by the "wave function"

$$\psi(E, \alpha; t) \equiv \langle E\alpha|\psi(t)\rangle.$$

The latter satisfies the Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} \psi(E, \alpha; t) = \langle E\alpha|H|\psi(t)\rangle = E\psi(E, \alpha; t).$$

We therefore have simply

$$\psi(E\alpha; t) = \psi(E\alpha; t_0)\, e^{-iE(t-t_0)/\hbar}.$$

It is thus very easy to find the motion of the vector representing the dynamical state of a quantum system if one knows its representation $\{E\alpha\}$ at the initial instant. In practice, the initial position of this vector is often given by its components in some other representation, $\{q\}$ for instance. The solution of its equation of motion is then readily obtained if one knows how to go over from the $\{q\}$ representation to a representation in which H is diagonal. Mathematically, the construction of the unitary matrix of this change of representation is a problem equivalent to the eigenvalue problem of H in the $\{q\}$ representation, in other words, to the solution in this representation of the time-independent Schrödinger equation.

---

[1]) In its original form, the Matrix Mechanics of Born, Heisenberg, and Jordan was a particular formulation of the Quantum Theory in the Heisenberg "representation", expressed in a representation where the energy is diagonal.

IV. QUANTUM STATISTICS

20. Incompletely Known Systems and Statistical Mixtures

When the dynamical state of a quantum system is incompletely known, certain predictions can still be made concerning its behavior if one has recourse to the usual methods of statistics. The discussion on this subject, begun in Chapter V (§ 16), may be easily translated into our new language.

The dynamical state of a quantum system is completely known if one has succeeded in determining precisely the variables of one of the complete sets of compatible variables associated with the system. It can then be represented by a certain vector $|\rangle$. When the information about the system is incomplete, one merely states that the system has certain probabilities $p_1, p_2, \ldots, p_m, \ldots$, of being in the dynamical states represented by the ket vectors $|1\rangle, |2\rangle, \ldots, |m\rangle, \ldots$, respectively. In other words, the dynamical state of the system must no longer be represented by a unique vector, but by a statistical mixture of vectors.

Suppose that we perform the measurement of a given quantity A on the system. The mean value $\langle A\rangle$ of the results of measurement has a probability p_m of being equal to $\langle A\rangle_m = \langle m|A|m\rangle / \langle m|m\rangle$. One can therefore write, assuming the vectors $|1\rangle, |2\rangle, \ldots, |m\rangle, \ldots$ to be normalized to unity,

$$\langle A\rangle = \sum_m p_m \langle m|A|m\rangle. \tag{VIII.62}$$

The mean value of an arbitrary function $F(A)$ is similarly obtained by replacing A by $F(A)$ in this formula; one easily derives therefrom the statistical distribution of the results of measurement of A.

21. The Density Operator [1])

It is especially convenient to describe the statistical mixture which was just described by means of the operator

$$\varrho = \sum_m |m\rangle\, p_m\, \langle m|. \tag{VIII.63}$$

In this expression, the vectors $|m\rangle$ are normalized to unity (but not

[1]) For a more complete presentation of the properties of the density operator, see U. Fano, *Description of States in Quantum Mechanics by Density Matrix and Operator Techniques*, Rev. Mod. Phys. **29** (1957) 74.

necessarily orthogonal) and the quantities p_m have the characteristic properties of statistical weights, namely

$$p_m \geqslant 0, \qquad \sum_m p_m = 1. \tag{VIII.64}$$

The operator ϱ is called *density operator* or *statistical operator*.

The average value of the observable A is the trace of ϱA:

$$\langle A \rangle = \operatorname{Tr} \varrho A. \tag{VIII.65}$$

Indeed

$$\operatorname{Tr} \varrho A = \sum_m p_m \operatorname{Tr} (|m\rangle \langle m|A).$$

To prove the equivalence of equations (VIII.62) and (VIII.65) it suffices to show that

$$\operatorname{Tr} (|m\rangle \langle m|A) = \langle m|A|m\rangle.$$

Now, since the operator $P_m \equiv |m\rangle \langle m|$ is a projector of trace unity [eq. (VII.88)],

$$\begin{aligned} \operatorname{Tr} P_m A &= \operatorname{Tr} P_m{}^2 A = \operatorname{Tr} P_m A P_m = \operatorname{Tr} |m\rangle \langle m|A|m\rangle \langle m| \\ &= \langle m|A|m\rangle \operatorname{Tr} P_m = \langle m|A|m\rangle. \end{aligned}$$

The same proof applied to the special case $A = 1$ yields the normalization condition

$$\operatorname{Tr} \varrho = 1.$$

This argument applies equally well to any function of the observable A and gives

$$\boxed{\langle F(A) \rangle = \operatorname{Tr} \varrho F(A).}$$

Knowing ϱ, it is thus possible to derive the statistical distribution of the results of measurement of A.

Very generally, if P_D is the projector upon the subspace spanned by the eigenvectors of A belonging to the eigenvalues located in a certain domain D of the spectrum of A, the probability w_D that the result of measurement belongs to the domain D is $\sum_m p_m \langle m|P_D|m\rangle$ [cf. eq. (VIII.5)], that is

$$\boxed{w_D = \operatorname{Tr} \varrho P_D.} \tag{VIII.66}$$

In particular, the probability of finding the system in the quantum state represented by the vector $|\chi\rangle$ (of norm unity) is

$$w_\chi = \mathrm{Tr}\,(\varrho|\chi\rangle\langle\chi|) = \langle\chi|\varrho|\chi\rangle. \qquad \text{(VIII.67)}$$

Since it is sufficient to give ϱ to be able to calculate all measurable physical quantities, average values, and statistical distributions of measurements, *we shall henceforth consider identical two statistical mixtures possessing the same density operator*: every quantized statistical mixture is exactly and completely defined by its density operator.

To complete the generalization of the postulates of § 2 to statistical mixtures, one still has to define which density operator represents the dynamical state of the system once the measurement is completed. We restrict ourselves, as in § 2, to the case of an ideal measurement. If the observation carried out on the system indicates that it is in an eigenstate of A belonging to the domain D defined above, the density operator after measurement is, to within a normalization constant, the projection $P_D\varrho P_D$ of the operator ϱ representing the statistical mixture before measurement. The constant is determined by the condition that this operator have a trace equal to unity. It is therefore equal to the inverse of the quantity $\mathrm{Tr}\,P_D\varrho P_D = \mathrm{Tr}\,\varrho P_D = w_D$. The (non-causal) evolution of the density operator in the course of the measurement, therefore, ultimately corresponds to the scheme [1]).

$$\varrho \rightarrow \begin{array}{l}\text{ideal measurement}\\ \text{yielding result } D\end{array} \rightarrow \frac{P_D\,\varrho P_D}{\mathrm{Tr}\,\varrho P_D}$$

22. Evolution in Time of a Statistical Mixture

We start out in the Schrödinger "representation". Suppose that at time t_0 the dynamical state of the system is represented by the mixture of vectors (of norm 1) $|1\rangle_0$, $|2\rangle_0$, ..., $|m\rangle_0$, ... with respective statistical weights $p_1, p_2, \ldots, p_m, \ldots$. Every component of the mixture evolves in time according to the law

$$|m\rangle_t = U(t, t_0)|m\rangle_0$$

[1]) To justify this extension of the postulate of filtering of a wave packet, one must have recourse to the detailed study of the mechanism of measurement in Quantum Mechanics. On this subject, see references cited in footnote p. 157; see also U. Fano, *loc. cit.*

and the system is represented at time t by the mixture of vectors $|1\rangle_t, |2\rangle_t, \ldots, |m\rangle_t, \ldots$ with the same respective statistical weights $p_1, p_2, \ldots, p_m, \ldots$ as the vectors occurring in the initial mixture. $U(t, t_0)$ is the unitary evolution operator defined in § 8.

From this we derive the causal law of motion of the density operator

$$\begin{aligned}\varrho_t &= \sum_m |m\rangle_t \, p_m \, {}_t\langle m| = \sum_m U(t, t_0)|m\rangle_0 \, p_m \, {}_0\langle m|U^\dagger(t, t_0) \\ &= U(t, t_0)\Big(\sum_m |m\rangle_0 \, p_m \, {}_0\langle m|\Big) \, U^\dagger(t, t_0) \\ &= U(t, t_0) \, \varrho_0 \, U^\dagger(t, t_0).\end{aligned}$$

The density operator at time t arises from the initial density operator by the unitary transformation $U(t, t_0)$.

Taking into account the equation of motion of operator U [eq. (VIII.32)] and its Hermitean conjugate, we have

$$i\hbar \frac{d}{dt} \varrho_t = [H, \varrho_t]. \tag{VIII.68}$$

This is the Schrödinger equation of the density operator. Although formally similar to the Heisenberg equation (VIII.40) from which it differs only in the sign of the commutator, it should not be confused with the latter. The quantities which occur in it are operators of the Schrödinger "representation".

One goes over from the Schrödinger "representation" to the Heisenberg "representation" by the unitary transformation $U^\dagger(t, t_0)$. Consequently, the density operator remains stationary in the Heisenberg "representation" ($\varrho_H = \varrho_0$), whereas the physical quantities are represented by observables moving according to the Heisenberg equation (VIII.40).

23. Characteristic Properties of the Density Operator

The density operator ϱ is a *positive definite, Hermitean* operator (cf. Ch. VII, § 8) of *trace equal to unity.*

Indeed, according to the very definition of ϱ [eq. (VIII.63)], for any $|u\rangle$

$$\langle u|\varrho|u\rangle = \sum_m p_m |\langle u|m\rangle|^2 \geqslant 0 \tag{VIII.69}$$

$$\mathrm{Tr}\, \varrho = \sum_m p_m \, \mathrm{Tr}\, (|m\rangle \langle m|) = \sum_m p_m = 1. \tag{VIII.70}$$

In fact, since all the p_m are positive and since (Schwarz inequality)

$$|\langle u|m\rangle|^2 \leqslant \langle u|u\rangle,$$

one has

$$\langle u|\varrho|u\rangle \leqslant \langle u|u\rangle. \tag{VIII.71}$$

In other words, the operator $(1-\varrho)$ is likewise positive definite.

In the general theory of Hilbert space, one shows that a positive definite Hermitean operator whose trace is finite is an observable and that its eigenvalue spectrum is entirely discrete. The eigenvalues of ϱ in fact all lie between 0 and 1.

Conversely, any positive definite Hermitean operator ϱ of trace 1 may be regarded as a density operator. Indeed, it is an observable which may be written

$$\varrho = \sum_n \varpi_n P_n, \tag{VIII.72}$$

an expression in which $\varpi_1, \varpi_2, \ldots, \varpi_n, \ldots$ designate the non-zero eigenvalues, and $P_1, P_2, \ldots, P_n, \ldots$ the projectors upon their respective subspaces. If none of the eigenvalues is degenerate, each P_n is an elementary projector $P_n = |\bar{n}\rangle\langle\bar{n}|$ and

$$\varrho = \sum_n \varpi_n |\bar{n}\rangle\langle\bar{n}|. \tag{VIII.73}$$

Since $\sum_n \varpi_n = \mathrm{Tr}\,\varrho$, and $\varpi_n = \langle\bar{n}|\varrho|\bar{n}\rangle \geqslant 0$, the ϖ_n have the properties of statistical weights:

$$\varpi_n \geqslant 0, \qquad \sum_n \varpi_n = 1.$$

ϱ is therefore the density operator of the mixture formed with the vectors $|\bar{n}\rangle$, each having the statistical weight ϖ_n, respectively [1]). This argument can be readily extended to the case where some of the eigenvalues of ϱ are degenerate.

24. Pure States

The density operator formalism allows the treatment of pure states as special cases of statistical mixtures.

If one knows with certainty that the system is in the pure state

---

[1]) In expression (VIII.73), the $|\bar{n}\rangle$ are mutually orthogonal. On the other hand, the $|m\rangle$ occurring in the definition (VIII.63) do not necessarily have this property.

$|\chi\rangle$, one can represent that state by a statistical mixture having $|\chi\rangle$ (assumed to be of norm unity) as its sole element; its density operator is the projector

$$\varrho_\chi = |\chi\rangle\langle\chi| \tag{VIII.74}$$

and one has:

$$\varrho_\chi^2 = \varrho_\chi. \tag{VIII.75}$$

Conversely, if a density operator is a projector, it represents a pure state; the latter is the state on which the projection is carried out.

One can give two other criteria permitting to recognize whether or not a density operator ϱ represents a pure state:

1) A density operator ϱ can be put in the form of a linear combination of projectors in several different ways: the expression (VIII.63) is not unique. But in order that the operator ϱ defined by eq. (VIII.63) represent a pure state, it is *necessary* (and sufficient) that all the $|m\rangle$ be equal to each other to within a phase. They then represent one and the same dynamical state, which is the pure state desired (Problem VIII.7).

2) Any density operator ϱ — i.e. any positive definite Hermitean operator of trace 1 — possesses the property $\mathrm{Tr}\,\varrho^2 \leqslant 1$. In order that it represent a pure state it is (necessary and) *sufficient* that (Problem VIII.8)

$$\mathrm{Tr}\,\varrho^2 = 1. \tag{VIII.76}$$

In conclusion, *it is always possible to represent the dynamical state of a system by its density operator, whether that state be completely or incompletely known.* In fact, the specification of this operator is sufficient to determine all physically measurable quantities which the Quantum Theory is in a position to furnish; eq. (VIII.66) plays the role which is played by eq. (VIII.5) when one uses vectors to represent the states. This procedure has the advantage of providing a uniform treatment for the pure states and the mixtures. Moreover, *the density operator representing the state of a system is defined in a unique manner*, while the vector representing a pure state is at best defined only to within a phase factor. Furthermore, the definition of the vector mixture representing an incompletely known dynamical state is fraught with a still higher degree of arbitrariness.

25. Classical and Quantum Statistics

In Classical Mechanics, a dynamical state is defined by a point in phase space; a statistical mixture of states is represented by a fluid in phase space whose density $\varrho_{\text{cl.}}$ at a point is equal to the probability of finding the system in the state defined by that point.

There is a remarkable parallelism between the classical density in phase $\varrho_{\text{cl.}}$ and the density operator ϱ of the Quantum Theory. $\varrho_{\text{cl.}}$ is a real, positive quantity whose integral over all phase space is equal to unity:

$$\iint \varrho_{\text{cl.}}\, \mathrm{d}q\, \mathrm{d}\boldsymbol{p} = 1, \tag{VIII.77}$$

while ϱ is a Hermitean operator whose eigenvalues are all positive (positive definite operator), and whose trace equals unity.

From a knowledge of $\varrho_{\text{cl.}}$ at a given instant, one deduces the mean value $\langle A\rangle_{\text{cl.}}$ of any function $A_{\text{cl.}}$ of the dynamical variables q and p, by integrating $\varrho_{\text{cl.}}\, A_{\text{cl.}}$ over all phase space

$$\langle A\rangle_{\text{cl.}} = \iint \varrho_{\text{cl.}}\, A_{\text{cl.}}\, \mathrm{d}q\, \mathrm{d}p. \tag{VIII.78}$$

The evolution in time of $\varrho_{\text{cl.}}$ is given by the equation

$$\frac{\partial \varrho_{\text{cl.}}}{\partial t} = \{H_{\text{cl.}}, \varrho_{\text{cl.}}\} \tag{VIII.79}$$

[this equation is not to be confused with eq. (VIII.42)].

The equations (VIII.78) and (VIII.79) are the classical analogues of eqs. (VIII.65) and (VIII.68), respectively. One goes over from the expressions of Classical Theory to those of the Quantum Theory by replacing ordinary quantities by observables, Poisson brackets by commutators (to within the coefficient $i\hbar$) and *the integration over all phase space by the trace.*

This new expression of the Correspondence Principle turns out to be extremely valuable when one seeks to extend the fundamental results of Classical Statistical Thermodynamics to the quantum domain. Most of the classical arguments can be taken over without change. We shall merely point out the main results here.

The state of a quantum system *in thermodynamic equilibrium at temperature* T is represented by the operator

$$\varrho = N\, \mathrm{e}^{-H/kT}; \tag{VIII.80}$$

H is the Hamiltonian of the system, k is the Boltzmann constant. N is a normalization constant adjusted in such a way that $\mathrm{Tr}\,\varrho = 1$. The various thermodynamic functions of such a system are calculated as in Classical Theory, by means of the *partition function*

$$Z(\mu) = \mathrm{Tr}\, e^{-\mu H}. \tag{VIII.81}$$

Thus the free energy $\mathscr{F}$, the entropy S, and the energy E are respectively given by the following expressions, written with $\mu = 1/kT$:

$$\mathscr{F} = -kT \ln Z, \tag{VIII.82}$$

$$S = k\left(\ln Z - \mu \frac{\partial}{\partial \mu} \ln Z\right), \tag{VIII.83}$$

$$E \equiv \langle H \rangle = -\frac{\partial}{\partial \mu} \ln Z. \tag{VIII.84}$$

More generally, the entropy of a system is given, in accordance with the Correspondence Principle, by the mean value of the operator $-k \ln \varrho$, that is

$$S = -k\,\mathrm{Tr}\,(\varrho \ln \varrho). \tag{VIII.85}$$

From this the equilibrium distribution (VIII.80) is easily derived: it is the distribution corresponding to a given mean value of the energy for which the entropy is a maximum.

EXERCISES AND PROBLEMS

1. Starting from the fundamental postulate on mean values, derive expression (VIII.5), which gives the probability law for the results of measurement of a given quantity.

2. Consider a quantum system possessing a classical analogue in one dimension. Derive from the commutation relation $[q, p] = i\hbar$ that the spectrum of p is entirely non-degenerate and continuous, and extends from $+\infty$ to $-\infty$. The suitably normalized eigenvectors of p form a complete orthonormal system in $\mathscr{E}$. Show that with a suitable choice of the phases of the vectors $|p'\rangle$ of this system, the action of the unitary operator $\exp(i\varpi q/\hbar)$ (ϖ an arbitrary constant) on these vectors yields

$$\exp(i\varpi q/\hbar)\,|p'\rangle = |p' + \varpi\rangle,$$

and that the operator q has as its matrix element:

$$\langle p'|q|p''\rangle = i\hbar\,\delta'(p' - p'').$$

Solve the eigenvalue problem of q in this representation.

3. The derivative of an operator $A(\xi)$ depending explicitly on a continuous parameter ξ is by definition

$$\frac{\mathrm{d}A}{\mathrm{d}\xi} = \lim_{\varepsilon \to 0} \frac{A(\xi+\varepsilon) - A(\xi)}{\varepsilon}.$$

Show that:

1) if $A(\xi)$ is a function of an observable or of several commuting observables, its derivative is obtained by means of the ordinary rules of differentiation. In particular, if O is an observable:

$$\frac{\mathrm{d}}{\mathrm{d}\xi}(\mathrm{e}^{\mathrm{i}O\xi}) = \mathrm{i}\,O\,\mathrm{e}^{\mathrm{i}O\xi};$$

2) if two operators are differentiable, then

$$\frac{\mathrm{d}}{\mathrm{d}\xi}(AB) = \frac{\mathrm{d}A}{\mathrm{d}\xi}B + A\frac{\mathrm{d}B}{\mathrm{d}\xi}.$$

In particular:

$$\frac{\mathrm{d}}{\mathrm{d}\xi}A^2 = \frac{\mathrm{d}A}{\mathrm{d}\xi}A + A\frac{\mathrm{d}A}{\mathrm{d}\xi};$$

3) if A is differentiable and possesses an inverse, one has

$$\frac{\mathrm{d}}{\mathrm{d}\xi}A^{-1} = -A^{-1}\frac{\mathrm{d}A}{\mathrm{d}\xi}A^{-1}.$$

4. Show that the operator $B(t)$ defined by the expression

$$B(t) = \mathrm{e}^{\mathrm{i}At}B_0\,\mathrm{e}^{-\mathrm{i}At},$$

where A and B_0 are operators independent of t, is a solution of the integral equation

$$B(t) = B_0 + \mathrm{i}[A, \int_0^t B(\tau)\,\mathrm{d}\tau].$$

Solving this equation by iteration, one obtains the expansion of $B(t)$ in a power series of t. From this, derive the identity between operators

$$\mathrm{e}^{\mathrm{i}A}B\,\mathrm{e}^{-\mathrm{i}A} = B + \mathrm{i}[A,B] + \frac{\mathrm{i}^2}{2!}[A,[A,B]] + \ldots + \frac{\mathrm{i}^n}{n!}\overbrace{[A,[A,\ldots[A,[A,B]]\ldots]]}^{n \text{ brackets}}.$$

N.B. We agree to regard $[A, B]$ as the operator resulting from the action of A upon B, and to denote this operator by $A\{B\}$; $A^n\{B\}$ then denotes the action of A repeated n times; following this notation

$$A^0\{B\} \equiv B, \qquad A\{B\} \equiv [A, B], \qquad A^2\{B\} \equiv [A, [A, B]], \text{ etc.}$$

and the identity is simply written

$$\mathrm{e}^{\mathrm{i}A}B\,\mathrm{e}^{-\mathrm{i}A} = \sum_{n=0}^{\infty}\frac{\mathrm{i}^n}{n!}A^n\{B\}.$$

5. Let $A(\xi)$ be an operator depending upon the continuous parameter ξ, $\mathrm{d}A/\mathrm{d}\xi$ its derivative with respect to ξ. Show the identity between operators

$$\mathrm{e}^{-\mathrm{i}A}\frac{\mathrm{d}}{\mathrm{d}\xi}(\mathrm{e}^{\mathrm{i}A}) = \mathrm{i}\sum_{n=0}^{\infty}\frac{(-\mathrm{i})^n}{(n+1)!}A^n\left\{\frac{\mathrm{d}A}{\mathrm{d}\xi}\right\}$$

(notation as in note of Problem 4).

6. If the operator $U(t)$, differentiable with respect to t, is unitary, the operator

$$H(t) \equiv \mathrm{i}\hbar\frac{\mathrm{d}U}{\mathrm{d}t}U^{\dagger}$$

is necessarily Hermitean.

Conversely, if $U(t)$ satisfies the equation

$$\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}t}U(t) = HU,$$

where H is a Hermitean operator possibly depending upon t, $U^{\dagger}U$ is independent of t, whereas $UU^{\dagger}$ is a solution of the equation

$$\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}t}UU^{\dagger} = [H, UU^{\dagger}].$$

In particular, if U is unitary when $t = t_0$, it remains so for all values of t.

7. Let $|1\rangle, |2\rangle, \ldots, |M\rangle$ be a sequence of vectors of norm 1 but not necessarily orthogonal. Show that the necessary and sufficient condition that the density operator $\varrho = \frac{1}{M}\sum_{i=1}^{M}|i\rangle\langle i|$ represent a pure state is that these M vectors are equal to within a phase.

8. Show that for the operator ϱ, Hermitean definite of trace 1, to represent a pure state, it is necessary and sufficient that $\mathrm{Tr}\,\varrho^2 = 1$.

Wisdom hath builded her house,
She hath hewn out her seven pillars:
She hath killed her beasts;
She hath mingled her wine;
She hath also furnished her table.
She hath sent forth her maidens:
She crieth upon the highest places
of the city,
Whoso is simple, let him turn in
hither!

(*Proverbs* ix: 1–4.)

PART TWO

SIMPLE SYSTEMS

CHAPTER IX

SOLUTION OF THE SCHRÖDINGER EQUATION BY SEPARATION OF VARIABLES. CENTRAL POTENTIAL

1. Introduction

The study of a physical system essentially consists in solving its time-independent Schrödinger equation. In particular, this eigenvalue equation enters directly in the two types of problems most frequently encountered in quantum physics, namely:

(*i*) the determination of the energy levels of bound states: they are the eigenvalues of the discrete spectrum of the Hamiltonian;

(*ii*) the determination of collision cross sections: as we shall prove further on (Ch. X) they are derived from the asymptotic form of the eigenfunctions for unbound states.

In Wave Mechanics, the Schrödinger equation is a second-order partial differential equation. For a one-dimensional system, the latter reduces to a differential equation; the study of the eigenvalue problem in this simple case was already made (Ch. III). The problem is generally much more difficult when the system has a larger number of dimensions. However, the symmetry properties the Hamiltonian may possess can facilitate its solution. In particular, it may happen that an appropriate change of variables leads to a partial differential equation whose variables are separable. The eigenvalue problem then splits into several simpler eigenvalue problems involving a smaller number of dimensions.

This occurs for a particle in a central potential, that is to say, in a potential depending only upon the distance r of the particle from a center of force, and not upon the direction of the vector $\boldsymbol{r}$ connecting that center with the particle. Since the Hamiltonian has spherical symmetry, the variables separate completely when one treats the problem in spherical polar coordinates. The solution of the Schrödinger equation, after separation of the angular variables, reduces to the solution of a differential equation involving only the radial variable, an equation which can always be integrated numerically.

The major portion of this chapter is devoted to the solution of the

Schrödinger equation of a particle in a central potential. The general treatment is given in Sec. I. Section II deals with the case of the free particle, and the particle in a "square" central potential.

Another very simple example of the separation of variables forms the subject of Sec. III, namely the center-of-mass motion of a system of particles. As in Classical Mechanics, this motion can be separated from the relative motion whenever the interaction depends only upon the relative positions of the particles with respect to each other.

I. PARTICLE IN A CENTRAL POTENTIAL. GENERAL TREATMENT

2. Expression of the Hamiltonian in Spherical Polar Coordinates

In this section, we intend to study the Schrödinger equation of a particle of mass m in a central potential $V(r)$. If $\boldsymbol{p}$ is the momentum of the particle, and $\boldsymbol{r}$ its position vector, the Hamiltonian is

$$H = \frac{\boldsymbol{p}^2}{2m} + V(r) \tag{IX.1}$$

and the time-independent Schrödinger equation is written

$$H\psi(\boldsymbol{r}) \equiv \left[-\frac{\hbar^2}{2m} \triangle + V(r) \right] \psi(\boldsymbol{r}) = E\psi(\boldsymbol{r}). \tag{IX.2}$$

Since the Hamiltonian has spherical symmetry, we study the problem in spherical polar coordinates.

The polar axis we choose is the z axis according to custom, and the cartesian coordinates (x, y, z) are given as functions of the polar coordinates (r, θ, φ) by the formulae (cf. Fig. IX.1)

$$\begin{aligned} x &= r \sin\theta \cos\varphi \\ y &= r \sin\theta \sin\varphi \\ z &= r \cos\theta. \end{aligned} \tag{I}$$

The expression of the potential energy V as a function of the polar coordinates is already given; we are to find the expression for the kinetic energy $\boldsymbol{p}^2/2m$, in other words to express in polar coordinates the differential operator:

$$-\frac{\hbar^2}{2m} \triangle \equiv -\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right).$$

This can be done directly by means of the transformation formulae (I) by the usual techniques of the differential calculus. The calculation is lengthy, but not particularly difficult; it will not be given here.

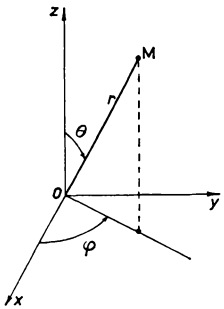


Fig. IX.1. Polar and cartesian coordinates.

However, in order to better grasp the physical significance of the result, we shall seek to express the kinetic energy $\mathbf{p}^2/2m$, not as function of the differential operators $\partial/\partial r$, $\partial/\partial\theta$, $\partial/\partial\varphi$ themselves, but as function of Hermitean operators constructed with these operators whose physical interpretation is more apparent.

Hence, rather than using the differential operator $\partial/\partial\varphi$ directly, it is more advisable to use the component along the z axis of the angular momentum which, according to eq. (V.49), has the explicit form:

$$l_z \equiv x\,p_y - y\,p_x = \frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial\varphi}. \tag{IX.3}$$

Since $V(r)$ does not depend upon φ, it is clear that l_z commutes with the potential energy. However, l_z also commutes with the kinetic energy $\mathbf{p}^2/2m$ as one can easily verify by using the definition of l_z itself, and the fundamental commutation relations (Problem IX.4) [1])

$$[r_i, p_j] = \mathrm{i}\hbar\delta_{ij}. \tag{IX.4}$$

Thus l_z commutes with the Hamiltonian H. Taking Ox, Oy as polar axes, one arrives at the same conclusions for l_x, l_y. Hence *the three components* l_x, l_y, l_z *of the angular momentum*:

$$\boldsymbol{l} \equiv \boldsymbol{r}\times\boldsymbol{p} = \frac{\hbar}{\mathrm{i}}(\boldsymbol{r}\times\boldsymbol{\nabla}) \tag{IX.5}$$

---

[1]) Henceforth we use the indices $i, j = 1, 2$, or 3 to label the components of the vectors along the axes Ox, Oy, or Oz, respectively; thus $r_1 \equiv x$, $p_1 \equiv p_x$, etc.

commute with the Hamiltonian. For this reason, we shall use these operators rather than the differential operators $\partial/\partial\theta$, $\partial/\partial\varphi$ themselves.

In the same spirit, we use the radial momentum

$$p_r \equiv \frac{\hbar}{\mathrm{i}} \frac{1}{r} \frac{\partial}{\partial r} r = \frac{\hbar}{\mathrm{i}} \left(\frac{\partial}{\partial r} + \frac{1}{r} \right) \tag{IX.6}$$

rather than the operator $(\hbar/\mathrm{i})\,(\partial/\partial r)$ which is not Hermitean (cf. Problem IX.1).

To establish this Hermitean property of p_r, let us examine under what condition the expression $\langle \psi, p_r \psi \rangle$, in which $\psi(\boldsymbol{r})$ is any square-integrable function, is real. One must have

$$0 = \langle \psi, p_r \psi \rangle - \langle \psi, p_r \psi \rangle^* \equiv \int [\psi^* (p_r \psi) - (p_r \psi)^* \psi] \, \mathrm{d}\boldsymbol{r}$$
$$= \frac{\hbar}{\mathrm{i}} \int_0^{\pi} \sin\theta \, \mathrm{d}\theta \int_0^{2\pi} \mathrm{d}\varphi \int_0^{\infty} \left[\frac{\partial}{\partial r} |r\psi|^2 \right] \mathrm{d}r.$$

Since $r\psi$ vanishes as $r \to \infty$, the integral with respect to r is equal to its value taken at the origin. The operator p_r is therefore Hermitean only if one restricts oneself to square-integrable functions subject to the condition [1]):

$$\lim_{r \to 0} r\,\psi(\boldsymbol{r}) = 0. \tag{IX.7}$$

It follows from the definition that p_r commutes with any function of θ and φ as well as with the three components of $\boldsymbol{l}$, but that, on the other hand,

$$[r, p_r] = \mathrm{i}\hbar. \tag{IX.8}$$

One has the operator identity

$$\boxed{\boldsymbol{p}^2 = p_r^2 + \frac{\boldsymbol{l}^2}{r^2} \qquad (r \neq 0)} \tag{IX.9}$$

according to which the action of the operators p^2 and $p_r^2 + (\boldsymbol{l}^2/r^2)$ on

[1]) p_r is Hermitean but is not an observable. Whatever the constant ϖ, the solution of the differential equation

$$p_r f(r) \equiv \frac{\hbar}{\mathrm{i}} \frac{1}{r} \frac{\mathrm{d}}{\mathrm{d}r} [r f(r)] = \varpi f(r)$$

is, to within a constant, equal to $\exp(\mathrm{i}\varpi r/\hbar)/r$; it never satisfies condition (IX.7); the eigenvalue problem of p_r has therefore no solution.

a function $\psi(r)$ yields the same result for all non-zero values of r.

To prove this, we use for the calculation of $\boldsymbol{l}^2$ the identity

$$(\boldsymbol{A}\times\boldsymbol{B})^2 = A^2B^2 - (\boldsymbol{A}\cdot\boldsymbol{B})^2,$$

by substituting for the vectors $\boldsymbol{A}$ and $\boldsymbol{B}$ the operators $\boldsymbol{r}$ and $\boldsymbol{p}$. However, since the components of $\boldsymbol{r}$ and $\boldsymbol{p}$ do not necessarily commute with each other, the above identity holds true only if one preserves the order of the operators, namely

$$\boldsymbol{l}^2 \equiv (\boldsymbol{r}\times\boldsymbol{p})\cdot(\boldsymbol{r}\times\boldsymbol{p}) = \sum_{i,j}(r_i p_j r_i p_j - r_i p_j r_j p_i).$$

Repeated application of the commutation relations (IX.4) allows us to rewrite this identity

$$\boldsymbol{l}^2 = r^2\boldsymbol{p}^2 - (\boldsymbol{r}\cdot\boldsymbol{p})^2 + \mathrm{i}\hbar(\boldsymbol{r}\cdot\boldsymbol{p}). \tag{IX.10}$$

However, since $r = (x^2+y^2+z^2)^{\frac{1}{2}}$,

$$x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y} + z\frac{\partial}{\partial z} = r\frac{\partial}{\partial r},$$

hence

$$\boldsymbol{r}\cdot\boldsymbol{p} = r\frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial r} = r\,p_r + \mathrm{i}\hbar$$

and, taking into account the commutation relation (IX.8),

$$(\boldsymbol{r}\cdot\boldsymbol{p})^2 - \mathrm{i}\hbar(\boldsymbol{r}\cdot\boldsymbol{p}) \equiv [(\boldsymbol{r}\cdot\boldsymbol{p}) - \mathrm{i}\hbar](\boldsymbol{r}\cdot\boldsymbol{p}) = rp_r(rp_r + \mathrm{i}\hbar) = r^2p_r^2.$$

The right-hand side of the identity (IX.10) is thus equal to

$$r^2(\boldsymbol{p}^2 - p_r^2).$$

Dividing through by r^2 term by term, one obtains the identity (IX.9), valid everywhere except possibly for $r=0$.

From (IX.9) we obtain the expression for the kinetic energy by dividing term by term by $2m$. We thus obtain the expression for the Hamiltonian in spherical polar coordinates:

$$H = \frac{p_r^2}{2m} + \frac{\boldsymbol{l}^2}{2mr^2} + V(r). \tag{IX.11}$$

Like the energy of the corresponding classical particle, it is the sum

of three terms, the "radial kinetic energy" $p_r^2/2m$, the "kinetic energy of rotation" $\boldsymbol{l}^2/2mr^2$ (mr^2 is the moment of inertia with respect to the origin), and the potential energy $V(r)$.

The direct calculation mentioned at the beginning of the paragraph would evidently have led to the same expression, $\boldsymbol{l}^2$ then appearing in its explicit form

$$\boldsymbol{l}^2 = -\frac{\hbar^2}{\sin^2\theta}\left[\sin\theta\,\frac{\partial}{\partial\theta}\left(\sin\theta\,\frac{\partial}{\partial\theta}\right)+\frac{\partial^2}{\partial\varphi^2}\right]. \tag{IX.12}$$

In conclusion, the Schrödinger equation in spherical polar coordinates is written

$$\left[\frac{p_r^2}{2m}+\frac{\boldsymbol{l}^2}{2mr^2}+V(r)\right]\psi(r,\theta,\varphi)=E\psi(r,\theta,\varphi). \tag{IX.13}$$

However, the solutions of this equation can be retained as solutions of the Schrödinger equation only after examination of their behavior at the origin. In fact, the validity at the origin of expression (IX.11) is not automatically assured for any function upon which operator H may act. We merely state here without proof that equation (IX.13) is equivalent to the Schrödinger equation in all of space, *including the origin*, as long as ψ statisfies condition (IX.7), the condition of hermiticity of p_r.

3. Separation of the Angular Variables. Spherical Harmonics

It is easily seen from their expressions (IX.11) and (IX.12) that H and $\boldsymbol{l}^2$ commute. This was to be anticipated; once H commutes with l_x, l_y, l_z, it commutes with any function of these operators, and with $\boldsymbol{l}^2$ in particular. The observables H and $\boldsymbol{l}^2$ have (at least) one common basis. We are thus led to solve the eigenvalue problem of H in two steps: first, to solve the eigenvalue problem of $\boldsymbol{l}^2$, and second, to search for eigensolutions of $\boldsymbol{l}^2$ satisfying the Schrödinger equation. The particular form of the potential $V(r)$ enters only into the second step of this method of solution.

In the search for a complete system of eigenfunctions of $\boldsymbol{l}^2$, the variable r plays the role of a simple parameter and may be momentarily omitted, since the operator $\boldsymbol{l}^2$ acts exclusively upon the angular variables θ and φ.

$\boldsymbol{l}^2$ commutes with each component of the angular momentum [eq. (V.70)], and in particular with l_z. One shows in function theory

that the eigenfunctions common to the differential operators $\boldsymbol{l}^2$ and l_z, defined respectively by expressions (IX.12) and (IX.3) are the *spherical harmonics* $Y_l^m(\theta, \varphi)$. The main properties of these functions are summarized in Appendix B (§ 10). Their construction will be given in detail when we come to the systematic study of angular momentum in Quantum Mechanics (Ch. XIII). They are labelled by the indices l and m; l can take on all integral positive or zero values, and m all integral values in the interval $(-l, +l)$, limits included. One has

$$\boldsymbol{l}^2 Y_l^m(\theta, \varphi) = l(l+1)\hbar^2\, Y_l^m(\theta, \varphi) \tag{IX.14}$$

$$\left.\begin{aligned} l_z Y_l^m(\theta, \varphi) &= m\hbar\, Y_l^m(\theta, \varphi) \\ (l = 0, 1, 2, \ldots, \infty; \qquad & m = -l, -l+1, \ldots, +l). \end{aligned}\right\} \tag{IX.15}$$

In the function space of square-integrable functions of (θ, φ), i.e. the space of square-integrable functions defined on the unit sphere, the spherical harmonics constitute a complete orthonormal set. One must realize that the scalar product is defined in that case as the integral over the unit sphere, and that the surface element is [1])

$$\mathrm{d}\Omega = \sin\theta\, \mathrm{d}\theta\, \mathrm{d}\varphi;$$

the orthonormality relations thus read

$$\int Y_l^{m*} Y_{l'}^{m'}\, \mathrm{d}\Omega \equiv \int_0^{2\pi} \mathrm{d}\varphi \int_0^{\pi} \sin\theta\, \mathrm{d}\theta\, Y_l^{m*}(\theta, \varphi)\, Y_{l'}^{m'}(\theta, \varphi) = \delta_{ll'}\, \delta_{mm'}. \tag{IX.16}$$

To each pair of quantum numbers (lm) corresponds a single spherical harmonic. Hence, the angular dependence of a function $\psi(r, \theta, \varphi)$ is completely determined if one requires this function to be a common eigenfunction of $\boldsymbol{l}^2$ and l_z corresponding to the respective eigenvalues $l(l+1)\hbar$ and $m\hbar$; $\psi(r, \theta, \varphi)$ is necessarily of the form $f(r)\, Y_l^m(\theta, \varphi)$.

4. The Radial Equation

We shall now tackle the second step of our method of solution of the Schrödinger equation. It consists in forming the eigenfunctions common to the commuting operators H, $\boldsymbol{l}^2$ and l_z. They are the solutions of the Schrödinger equation of the type

$$\psi_l^m(r, \theta, \varphi) = Y_l^m(\theta, \varphi)\, \chi_l(r). \tag{IX.17}$$

---

[1]) See the discussion at the end of Ch. VII, § 13. Similar precautions must be taken when writing the closure relation in the (θ, φ) representation [eq. (B.88)].

Since ψ_l^m is a solution of eq. (IX.13) and Y_l^m is an eigenfunction of $\boldsymbol{l}^2$ [eq. (IX.14)], $\chi_l(r)$ is a solution of the second-order, linear differential equation

$$\left[\frac{p_r^2}{2m} + \frac{l(l+1)\hbar^2}{2mr^2} + V(r) - E\right]\chi_l(r) = 0 \qquad \text{(IX.18)}$$

where $p_r^2 \equiv -\hbar^2 \frac{1}{r}\frac{\mathrm{d}^2}{\mathrm{d}r^2} r$.

It is convenient to set

$$y_l(r) = r\chi_l(r) \qquad \text{(IX.19)}$$

and to replace eq. (IX.18) by the equivalent radial equation

$$\left[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}r^2} + l(l+1)\frac{\hbar^2}{2mr^2} + V(r) - E\right] y_l(r) = 0; \qquad \text{(IX.20)}$$

note the resemblance to the one-dimensional Schrödinger equation. Note also that the norm of ψ_l^m is given, after integration over angles, by the expression

$$\langle \psi_l^m, \psi_l^m \rangle = \int_0^\infty r^2 |\chi_l(r)|^2 \,\mathrm{d}r = \int_0^\infty |y_l(r)|^2 \,\mathrm{d}r \qquad \text{(IX.21)}$$

and that the Hermiticity condition (IX.7) of p_r is equivalent to the condition

$$y_l(0) = 0. \qquad \text{(IX.22)}$$

Not all the solutions of the radial equation (IX.20) should be kept. In order that ψ_l^m be acceptable as an eigenfunction, y_l must satisfy certain conditions of regularity. Indeed, it is necessary:

(*a*) to make sure by examining the behavior of y_l at the origin that ψ_l^m is actually a solution of the Schrödinger equation in all of space, including the origin;

(*b*) that this solution be normalizable (in the general sense defined in Ch. V, § 9).

In order to make these conditions of regularity specific, we shall carry out a systematic study of the behavior of the solutions of eq. (IX.20) at the origin.

We suppose that $V(r)$ is bounded in any finite interval, except possibly at the origin, where it may have a singularity in $1/r$.

These conditions are fulfilled in all cases of practical interest. With this assumption, eq. (IX.20) possesses a "regular" solution R_l (defined to within a constant) which vanishes at the origin as r^{l+1}, and all other solutions of this equation behave at the origin as $(1/r)^l$. [1])

To show this, we assume $V(r)$ to be analytic in the vicinity of the origin, and we seek a particular solution of the form

$$r^s(1 + a_1 r + a_2 r^2 + \ldots).$$

Substituting this expansion, the second derivative which can be derived from it, and the Taylor expansion of $V(r)$ into eq. (IX.20), and writing down that the coefficients of successive powers of r occurring in the left-hand side vanish, one obtains an infinite set of equations whose first (characteristic equation)

$$s(s-1) - l(l+1) = 0$$

yields s, and the following equations yield $a_1, a_2, \ldots$ successively. In the present case, the equation in s has two solutions, $l+1$, and $-l$. If $s = l+1$, the calculation of the coefficients of the series can be carried out to all orders and leads to the "regular" solution R_l. If $s = -l$, one faces an impossibility and the method fails. However, if R_l is a solution of eq. (IX.20), it is easy to see that the function

$$R_l(r) \int^r \frac{dr'}{R_l^2(r')}$$

which behaves as $(1/r)^l$ at the origin, is also a solution of this equation. The general solution is a linear combination of these two particular solutions.

Now, any solution in $(1/r)^l$ must be rejected since it does not satisfy at least one of the conditions (a) and (b).

Indeed, if $l \neq 0$, the integral of the square of its modulus diverges at the origin and, according to eq. (IX.21), the function ψ_l^m constructed with this solution does not belong to Hilbert space [condition (b)]. This divergence at the origin persists when one forms the eigendiffer-

[1]) This result was to be expected since the general solution of eq. (IX.20) in the vicinity of the origin is an approximate solution of the equation

$$\left[\frac{d^2}{dr^2} - \frac{l(l+1)}{r^2}\right] v(r) = 0,$$

an equation whose general solution is $v(r) = ar^{l+1} + br^{-l}$ (a and b are arbitrary constants).

ential of $\psi_l{}^m$. This solution must therefore be rejected in any case, whether the eigenvalue E belongs to the discrete or the continuous spectrum.

The foregoing argument does not apply when $l=0$. But in that case, the corresponding wave function ψ_0 does not satisfy the Schrödinger equation [condition (a)]. In fact, ψ_0 behaves as $(1/r)$ at the origin, and since [eq. (A.12)] $\triangle(1/r) = -4\pi\delta(\mathbf{r})$,

$$(H-E)\,\psi_0 = \frac{2\pi\hbar^2}{m}\,\delta(\mathbf{r}).$$

One must therefore keep only the so-called "regular" solutions, that is, the solutions satisfying the condition at the origin (IX.22). With such a solution we can be sure that the function $\psi_l{}^m$ is a solution of the Schrödinger equation everywhere, including the origin [condition (a)]; moreover, since the normalization integral converges at the origin, the condition that $\psi_l{}^m$ or its eigendifferential belong to Hilbert space [condition (b)] depends solely upon the behavior of this solution at infinity.

Supplemented by the condition at the origin

$$y_l(0) = 0 \tag{IX.22}$$

the radial equation (IX.20) is the Schrödinger equation of a particle in one dimension of mass m, subjected to the potential $V(r) + [l(l+1)\hbar^2/2mr^2]$ in the region $(0, \infty)$, and to an infinitely repulsive potential in the region $(-\infty, 0)$. The solution of the three-dimensional Schrödinger equation is thus reduced to the solution of a one-dimensional Schrödinger equation. The properties stated in Ch. III (property of the Wronskian, asymptotic behavior of solutions, orthogonality relations, etc.) remain valid in that case in spite of the singularity in $l(l+1)/r^2$ of the "equivalent potential" at the origin.

5. Eigensolutions of the Radial Equation. Nature of the Spectrum

The nature of the energy spectrum and of the eigensolutions of the radial equation (IX.20) belonging to a given value of l may be deduced from the asymptotic behavior of the solutions of this equation which are regular at the origin. The discussion of Ch. III, § 10 may be taken over unchanged.

Suppose, for instance, that $V(r)$ approaches zero asymptotically more rapidly than $1/r$:

$$\lim_{r\to\infty} rV(r) = 0.$$

The energy spectrum contains two parts:

(*i*) if $E<0$, the solution which is regular at the origin grows indefinitely in absolute value as $\exp(\varkappa r)$ ($\varkappa=\sqrt{-2mE}/\hbar$), except for certain discrete values $E_l^{(1)}, E_l^{(2)}, \ldots$, for which

$$y_l \underset{r\to\infty}{\sim} e^{-\varkappa r}.$$

These values are the only possible eigenvalues. To each of them corresponds a radial function of finite norm.

(*ii*) if $E>0$, the solution which is regular at the origin oscillates indefinitely according to the law

$$y_l \underset{r\to\infty}{\sim} \sin(kr - \tfrac{1}{2}l\pi + \delta_l) \qquad \left(k = \frac{\sqrt{2mE}}{\hbar}\right).$$

It is acceptable as an eigensolution for any E, and represents an unbound eigenstate. The constant δ_l is called the *phase shift* (the additional term $-\frac{1}{2}l\pi$ has been added so that $\delta_l = 0$ when $V(r) = 0$; cf. § 7 below). δ_l is a very important quantity: it characterizes the asymptotic behavior of the regular solution and, as such, often enters into collision problems (Ch. X).

If V approaches zero as $1/r$ or more slowly (but monotonically) when $r\to\infty$, it is not possible to write the asymptotic forms as simply, but the essential result concerning the nature of the spectrum persists. It is an entirely non-degenerate spectrum consisting of a continuous portion, the half-axis of positive energies, and a (denumerably infinite) set of negative, discrete values.

It remains to be shown that for a given value of l, the ensemble of eigenfunctions $y_l(r)$ thus constructed forms a complete set, in the sense that any square-integrable function of r defined on the half-axis $(0, \infty)$ can be expanded in a series of these eigenfunctions. We shall assume that this actually holds true for all potential shapes we shall consider; otherwise, the Hamiltonian H would not be an observable.

6. Conclusions

In conclusion, we note that the observables H, $\boldsymbol{l}^2$, and l_z form a complete set of commuting observables. The problem of constructing

the eigenfunctions common to H, $\boldsymbol{l}^2$ and l_z amounts to separating in the Schrödinger equation, the angular variables from the radial variable. If one fixes the eigenvalues $l(l+1)\hbar^2$ and $m\hbar$, of $\boldsymbol{l}^2$ and l_z, respectively, these eigenfunctions are of the form

$$\psi_l{}^m = r^{-1} y_l(r)\ Y_l{}^m(\theta, \varphi) \tag{IX.23}$$

where $y_l(r)$ is the solution of the radial equation (IX.20) which vanishes at the origin and remains bounded in all space.

One usually says that such an eigenfunction represents a state of angular momentum l or, more specifically, that the particle has an angular momentum l with a component m along the z axis. We recall that l and m are integers and that

$$l \geqslant 0, \qquad -l \leqslant m \leqslant l.$$

According to the traditional terminology of spectroscopy, l is the *azimuthal quantum number*, and m the *magnetic quantum number*. Tradition also requires that one label the states of lowest angular momenta by particular letters of the alphabet rather than by the value of their azimuthal quantum number: to the values 0, 1, 2, 3, 4, 5, ... of the angular momentum correspond the letters s, p, d, f, g, h, ..., respectively.

The nature of the spectrum of H depends upon the behavior at infinity of $V(r)$. In particular, if $V(r)$ approaches zero (monotonically), the energy spectrum consists of a certain number of discrete, negative values, and of the (continuous) ensemble of positive values.

Each of the values of the *continuous spectrum* is *infinitely degenerate*. Indeed, there exists an eigenfunction of positive energy E for all possible values (lm) of the angular momentum.

The energies of the *discrete spectrum* E_{kl} can be labelled by means of two indices, the azimuthal quantum number l and a radial quantum number k which serves to label the various eigenvalues of the radial equation belonging to a given angular momentum. There exists no *a priori* reason why the radial equations belonging to different values of l should have common eigenvalues: *in general*, the eigenvalues E_{kl} are all distinct, and they are respectively $(2l+1)$-*fold degenerate* since to each of them there correspond as many linearly independent eigenfunctions as there are possible values for the magnetic quantum number:

$$-l, -l+1, \ldots, +l.$$

For some very special shapes of the potential $V(r)$, it may happen that certain ones of these values E_{kl} coincide; in that case, the degeneracy is greater. We shall encounter this type of accidental degeneracy in the treatment of the hydrogen atom (Ch. XI) and of the three-dimensional, isotropic harmonic oscillator (Ch. XII).

II. CENTRAL SQUARE WELL POTENTIAL. FREE PARTICLE

7. Spherical Bessel Functions

If there exist regions of the interval $(0, \infty)$ where the potential $V(r)$ is constant:

$$V(r) = V_0 = \text{const.},$$

the radial equation assumes a particularly simple form, and its general solution is a linear combination of functions well-known in function theory, the spherical Bessel functions.

Let us first of all suppose that $E > V_0$. If one puts

$$k = \frac{\sqrt{2m(E-V_0)}}{\hbar}, \quad \varrho = kr, \tag{IX.24}$$

eq. (IX.20) can be written in the form

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}\varrho^2} + \left(1 - \frac{l(l+1)}{\varrho^2}\right)\right] y_l = 0.$$

Likewise, the radial function $f_l = y_l/r$, considered as a function of φ, is a solution of the "spherical Bessel differential equation":

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}\varrho^2} + \frac{2}{\varrho}\frac{\mathrm{d}}{\mathrm{d}\varrho} + \left(1 - \frac{l(l+1)}{\varrho^2}\right)\right] f_l = 0. \tag{IX.25}$$

The general solution of eq. (IX.25) is a linear combination of two particular solutions. The most commonly used particular solutions are described in Appendix B (§ 6). They are the functions j_l, n_l [1]), $h_l^{(+)}$, $h_l^{(-)}$. j_l is the only solution that is regular at the origin (behaving as ϱ^l); the three others have a pole of order $l+1$ there. j_l and n_l are real functions and behave like standing waves at infinity

$$j_l(\varrho) \underset{\varrho\to\infty}{\sim} \frac{\sin(\varrho - \frac{1}{2}l\pi)}{\varrho}, \qquad n_l(\varrho) \underset{\varrho\to\infty}{\sim} \frac{\cos(\varrho - \frac{1}{2}l\pi)}{\varrho}. \tag{IX.26}$$

[1]) Most authors designate by n_l the same function with the opposite sign.

The functions $h_l^{(+)} \equiv n_l + \mathrm{i}j_l$ and $h_l^{(-)} \equiv n_l - \mathrm{i}j_l$ behave asymptotically as outgoing and incoming waves, respectively:

$$h_l^{(+)} \underset{\varrho\to\infty}{\sim} \frac{\mathrm{e}^{\mathrm{i}(\varrho - \frac{1}{2}l\pi)}}{\varrho}, \qquad h_l^{(-)} \underset{\varrho\to\infty}{\sim} \frac{\mathrm{e}^{-\mathrm{i}(\varrho - \frac{1}{2}l\pi)}}{\varrho}. \tag{IX.27}$$

In the case where $E < V_0$, one puts

$$\varkappa = \frac{\sqrt{2m(V_0 - E)}}{\hbar}, \tag{IX.28}$$

and everything stated thus far remains valid provided we replace k everywhere by $\mathrm{i}\varkappa$. In particular, the asymptotic forms (IX.26) and (IX.27) remain valid. The only radial solution bounded at infinity is the function $h_l^{(+)}(\mathrm{i}\varkappa r)$; it approaches zero exponentially. More precisely, the function $\mathrm{i}^l\, h_l^{(+)}(\mathrm{i}\varkappa r)$ is a real function, equal to the product of a polynomial of degree l in $(1/\varkappa r)$ and $\exp(-\varkappa r)/\varkappa r$, and its asymptotic form is

$$\mathrm{i}^l\, h_l^{(+)}(\mathrm{i}\varkappa r) \underset{r\to\infty}{\sim} \frac{\mathrm{e}^{-\varkappa r}}{\varkappa r}. \tag{IX.29}$$

8. Free Particle. Plane Waves and Free Spherical Waves

The foregoing considerations apply in particular to the free particle. In that case $V(r) = 0$ over the entire interval $(0, \infty)$, and the Hamiltonian reduces to its kinetic energy term:

$$H = \frac{\boldsymbol{p}^2}{2m}.$$

Let us therefore look for eigensolutions common to H, $\boldsymbol{l}^2$ and l_z. A solution of angular momentum (lm) and energy E is a function of the form

$$Y_l^m(\theta, \varphi) f_l(r)$$

in which f_l is the solution of eq. (IX.25) bounded over the entire interval $(0, \infty)$.

If $E < 0$, the only solution bounded at infinity, $h_l^{(+)}(\mathrm{i}\varkappa r)$, has a pole of order $l+1$ at the origin. The eigenvalue problem has no solution; as was to be expected, there exists no eigenstate of negative energy.

If $E > 0$, equation (IX.25) has one and only one solution bounded everywhere: the function $j_l(kr)$. There thus exists an eigensolution of

angular momentum (lm) for every positive value $E=\hbar^2k^2/2m$ of the energy, namely the function

$$Y_l^m(\theta, \varphi)\, j_l(kr). \tag{IX.30}$$

Every eigensolution thus formed can be labelled by the two discrete indices l, m and by the continuous index k which can assume any value in the interval $(0, \infty)$; the ensemble of these spherical waves forms a complete orthonormal set (Problem IX.3).

The collection of plane waves $\exp(i\boldsymbol{k}\cdot\boldsymbol{r})$ forms another complete, orthonormal set of eigenfunctions of the energy of the free particle. They are the eigenfunctions common to the observables p_x, p_y, p_z, that is to say the solutions corresponding to a well-defined value of the momentum $\boldsymbol{p}$. Every plane wave is defined by three continuous parameters, the three components of the vector $\boldsymbol{k}$, which can take on all values between $-\infty$ and $+\infty$.

The wave $\exp(i\boldsymbol{k}\cdot\boldsymbol{r})$ represents a free particle of momentum $\hbar\boldsymbol{k}$ and of energy $E=\hbar^2k^2/2m$. On the other hand, it does not represent a well-defined angular momentum state, just as the spherical wave (IX.30) does not represent a well-defined momentum state. This is not surprising since the three components of the momentum p_x, p_y, p_z do not simultaneously commute with $\boldsymbol{l}^2$ and l_z.

9. Expansion of a Plane Wave in Spherical Harmonics

Every energy eigenvalue of the free particle is infinitely degenerate. Since the spherical waves (IX.30) form a complete set, the denumerable set of spherical waves corresponding to a given value of the wave number k, spans the space of the eigenfunctions of energy $E=\hbar^2k^2/2m$; therefore, the plane wave $\exp(i\boldsymbol{k}\cdot\boldsymbol{r})$ may be expanded in a series of these functions:

$$e^{i\boldsymbol{k}\cdot\boldsymbol{r}} = \sum_{l=0}^{\infty} \sum_{m=-l}^{+l} a_{lm}(\boldsymbol{k})\, Y_l^m(\theta, \varphi)\, j_l(kr). \tag{IX.31}$$

If one chooses the z axis along $\boldsymbol{k}$, this wave can be written in the form $\exp(ikr\cos\theta)$; it is independent of φ and the expansion (IX.31) contains only the terms $m=0$ [1]). Let

$$\varrho = kr, \qquad u = \cos\theta.$$

[1]) Indeed:

$$l_z \exp(ikr\cos\theta) = \frac{\hbar}{i}\frac{\partial}{\partial\varphi}\exp(ikr\cos\theta) = 0.$$

The expansion of the plane wave reduces to an expansion in a series of Legendre polynomials [cf. eq. (B.94)]

$$e^{i\varrho u} = \sum_{l=0}^{\infty} c_l\, j_l(\varrho)\, P_l(u). \tag{IX.32}$$

To determine the coefficients c_l, one may proceed as follows. Differentiating the series (IX.32) term by term with respect to ϱ, we obtain

$$iue^{i\varrho u} = \sum_l c_l \frac{dj_l}{d\varrho} P_l. \tag{IX.33}$$

But this expansion is also written, taking into account the recursion relation (B.78) of the Legendre polynomials,

$$iue^{i\varrho u} = i\sum_l c_l j_l u P_l = i\sum_l \left(\frac{l+1}{2l+3} c_{l+1} j_{l+1} + \frac{l}{2l-1} c_{l-1} j_{l-1}\right) P_l. \tag{IX.34}$$

Equating the coefficients of P_l in the expansions (IX.33) and (IX.34) and using the recursion relations (B.53) and (B.54) of the spherical Bessel functions, one obtains the relations

$$l\left(\frac{1}{2l+1} c_l - \frac{i}{2l-1} c_{l-1}\right) j_{l-1}(\varrho) = (l+1)\left(\frac{1}{2l+1} c_l + \frac{i}{2l+3} c_{l+1}\right) j_{l+1}(\varrho).$$

In order that they be satisfied for any ϱ, it is necessary and sufficient that the expressions between brackets all vanish, or that

$$\frac{1}{2l+3} c_{l+1} = \frac{i}{2l+1} c_l \qquad (l = 0, 1, 2, \ldots, \infty),$$

that is

$$c_l = (2l+1) i^l c_0.$$

The coefficient c_0 is obtained by writing the expansion $\exp(i\varrho u)$ for $\varrho = 0$; since $j_l(0) = \delta_{l0}$, $c_0 = 1$.

In conclusion, the expansion of the plane wave may be written

$$\boxed{e^{ikz} \equiv e^{ikr\cos\theta} = \sum_{l=0}^{\infty} (2l+1)\, i^l\, j_l(kr)\, P_l(\cos\theta).} \tag{IX.35}$$

To obtain this expansion in any polar coordinate system, one notes that the angle θ occurring in the expansion (IX.35) is the angle between

the vectors $\boldsymbol{k}$ and $\boldsymbol{r}$. Let us denote by $\hat{\boldsymbol{k}}$ and $\hat{\boldsymbol{r}}$ the respective angular coordinates of these vectors. According to the addition theorem of spherical harmonics [eq. (B.98)]

$$P_l(\cos\theta) = \frac{4\pi}{2l+1} \sum_{m=-l}^{+l} Y_l^{m*}(\hat{\boldsymbol{k}})\, Y_l^m(\hat{\boldsymbol{r}}).$$

Substitution of this expression in the expansion (IX.35) yields

$$e^{i\boldsymbol{k}\cdot\boldsymbol{r}} = 4\pi \sum_{l=0}^{\infty} \sum_{m=-l}^{+l} i^l j_l(kr)\, Y_l^{m*}(\hat{\boldsymbol{k}})\, Y_l^m(\hat{\boldsymbol{r}}). \qquad \text{(IX.36)}$$

10. Study of the Spherical Square Well

By way of an illustration of the problem of a particle in a central force field, we treat the "square well" (Fig. IX.2)

$$V(r) = \begin{cases} -V_0 & r<a \\ 0 & r>a. \end{cases} \qquad \text{(IX.37)}$$

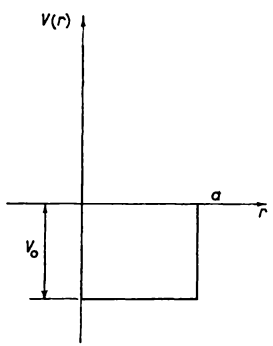


Fig. IX.2. Spherical square well.

The solution of the radial equation is entirely analogous to that of the one-dimensional square well. One knows how to write down the general solution of the Schrödinger equation in each of the two regions $(0, a)$ and (a, ∞): it is a linear combination of spherical Bessel functions. The conditions of regularity at the origin and at infinity, and the continuity condition of the function and of its logarithmic derivative at the point $r=a$ allow us to determine the acceptable solutions.

Let E be the energy of the particle, and put $K=[2m(E+V_0)]^{\frac{1}{2}}/\hbar$. In the inner region $(0<r<a)$ the radial equation reads

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}r^2}+\frac{2}{r}\frac{\mathrm{d}}{\mathrm{d}r}+\left(K^2-\frac{l(l+1)}{r^2}\right)\right] f_l(r)=0.$$

Setting $\varrho=Kr$, one again arrives at equation (IX.25). There is only one solution that is regular at the origin:

$$Aj_l(Kr) \qquad (A \text{ is a normalization constant}).$$

In the outer region $(r>a)$, the Schrödinger equation is that of a free particle. Two cases must be considered according to whether $E \lessgtr 0$.

A) $E<0$. *Discrete spectrum, bound states.*

Let us put $\varkappa=(-2mE)^{\frac{1}{2}}/\hbar$. The only solution bounded at infinity is the exponentially decreasing solution $Bh_l^{(+)}(\mathrm{i}\varkappa r)$ characteristic of a bound state. The continuity condition of the function for $r=a$ fixes the ratio B/A. The continuity of the logarithmic derivative yields

$$\frac{1}{h_l^{(+)}(\mathrm{i}\varkappa r)}\left(\frac{\mathrm{d}}{\mathrm{d}r}h_l^{(+)}(\mathrm{i}\varkappa r)\right)\Bigg|_{r=a}=\frac{1}{j_l(Kr)}\frac{\mathrm{d}}{\mathrm{d}r}j_l(Kr)\Bigg|_{r=a}. \qquad \text{(IX.38)}$$

This condition can be fulfilled only for certain discrete values of E. It determines the energy levels of the bound states of the particle in the well under study. If one deals with s states $(l=0)$, this equation is simply

$$-\varkappa a = Ka\cot Ka. \qquad \text{(IX.38 } a\text{)}$$

This is almost exactly the equation (III.18) for the one-dimensional problem of Ch. III, § 6. The discussion of the number of roots of the equation, and the number of nodes of the solutions can be taken over here without much change. An analogous discussion applies for the levels of higher angular momentum (Problem IX.5).

B) $E>0$. *Continuous spectrum, unbound states.*

Let $k=(2mE)^{\frac{1}{2}}/\hbar$. The general solution of the Schrödinger equation in the outer region is bounded everywhere. It is a linear combination of $j_l(kr)$ and $n_l(kr)$. The continuity conditions at the point $r=a$ fix the coefficients of the linear combination that is acceptable as eigen-

function. To each value of E, therefore, corresponds one and only one wave function (to within a constant).

If we write the outer solution in the form

$$B[\cos \delta_l \, j_l(kr) + \sin \delta_l \, n_l(kr)], \tag{IX.39}$$

the continuity condition of the function at $r=a$ fixes the ratio B/A. δ_l is determined by the continuity condition of the logarithmic derivative

$$\frac{Kj_l'(Ka)}{j_l(Ka)} = k \frac{\cos \delta_l \, j_l'(ka) + \sin \delta_l \, n_l'(ka)}{\cos \delta_l \, j_l(ka) + \sin \delta_l \, n_l(ka)}. \tag{IX.40}$$

It is a real quantity. δ_l is in fact the phase shift of the spherical wave of angular momentum l. Using expressions (IX.26), it is indeed easily verified that the asymptotic form of the solution (IX.39) is

$$B \frac{\sin (kr - \frac{1}{2} l\pi + \delta_l)}{kr}.$$

In the case of an s wave, equation (IX.40) takes on the very simple form

$$K \cot Ka = k \cot (ka + \delta_0). \tag{IX.40 a}$$

III. TWO-BODY PROBLEMS. SEPARATION OF THE CENTER-OF-MASS MOTION

11. Separation of the Center-of-Mass Motion in Classical Mechanics

The study of a system of two particles in Quantum Mechanics is a six-dimensional problem. However, it splits up into two three-dimensional problems, that of a free particle, and that of a particle in a static potential when the two particles feel no interaction other than their mutual interaction, and the latter depends only upon their relative position $\boldsymbol{r} = \boldsymbol{r}_1 - \boldsymbol{r}_2$. Let us denote by m_1, m_2 the masses, by $\boldsymbol{p}_1$, $\boldsymbol{p}_2$ the momenta, and by $\boldsymbol{r}_1$, $\boldsymbol{r}_2$ the respective positions of these two particles. The Hamiltonian of the system under study is of the form

$$H \equiv \frac{\boldsymbol{p}_1^2}{2m_1} + \frac{\boldsymbol{p}_2^2}{2m_2} + V(\boldsymbol{r}_1 - \boldsymbol{r}_2). \tag{IX.41}$$

The treatment consists in separating the motion of the center of mass

from the relative motion, in complete analogy with the corresponding classical treatment.

Let us briefly recall how the classical procedure appears in the Hamiltonian formalism. Let us write

$$\begin{aligned} M &= m_1+m_2, \quad \boldsymbol{R} = \frac{m_1\boldsymbol{r}_1+m_2\boldsymbol{r}_2}{m_1+m_2}, \quad \boldsymbol{P} = \boldsymbol{p}_1+\boldsymbol{p}_2, \\ m &= \frac{m_1m_2}{m_1+m_2}, \quad \boldsymbol{r} = \boldsymbol{r}_1-\boldsymbol{r}_2, \qquad \boldsymbol{p} = \frac{m_2\boldsymbol{p}_1-m_1\boldsymbol{p}_2}{m_1+m_2}. \end{aligned} \tag{II}$$

According to the change of dynamical variables (II), the motion of the two particles can be pictured as the motion of two fictitious particles. One is the *center of mass*, whose position is $\boldsymbol{R}$; its momentum $\boldsymbol{P}$ is the total momentum, and its mass M the total mass of the system. The other is the particle associated with the relative motion; its position $\boldsymbol{r}$ is the relative position of the first particle with respect to the second, and its velocity $\boldsymbol{p}/m$ is equal to its relative velocity $(\boldsymbol{p}_1/m_1)-(\boldsymbol{p}_2/m_2)$; the mass m of this *relative particle* is called the *reduced mass*.

We indicate in passing some noteworthy properties of the transformation (II):

$$m_1m_2 = mM \tag{IX.42a}$$

$$\frac{p_1^2}{2m_1}+\frac{p_2^2}{2m_2} = \frac{p^2}{2m}+\frac{P^2}{2M} \tag{IX.42b}$$

$$m_1r_1^2+m_2r_2^2 = mr^2+MR^2 \tag{IX.42c}$$

$$\boldsymbol{p}_1\cdot\boldsymbol{r}_1+\boldsymbol{p}_2\cdot\boldsymbol{r}_2 = \boldsymbol{p}\cdot\boldsymbol{r}+\boldsymbol{P}\cdot\boldsymbol{R} \tag{IX.42d}$$

$$\boldsymbol{l}_1+\boldsymbol{l}_2 = \boldsymbol{l}+\boldsymbol{L}. \tag{IX.42e}$$

In equation (IX.42*e*), we have introduced the angular momenta l_1, l_2 of the two particles, the angular momentum of the relative particle $\boldsymbol{l}=\boldsymbol{r}\times\boldsymbol{p}$ and that of the center of mass $\boldsymbol{L}=\boldsymbol{R}\times\boldsymbol{P}$.

One readily verifies that this change of variable conserves the Poisson brackets: it is a canonical transformation. Hence the equations of motion of the new variables are the canonical equations obtained from the Hamiltonian function considered as a function of the new variables, namely:

$$H = \frac{\boldsymbol{P}^2}{2M}+\frac{\boldsymbol{p}^2}{2m}+V(\boldsymbol{r}). \tag{IX.43}$$

One finds

$$\dot{\boldsymbol{R}} = \frac{\boldsymbol{P}}{M}, \qquad \dot{\boldsymbol{P}} = 0,$$

$$\dot{\boldsymbol{r}} = \frac{\boldsymbol{p}}{m}, \qquad \dot{\boldsymbol{p}} = -\operatorname{grad} V.$$

The equations of motion of the center of mass and of the relative particle are completely separated. The center-of-mass motion is a uniform rectilinear motion, namely the motion of a free particle of mass M. The motion of the relative particle is that of a particle of mass m in the potential $V(\boldsymbol{r})$.

12. Separation of the Center-of-Mass Motion of a Quantized Two-Particle System

In order to treat the same problem in Quantum Mechanics, one likewise introduces the new dynamical variables $\boldsymbol{r}$, $\boldsymbol{R}$, $\boldsymbol{p}$, and $\boldsymbol{P}$ defined as functions of the old variables by equations (II). The Hamiltonian, given by expression (IX.41) as function of the old variables, takes the form (IX.43) as function of the new variables. In fact, the new variables obey the same commutation relations as if they represented two particles of positions $\boldsymbol{r}$ and $\boldsymbol{R}$ and momenta $\boldsymbol{p}$ and $\boldsymbol{P}$, respectively, the only non-zero commutators being

$$[r_j, p_j] = \mathrm{i}\hbar, \quad [R_j, P_j] = \mathrm{i}\hbar. \qquad (j = x, y, z)$$

All these properties are algebraic properties which can be easily checked starting from equations (II).

One may also verify that eqs. (IX.42) remain valid in the Quantum Theory – including eq. (IX.42*e*) – without the necessity of changing the order of the operators occurring there.

When expressed as a function of the new variables, H is the sum of two terms:

$$H = H_r + H_R,$$

the first of which,

$$H_R = \frac{\boldsymbol{P}^2}{2M},$$

depends only on the center-of-mass variables, and the second,

$$H_r = \frac{\boldsymbol{p}^2}{2m} + V(r),$$

only on the variables of the relative particle. The vectors formed by the tensor product of the eigenvectors of H_r and the eigenvectors of H_R constitute a complete set of eigenvectors of H.

Hence the Schrödinger equation in the representation $\{\boldsymbol{R}, \boldsymbol{r}\}$ reads

$$\left[\left(-\frac{\hbar^2}{2M}\triangle_R\right)+\left(-\frac{\hbar^2}{2m}\triangle_r+V(\boldsymbol{r})\right)\right]\Psi(\boldsymbol{R},\boldsymbol{r})=E\Psi(\boldsymbol{R},\boldsymbol{r}), \quad \text{(IX.44)}$$

where $\triangle_R$ and $\triangle_r$ designate the Laplacians relative to the coordinates $\boldsymbol{R}$ and $\boldsymbol{r}$, respectively [1]). This equation possesses a complete set of eigensolutions of the form

$$\Psi(\boldsymbol{R},\boldsymbol{r})=\Phi(\boldsymbol{R})\,\varphi(\boldsymbol{r}),$$

the functions Φ and φ respectively satisfying the separate equations

$$H_R\,\Phi(\boldsymbol{R})\equiv\left(-\frac{\hbar^2}{2M}\triangle_R\right)\Phi(\boldsymbol{R})=E_R\,\Phi(\boldsymbol{R})$$

$$H_r\,\varphi(\boldsymbol{r})\equiv\left(-\frac{\hbar^2}{2m}\triangle_r+V(\boldsymbol{r})\right)\varphi(\boldsymbol{r})=E_r\,\varphi(\boldsymbol{r}).$$

The energy eigenvalue of the overall system is the sum of the energy eigenvalues of the partial systems:

$$E=E_R+E_r.$$

More generally, if at the initial time t_0 the wave function is a product of two factors, $F(\boldsymbol{R})\,f(\boldsymbol{r})$, this factorization property is conserved in the course of time. The function $F(\boldsymbol{R})$ moves like a wave packet representing a free particle of mass M, and the function $f(\boldsymbol{r})$ like the wave function representing a particle of mass m in the potential $V(\boldsymbol{r})$. [2])

In practice, the solution of our original two-body problem therefore

---

[1]) This equation could just as well have been obtained directly from the Schrödinger equation in the representation $\{\boldsymbol{r}_1, \boldsymbol{r}_2\}$, by making in this partial differential equation the change of variables

$$(\boldsymbol{r}_1, \boldsymbol{r}_2)\rightarrow(\boldsymbol{r}, \boldsymbol{R}).$$

[2]) If $m_1 \ll m_2$, $m \approx m_1$ and $M \approx m_2$ (example: hydrogen atom, m_1 = electron mass, m_2 = proton mass); if $m_1 \approx m_2$, $m \approx \frac{1}{2}m_1$ and $M \approx 2m_1$ (example: the deuterium nucleus, m_1 = proton mass, m_2 = neutron mass).

reduces to that of a single particle in the potential $V(\boldsymbol{r})$, a problem we have learned to solve for the case where this potential is a central potential.

13. Extension to Systems of more than Two Particles

The separation of the center-of-mass motion very generally applies to systems of several particles whenever the interaction potential V depends upon the relative position of the particles and not on their absolute position; in other words, *whenever the interaction is invariant under an overall translation of all particles.*

Indeed, let us consider a quantum system of $(N+1)$ particles whose Hamiltonian H has the stated invariance property. It is always possible to carry out the reduction to the center of mass for two among them, i.e. to replace their dynamical variables by those of their center of mass and their relative particle. This procedure may be continued; having carried out a first reduction with two particles, one can perform this same reduction for the center of mass of these two particles and a third particle. After these two successive changes of variables, the three particles are replaced by: the "relative particle" of particles 1 and 2, the "particle" associated with the relative motion of the center of mass of particles 1 and 2 with respect to particle 3, and the center of mass of all three particles. In general, it is possible, after N such successive changes, to replace the $(N+1)$ particles by N "relative particles" and by the center of mass of the assembly of the $(N+1)$ particles. This reduction to the center of mass of the assembly can be achieved in a variety of ways, either by proceeding step by step, particle by particle – with $\frac{1}{2}(N+1)!$ possible variants – or by splitting the $(N+1)$ particles into two groups of N_1 and N_2 particles, respectively; one then performs the reduction to the center of mass in each of the two groups, and then replaces their centers of mass $\boldsymbol{R}_1$, $\boldsymbol{R}_2$ by the "relative particle" $\boldsymbol{R}_1-\boldsymbol{R}_2$ and the center of mass of the ensemble $(M_1\boldsymbol{R}_1+M_2\boldsymbol{R}_2)/(M_1+M_2)$; or else by splitting up the $(N+1)$ particles into three groups, and so forth.

Denote by $\boldsymbol{r}_i$, $\boldsymbol{p}_i$, and m_i the position, momentum, and the mass of the ith particle, respectively; by $\boldsymbol{R}$, $\boldsymbol{P}$, and M the corresponding quantities associated with the center of mass of the assembly of $(N+1)$ particles:

$$M=\sum_{i=1}^{N+1} m, \qquad \boldsymbol{P}=\sum_{i=1}^{N+1} \boldsymbol{p}_i, \qquad \boldsymbol{R}=\frac{1}{M}\sum_{i=1}^{N+1} m_i \boldsymbol{r}_i;$$

and by $\boldsymbol{\rho}_j$, $\boldsymbol{k}_j$, μ_j $(j=1, 2, \ldots, N)$, the corresponding quantities associated with the jth "relative particle" introduced during one of the reductions to the center of mass described above. Since such

a reduction is the succession of N reductions to the center of mass of two particles, the equalities (IX.42) may be easily extended to give

$$m_1 m_2 \dots m_{N+1} = M\mu_1 \mu_2 \dots \mu_N \tag{IX.45a}$$

$$\sum_{i=1}^{N+1} \frac{p_i^2}{2m_i} = \frac{P^2}{2M} + \sum_{j=1}^{N} \frac{\boldsymbol{k}_j^2}{2\mu_j} \tag{IX.45b}$$

$$\sum_{i=1}^{N+1} m_i r_i^2 = MR^2 + \sum_{j=1}^{N} \mu_j \varrho_j^2 \tag{IX.45c}$$

$$\sum_{i=1}^{N+1} \boldsymbol{p}_i \cdot \boldsymbol{r}_i = \boldsymbol{P} \cdot \boldsymbol{R} + \sum_{j=1}^{N} \boldsymbol{k}_j \cdot \boldsymbol{\rho}_j \tag{IX.45d}$$

$$\sum_{i=1}^{N+1} (\boldsymbol{r}_i \times \boldsymbol{p}_i) = \boldsymbol{R} \times \boldsymbol{P} + \sum_{j=1}^{N} (\boldsymbol{\rho}_j \times \boldsymbol{k}_j). \tag{IX.45e}$$

On the other hand, and for the same reason, the new dynamical variables obey the commutation relations characteristic of the dynamical variables of a quantum system of $(N+1)$ particles. Finally, it is clear that the potential V depends only on the relative coordinates $\boldsymbol{\rho}_1, \boldsymbol{\rho}_2, \dots, \boldsymbol{\rho}_N$, and that the total kinetic energy is the sum of the kinetic energy of the center of mass $P^2/2M$ and the kinetic energies of the relative particles [eq. (IX.45b)]; namely

$$H = \frac{P^2}{2M} + \left[\sum_{j=1}^{N} \frac{\boldsymbol{k}_j^2}{2\mu_j} + V(\boldsymbol{\rho}_1, \dots, \boldsymbol{\rho}_N) \right].$$

Thus the Hamiltonian separates as in the case of two particles, and the solution of the $(N+1)$-body problem reduces to that of an N-body problem.

All the properties which were just stated are independent of the particular way adopted for the reduction to the center of mass. In particular, whatever the choice of the N "relative particles" we introduce, the product of their reduced masses $\mu_1 \mu_2 \dots \mu_N$, the sum of their kinetic energies $\sum_j (\boldsymbol{k}_j^2/2\mu_j)$, and the sum of their angular momenta $\sum_j (\boldsymbol{\rho}_j \times \boldsymbol{k}_j)$ remain unchanged [eqs. (IX.45a), (IX.45b) and (IX.45e)] (Problem IX.7).

EXERCISES AND PROBLEMS

1. Show that the Hermitean radial momentum p_r defined by eq. (IX.6) satisfies the equation

$$p_r = \tfrac{1}{2} \left[\frac{\boldsymbol{r}}{r} \cdot \boldsymbol{p} + \boldsymbol{p} \cdot \frac{\boldsymbol{r}}{r} \right].$$

2. Given a particle in a central potential $V(r)$ having a certain number of

bound states, show that the ground state is necessarily an s state. More generally, show that if there exists a bound state of angular momentum L, there exists at least one bound state corresponding to each of the values l of the angular momentum such that $l < L$, and that, E_l being the lowest energy level one can obtain with angular momentum l, one necessarily has

$$E_0 < E_1 < \ldots < E_L.$$

3. Write the orthogonality relation and closure relation explicitly, and show that the eigenfunctions of the free particle

$$k\left(\frac{2}{\pi}\right)^{\frac{1}{2}} Y_l^m(\theta, \varphi)\, j_l(kr)$$

depending upon the continuous index $k (0 < k < \infty)$ and the integral indices l and m $(l \geqslant 0, -l \leqslant m \leqslant l)$, form an orthonormal and complete set.

For this purpose, derive the relation

$$\int_0^\infty j_l(kr)\, j_l(k'r)\, r^2\, \mathrm{d}r = \tfrac{1}{2}\frac{\pi}{k^2}\, \delta(k-k')$$

being aware that the function $\delta(\boldsymbol{r} - \boldsymbol{r}') \equiv \delta(x - x')\, \delta(y - y')\, \delta(z - z')$ is expressed in polar coordinates by:

$$\delta(\boldsymbol{r}-\boldsymbol{r}') = [(r^2 \sin\theta)^{-1}\, \delta(r-r')\, \delta(\theta-\theta')\, \delta(\varphi-\varphi')].$$

Show that if (k, θ_k, φ_k) are the polar coordinates of the vector $\boldsymbol{k}$,

$$\int \exp(-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r})\, Y_l^m(\theta, \varphi)\, j_l(k'r)\, \mathrm{d}\boldsymbol{r} = \frac{2\pi^2}{k^2}\, (-\mathrm{i})^l\, Y_l^m(\theta_k, \varphi_k)\, \delta(k-k').$$

4. Calculate the commutators of each of the components of $\boldsymbol{r}$ and of $\boldsymbol{p}$ with the component $(\boldsymbol{u}\cdot\boldsymbol{l})$ of the angular momentum $\boldsymbol{l} = \boldsymbol{r} \times \boldsymbol{p}$ along the unit vector $\boldsymbol{u}$. Show that they can be written in the abbreviated form

$$[(\boldsymbol{u}\cdot\boldsymbol{l}), \boldsymbol{p}] = \frac{\hbar}{\mathrm{i}}\, (\boldsymbol{u}\times\boldsymbol{p}), \qquad [(\boldsymbol{u}\cdot\boldsymbol{l}), \boldsymbol{r}] = \frac{\hbar}{\mathrm{i}}\, (\boldsymbol{u}\times\boldsymbol{r}).$$

From this, show that every component of $\boldsymbol{l}$ commutes with the scalar quantities $\boldsymbol{p}^2$, $\boldsymbol{r}^2$, and $\boldsymbol{r}\cdot\boldsymbol{p}$.

5. What relation must the characteristic parameters V_0 and a of the square well of § 10 satisfy in order that there be no bound s state? That there be a given number of bound s states? Is there a connection between the number of bound s states, and the number of nodes of the radial function belonging to zero energy? Answer the same questions for an arbitrary l.

6. Let us treat the radial equation of a particle in a central potential $V(r)$ by the WKB-method. In order that the method be justified, it is necessary not only that the variation of $V(r)$ over a region of the order of the wavelength be small, but also that $l \gg 1$. One notes empirically — and one can give some theoretical justification for this (cf. Langer, *loc. cit.*, footnote p. 231) -- that the method can also yield good results for small values of l if one replaces $l(l+1)$ by $(l+\frac{1}{2})^2$ in the centrifugal-barrier term $l(l+1)/r^2$ of the radial equation. Show that if this modification of the radial equation is carried out beforehand, the WKB method gives correctly, for any l:

(*i*) the asymptotic form $\sin(kr - \frac{1}{2}l\pi)/r$ of the free spherical wave $[V(r) = 0]$;

(*ii*) the spectrum of the hydrogen atom $[V(r) = -e^2/r]$;

(*iii*) the spectrum of the isotropic harmonic oscillator $[V(r) = \frac{1}{2}m\omega^2 r^2]$.

(The rigorous solutions of the eigenvalue problems (*ii*) and (*iii*) are given in Chapters XI and XII, respectively.)

7. One performs on a quantum system of $(N+1)$ particles the separation of the center-of-mass motion in two different ways. The position vectors $\boldsymbol{\rho}_1', \ldots, \boldsymbol{\rho}_N'$, of the N "relative particles" introduced in the second case are deduced from the position vectors $\boldsymbol{\rho}_1, \ldots, \boldsymbol{\rho}_N$ of the N "relative particles" introduced in the first case by a linear transformation:

$$\boldsymbol{\rho}_j' = \sum_{k=1}^{N} A_{jk}\, \boldsymbol{\rho}_k.$$

Let $\mu_1, \ldots, \mu_N$ and $\mu_1', \ldots, \mu_N'$ be the reduced masses of these relative particles. Show that the $(N \times N)$ matrix having matrix elements

$$U_{jk} = \mu_j'^{\frac{1}{2}}\, A_{jk}\, \mu_k^{-\frac{1}{2}}$$

is an orthogonal matrix.

8. Consider the two-dimensional Schrödinger equation for the case where the potential energy $V(r)$ depends only upon the radial variable ($x = r\cos\theta$, $y = r\sin\theta$). Prove the identity

$$\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} = \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2}.$$

Deduce from this that there is a complete set of eigenfunctions of the form

$$\psi(r, \theta) = f(r) \exp(il\theta)$$

whose radial part is that solution of the equation

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}r^2} + \frac{1}{r}\frac{\mathrm{d}}{\mathrm{d}r} - \frac{l^2}{r^2} + \frac{2m}{\hbar^2}[E - V(r)]\right] f(r) \equiv 0$$

which vanishes at $r = 0$.

N.B. — If $V(r) = 0$, this regular solution is the Bessel function $J_{|l|}(kr)$, with $k = (2mE)^{\frac{1}{2}}/\hbar$.

CHAPTER X

SCATTERING PROBLEMS. CENTRAL POTENTIAL AND PHASE SHIFT METHOD

1. Introduction

This chapter is devoted to the elementary concepts concerning collision problems. The results of collision experiments are expressed by means of quantities which are called cross sections and are directly related to the asymptotic behavior of the stationary solutions of the Schrödinger equation. After giving a general definition of cross sections, we devote the major portion of Sec. I to establishing this *connection between cross sections and asymptotic form*, for the simple case of the scattering of a particle by a potential approaching zero sufficiently rapidly at large distances (faster than $1/r$). We then show how the technique of separating the center-of-mass motion permits the extension of this treatment to collisions of two interacting particles.

All the remainder of this chapter is devoted to the scattering of a particle by a *central* potential, and to the method of solution known as the *phase-shift method*. The method is outlined in Sec. II. It is particularly useful when the potential has a finite range; very important properties of the phase shifts, which are particularly easy to demonstrate in this case, are discussed in Sec. III. Resonance phenomena may occur in quantum-mechanical collisions as in any wave-propagation problem; the discussion and interpretation of scattering resonances form the subject of Sec. IV. In the fifth and last section some expressions for phase shifts of current interest are given, in particular two approximate formulae, the Born formula and the Bethe formula, or effective range formula.

I. CROSS SECTIONS AND SCATTERING AMPLITUDES

2. Definition of Cross Sections

Consider a typical scattering experiment. A target is struck by a beam of monoërgic particles. Let J be the magnitude of the incident flux, i.e. the number of incident particles crossing per unit time a

unit surface placed perpendicular to the direction of propagation, and at rest with respect to the target. If P is the number of particles per unit volume in the incident beam, and v the velocity of the incident particles relative to the target,

$$J = Pv.$$

Under the conditions of the experiment, P is so small that one can neglect the mutual interaction of the incident particles; they therefore undergo their collisions independently of each other. By means of appropriate counters one measures the number $\mathcal{N}$ of particles scattered per unit time into the solid angle $\mathrm{d}\Omega$ located in the direction $\Omega \equiv (\theta, \varphi)$. $\mathcal{N}$ is directly proportional to the incident current:

$$\mathcal{N} = J\Sigma(\Omega)\,\mathrm{d}\Omega.$$

The quantity $\Sigma(\Omega)$ which has the dimensions of a surface, is a characteristic parameter of the collision of the particle with the target; it is the scattering cross section of the particle by the target in the direction Ω.

In most practical cases, the target is made up of a large number N of atomic or nuclear scattering centers, and the distances between these atoms or atomic nuclei are sufficiently large with respect to the wavelength of the incident particles so that one may neglect all coherence between the waves scattered by each of them [1]). Each scattering center then acts as if it were alone. Moreover, if the target is sufficiently thin so that one may neglect multiple scattering, $\mathcal{N}$ is directly proportional to N and we have

$$\mathcal{N} = JN\,\sigma(\Omega)\,\mathrm{d}\Omega.$$

The area $\sigma(\Omega)$ is called the scattering cross section of the particle by the scattering center in the direction Ω or, for short, *differential scattering cross section.*

The total number of particles scattered in unit time is obtained by integrating over angles. It is equal to $JN\sigma_{\mathrm{tot.}}$, where

$$\sigma_{\mathrm{tot.}} = \int \sigma(\Omega)\,\mathrm{d}\Omega$$

is the *total scattering cross section.*

[1]) This circumstance is not always realized. Important exceptions to this rule are the diffraction phenomena in crystals: electron diffraction, thermal neutron diffraction, or X-ray diffraction.

In nuclear physics, where the scattering centers have linear dimensions of the order of 10^{-13} to 10^{-12} cm, the cross sections are usually measured in barns or millibarns:

$$1 \text{ barn} = 10^{-24} \text{ cm}^2, \qquad 1 \text{ mb} = 10^{-27} \text{ cm}^2.$$

In the foregoing, we have implicitly assumed that the only possible collisions are elastic collisions, i.e. collisions during which the quantum state of the scatterer does not change and where, *a fortiori*, there is no energy transfer to the internal degrees of freedom of the scatterer. We confine ourselves to this type of collision for the moment. Moreover, rather than treat the scatterer, atom or nucleus, in all its complexity, we represent it by a static potential $V(\boldsymbol{r})$ depending on the coordinate $\boldsymbol{r}$ of the particle.

3. Stationary Scattering Wave

We shall therefore consider the scattering of a particle of mass m by a potential $V(\boldsymbol{r})$. In this chapter we shall limit ourselves to potentials $V(\boldsymbol{r})$ which tend to zero more rapidly than $1/r$ as $r \to \infty$. The scattering by a Coulomb potential will be treated in Chapter XI.

Let E be the energy, and $\boldsymbol{p} = \hbar\boldsymbol{k}$ the initial momentum of the particle. One can relate the cross section $\sigma(\Omega)$ to the solution of the Schrödinger equation

$$\left[-\frac{\hbar^2}{2m} \triangle + V(\boldsymbol{r}) \right] \psi_{\boldsymbol{k}}(\boldsymbol{r}) = E\psi_{\boldsymbol{k}}(\boldsymbol{r})$$

whose behavior at infinity is of the form

$$e^{i\boldsymbol{k}\cdot\boldsymbol{r}} + f(\Omega)\frac{e^{ikr}}{r}. \tag{X.1}$$

We assume [1]) without further discussion that one and only one solution of this type exists for each value of $\boldsymbol{k}$. We shall call this solution $\psi_{\boldsymbol{k}}(\boldsymbol{r})$ the stationary scattering wave, with wave vector $\boldsymbol{k}$.

The two terms of the asymptotic form are easily interpreted

---

[1]) We shall prove it in Sec. II for the particular case where V is a central potential sufficiently regular at the origin.

if one uses the definition (IV.9) of the current density vector [1])

$$\boldsymbol{J}(\boldsymbol{r}) = \frac{\hbar}{2mi}\left[\psi^*(\boldsymbol{r})\,(\boldsymbol{\nabla}\psi(\boldsymbol{r})) - (\boldsymbol{\nabla}\psi(\boldsymbol{r}))^*\,\psi(\boldsymbol{r})\right].$$

The plane wave term $\exp(i\boldsymbol{k}\cdot\boldsymbol{r})$ represents a wave of unit density and of current density $\hbar\boldsymbol{k}/m$. Retaining only the lowest order in $1/r$, the term $f(\Omega)\exp(ikr)/r$ represents a wave of density $|f(\Omega)|^2/r^2$ and of current density directed along the direction Ω *toward increasing* r (outgoing wave) and equal to $(\hbar k/m)(|f(\Omega)|^2/r^2)$. In fact, since the effect of the potential $V(\boldsymbol{r})$ can be neglected in the asymptotic region, one can, according to the classical approximation (cf. Ch. VI, § 4) interpret the term $\exp(i\boldsymbol{k}\cdot\boldsymbol{r})$ as a beam of mono-energetic particles of momentum $\hbar\boldsymbol{k}$ and of density 1, representing the incident beam. The term $f(\Omega)\exp(ikr)/r$ is interpreted as a beam of particles emitted radially from the scattering center and represents the beam of scattered particles.

In accordance with this interpretation, one can calculate the number of particles emitted per unit time into the solid angle $d\Omega$ located in the direction Ω: it is equal to the flux of scattered particles through a spherical surface element of very large radius which subtends the solid angle $(\Omega, \Omega+d\Omega)$, namely $(\hbar k/m)|f(\Omega)|^2 d\Omega$. Upon dividing by the incident flux $J=\hbar k/m$, one obtains the scattering cross section

$$\boxed{\sigma(\Omega) = |f(\Omega)|^2} \qquad \text{(X.2)}$$

$f(\Omega)$ is called the *scattering amplitude.*

4. Representation of the Scattering Phenomenon by a Bundle of Wave Packets [2])

The very intuitive argument given above is incorrect for two reasons.

---

[1]) One can define a Hermitean operator $\boldsymbol{J}(\boldsymbol{r}_0)$ representing the current at the point $\boldsymbol{r}_0$:

$$\boldsymbol{J}(\boldsymbol{r}_0) = \frac{1}{2m}\left[\boldsymbol{p}\delta(\boldsymbol{r}-\boldsymbol{r}_0) + \delta(\boldsymbol{r}-\boldsymbol{r}_0)\boldsymbol{p}\right];$$

the current defined above is the average value of this operator with respect to a given quantum state.

[2]) The treatment of the scattering problems we give in §§ 4, 5, and 6 is essentially taken from an outline of the scattering theory by Chew and Low. The calculation of § 16 has also been borrowed from this work.

In the first place, the current density vector is not simply the sum of the current of the incident plane wave and that of the scattered wave. One must add to these contributions the interference term between $\exp(i\mathbf{k}\cdot\mathbf{r})$ and $f(\Omega)[\exp(ikr)/r]$. In the foregoing argument the interferences between incident and scattered wave were deliberately ignored.

In the second place, the representation of the physical situation by the stationary scattering wave

$$\psi_{\mathbf{k}}(\mathbf{r})\, e^{iEt/\hbar} \tag{X.3}$$

is an idealization. In reality, each particle participating in the scattering must be represented by a wave packet formed by the superposition of stationary waves of the type (X.3) corresponding to wave vectors of *magnitude* and *direction* slightly different from $\mathbf{k}$. This packet is constructed so as to correctly fulfill the initial conditions. Because of the spread in momentum directions, it is bounded transversely such that its transverse dimensions do not exceed those of the diaphragm or diaphragms of the apparatus producing the incident beam. Because of the spread in energy, it is limited longitudinally, and its center moves toward the target along a straight line with a velocity equal to its group velocity $\mathbf{v}=\hbar\mathbf{k}/m$.

We specify each such initial trajectory by the position $\mathbf{b}$ of its point of intersection with the plane (S) passing through the scattering center perpendicularly to the direction of propagation. If we designate by t_0 the instant at which the center of the packet would have passed the plane (S) had its motion not been modified by the presence of the potential, the motion of the center of the packet before collision follows the law

$$\langle\mathbf{r}\rangle = \mathbf{b}+\mathbf{v}(t-t_0) \qquad (t \ll t_0).$$

The incident beam is in reality a beam of wave packets of this type moving parallel to each other at the velocity $\mathbf{v}$, and differing from each other only in the values of the parameters $\mathbf{b}$ and t_0 which fix the motion of their respective centers before collision.

In what follows these characteristic lengths enter into play (Fig. X.1)

$\lambda\!\!\!^{-}=(\hbar/mv)=$ mean wavelength of the incident packet;
$d, l=$ transverse and longitudinal dimensions of the incident packet;
$a=$ extension of the scattering region;
$D=$ distance from the counting instruments to the scattering region.

Through the uncertainty relations, d and l are related, respectively, to the spread in direction and in energy of the incident wave packet.

Since we assume good definition of the incident direction of propagation and energy, we must have

$$\lambda\!\!\!^{-} \ll d \quad \text{and} \quad \lambda\!\!\!^{-} \ll l. \tag{X.4}$$

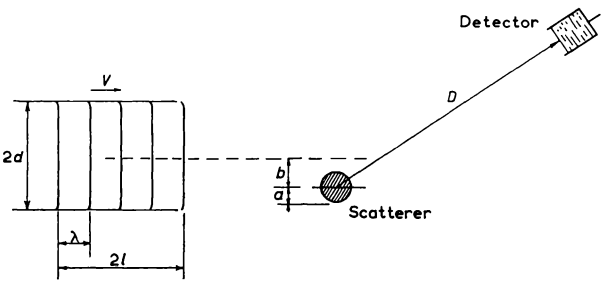


Fig. X.1. Schematic of a scattering experiment and of the characteristic lengths entering into this phenomenon.

In order that the collision phenomenon not depend critically upon the particular form of the wave packet, it is furthermore necessary that its dimensions be very much larger than those of the scattering region, i.e. of the region of space where a non-zero potential prevails, namely

$$a \ll d, l \tag{X.5}$$

($a \approx 10^{-8}$ cm if the scatterer is an atom; $a \approx 10^{-12}$ cm if it is an atomic nucleus).

Now, if the impact parameter is greater than the lateral extension of the packet ($b > d$) the incident wave packet never enters the scattering region and propagates during the entire duration of the phenomenon as a free wave packet. If on the other hand $b < d$, it enters the scattering region at some time $t_1 \approx t_0 - (l/v)$. The collision proper begins at that moment. After a sufficiently long time, the wave packet is again entirely outside the scattering region; it is then generally composed of two additive terms: a transmitted wave packet whose form and law of propagation are essentially the same as those of the incident wave packet, and a wave packet scattered in directions different from the incident direction [1]) (Fig. X.2).

The detection of scattered particles is carried out by placing a suitable counting system (counters, photographic plates, etc.) in a given direction $\Omega = (\theta, \varphi)$ at a distance of order D from the scattering center. This distance must not be too great if one wants

[1]) The phenomenon is analogous to the phenomena of reflection and transmission of one-dimensional waves studied in Ch. III (§§ 3, 6, and 7).

the spreading of the wave packet to remain negligible during the experiment (cf. Ch. VI, § 3):

$$\sqrt{\lambda\!\!\!^{-}D} \ll d, l. \tag{X.6}$$

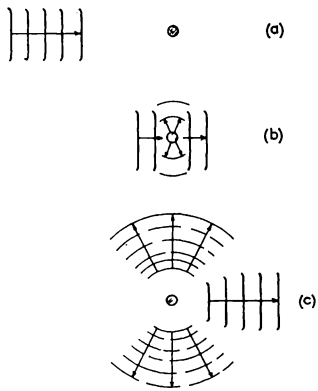


Fig. X.2. Stages of the scattering phenomenon of a wave packet: (*a*) before collision; (*b*) during collision; (*c*) after collision. (Shaded area: region where a non-zero potential prevails.)

With this qualification, the detection must be made sufficiently far away in order that the wave propagation one detects be unaffected by the presence of the scattering center, namely

$$a, \lambda\!\!\!^{-} \ll D \tag{X.7}$$

and that the detector be under no circumstances triggered by the transmitted wave:

$$d \ll D \sin\theta. \tag{X.8}$$

[*N.B.* – The forward-scattered wave ($\theta = 0$) can never be separated from the transmitted wave.]

Combining conditions (X.4)–(X.8), we finally obtain the following inequalities [1]):

$$a, \sqrt{\lambda\!\!\!^{-}D} \ll l \tag{X.9 a}$$

$$a, \sqrt{\lambda\!\!\!^{-}D} \ll d \ll D. \tag{X.9 b}$$

[1]) In atomic or nuclear physics, d is at most equal to the size of the entrance diaphragm of the incident particles, $d \approx 1$ mm; l may be appreciably larger; D is of the order of 1 m. With $a \approx 10^{-8}$ cm and $\lambda\!\!\!^{-} \approx 10^{-8}$ cm, at most, this yields $\sqrt{\lambda\!\!\!^{-}D} \approx 10^{-3}$ cm and $l/a \geq d/a \approx 10^{7}$, $l/\sqrt{\lambda\!\!\!^{-}D} \geq d/\sqrt{\lambda\!\!\!^{-}D} \approx 10^{2}$, $d/D \approx 10^{-3}$. The conditions (X.9) are thus amply fulfilled.

In order that the cross section $\sigma(\Omega)$ actually be given by formula (X.2), it is necessary that the measuring device satisfy conditions (X.9).

Moreover, the wave packets must be sufficiently well defined in direction and in energy so that one can assign to them a well-defined scattering amplitude: $f(\Omega)$ must remain practically constant in *modulus* and in *argument* when the energy varies by an amount of order $\delta E \approx \hbar v/l$, and the angle of incidence by an amount of order $\lambda\!\!\!^{-}/d$ about their respective mean values.

The proof of this result forms the subject of the next two sections.

5. Scattering of a Wave Packet by a Potential

Under these circumstances ($l, d \gg a$), the evolution of the wave packets of the incident beam is practically independent of their particular form. We shall assume that they all have the same shape, each being characterized by the parameters $\boldsymbol{b}, t_0$ determining the motion of its center. Without loss of generality, we take $t_0 = 0$.

To define the form of the incident packet, let us introduce the function $\chi(\boldsymbol{\rho})$ normalized to unity:

$$\int |\chi(\boldsymbol{\rho})|^2 \, \mathrm{d}\boldsymbol{\rho} = 1. \tag{X.10}$$

Denote by $A(\boldsymbol{\varkappa})$ its Fourier transform defined with a norm such that:

$$\chi(\boldsymbol{\rho}) = \int A(\boldsymbol{\varkappa}) \, \mathrm{e}^{\mathrm{i}\boldsymbol{\varkappa}\cdot\boldsymbol{\rho}} \, \mathrm{d}\boldsymbol{\varkappa}. \tag{X.11}$$

By hypothesis, $\chi(\boldsymbol{\rho})$ is a real function whose only appreciable values occur when $\boldsymbol{\rho}$ is located in a region of length l and of width d about the point $\boldsymbol{\rho} = 0$. Likewise, $A(\boldsymbol{\varkappa})$ is real and has appreciable values only in a region of length $(1/l)$ and width $(1/d)$ about the point $\boldsymbol{\varkappa} = 0$. To simplify the discussion, we assume $d \approx l$.

A long time before the collision ($t \ll -l/v$) the wave packet under study $\Psi_{\boldsymbol{b}}(\boldsymbol{r}, t)$ must reduce to the free wave packet $\Phi_{\boldsymbol{b}}(\boldsymbol{r}, t)$ whose center moves according to the law

$$\langle \boldsymbol{r} \rangle = \boldsymbol{b} + \boldsymbol{v}t$$

and whose form at the instant $t = 0$ is given by

$$\begin{aligned} \Phi_{\boldsymbol{b}}(\boldsymbol{r}, 0) &= \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{r}-\boldsymbol{b})} \, \chi(\boldsymbol{r}-\boldsymbol{b}) \\ &= \int A(\boldsymbol{k}'-\boldsymbol{k}) \, \mathrm{e}^{\mathrm{i}\boldsymbol{k}'\cdot(\boldsymbol{r}-\boldsymbol{b})} \, \mathrm{d}\boldsymbol{k}'. \end{aligned}$$

In fact

$$\Phi_{\boldsymbol{b}}(\boldsymbol{r}, t) = \int A(\boldsymbol{k}'-\boldsymbol{k}) \, \mathrm{e}^{\mathrm{i}[\boldsymbol{k}'\cdot(\boldsymbol{r}-\boldsymbol{b}) - (E't/\hbar)]} \, \mathrm{d}\boldsymbol{k}'. \tag{X.12}$$

If one neglects its spreading, this free wave packet is equally well represented by the expression (cf. Ch. VIII, § 18)

$$\Phi_{\boldsymbol{b}}(\boldsymbol{r}, t) \approx \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{b}} \, \mathrm{e}^{\mathrm{i}[\boldsymbol{k}\cdot\boldsymbol{r} - (Et/\hbar)]} \, \chi(\boldsymbol{r} - \boldsymbol{v}t - \boldsymbol{b}) \tag{X.13}$$

obtained by replacing in the integral of the right-hand side of (X.12) the energy $E' = \hbar^2 k^2/2m$ by the first two terms of its expansion in powers of $(\boldsymbol{k}' - \boldsymbol{k})$,

$$E' = E + \hbar \boldsymbol{v} \cdot (\boldsymbol{k}' - \boldsymbol{k}).$$

The wave packet Ψ_b is obtained by substituting the stationary scattering wave $\psi_{\boldsymbol{k}'}(\boldsymbol{r})$ for the plane wave $\exp(i\boldsymbol{k}' \cdot \boldsymbol{r})$ in the integral (X.12):

$$\Psi_b(\boldsymbol{r}, t) \approx \int A(\boldsymbol{k}' - \boldsymbol{k})\, e^{-i\boldsymbol{k}' \cdot \boldsymbol{b}}\, \psi_{\boldsymbol{k}'}(\boldsymbol{r})\, e^{-iE't/\hbar}\, d\boldsymbol{k}'. \tag{X.14}$$

Indeed, this is actually a solution of the Schrödinger equation since it is a superposition of solutions of that equation. It is therefore sufficient to show that it is identical to the free wave packet Φ_b before the collision.

Since the function $A(\boldsymbol{k}' - \boldsymbol{k})$ possesses a pronounced peak about the point $\boldsymbol{k}' = \boldsymbol{k}$, the only important contribution to the integral (X.14) comes from the region surrounding that point. When $t \ll -l/v$, the phase of the integrand varies very rapidly in this region because of the exponential $\exp(iE't/\hbar)$, and the integral is practically zero except for values of r for which this phase is stationary; this can only happen for r of order $v|t|$, in other words for regions of configuration space where $\psi_{\boldsymbol{k}'}(\boldsymbol{r})$ can be replaced by its asymptotic form

$$\psi_{\boldsymbol{k}'}(\boldsymbol{r}) \sim e^{i\boldsymbol{k}' \cdot \boldsymbol{r}} + f_{\boldsymbol{k}'}(\Omega)\, \frac{e^{ik'r}}{r}. \tag{X.15}$$

Substitution of this expression in the integral (X.14) yields

$$\Psi_b(\boldsymbol{r}, t) \underset{t \to -\infty}{\sim} \Phi_b(\boldsymbol{r}, t) + \Psi_b^{(d)}(\boldsymbol{r}, t) \tag{X.16}$$

with

$$\Psi_b^{(d)} = \int A(\boldsymbol{k}' - \boldsymbol{k})\, e^{-i\boldsymbol{k}' \cdot \boldsymbol{b}}\, f_{\boldsymbol{k}'}(\Omega)\, \frac{e^{i[k'r - (E't/\hbar)]}}{r}\, d\boldsymbol{k}'. \tag{X.17}$$

When $t \to -\infty$, the phase of the integrand of the right-hand side of (X.17) cannot be stationary in the region $\boldsymbol{k}' = \boldsymbol{k}$ and the integral $\Psi_b^{(d)}$ is practically zero for any r. The wave packet Ψ_b is thus actually identical to the free wave packet, in this limit.

Consider now the motion of the wave packet Ψ_b in the detection region ($r \gtrsim D$). In this region of space the substitution of its asymptotic form (X.15) for $\psi_{\boldsymbol{k}'}$ is certainly justified. Expression (X.16) therefore holds again.

We assume that the spreads in angle and energy are so small that $f_{\boldsymbol{k}'}(\Omega)$ remains practically constant in the region of extension $1/d \approx 1/l$ surrounding the point $\boldsymbol{k}' = \boldsymbol{k}$, and that we can replace in

the integral (X.17) the modulus of $f_{k'}(\Omega)$ by its value at the point k, and its argument by the first two terms of its expansion:

$$\begin{aligned} &\arg f_{k'}(\Omega) \approx \arg f_k(\Omega) + (k'-k)\cdot s(\Omega) \\ &s(\Omega) = \operatorname{grad}_k [\arg f_k(\Omega)] \qquad (s \ll d, l). \end{aligned} \tag{X.18}$$

Likewise, we replace the arguments of the other factors by the first terms of their expansion

$$k' \approx k + u\cdot(k'-k), \qquad E' \approx E + \hbar v\cdot(k'-k) \tag{X.19}$$

($u \equiv v/v$ = unit vector along the incident direction). Considerations of stationary phase analogous to those for Φ_b [cf. eq. (X.13)] lead us to the equation

$$\Psi_b^{(d)} \approx e^{-ik\cdot b} f_k(\Omega) \frac{e^{i[kr-(Et/\hbar)]}}{r} \chi[u(r-vt)+s-b]. \tag{X.20}$$

Whether $\Psi_b^{(d)}$ can be neglected or not depends upon whether the impact parameter b is larger or smaller than d.

If $b > d$, the argument of χ lies at all times and for all r in the region where this function is negligible; $\Psi_b^{(d)}$ always remains practically zero, and the wave packet propagates as a simple free wave packet.

If $b < d$, i.e. if the incident packet actually enters the scattering region during its motion, χ is appreciable in a spherical shell of thickness of order l on either side of the sphere $r = vt$: $\Psi_b^{(d)}$, which had remained practically zero before the collision, represents an outgoing wave packet of velocity v after the collision, i.e. a packet of spherical waves traveling away from the scatterer with radial velocity v. At the time $t \approx D/v$, the wave $\Psi_b^{(d)}$ enters the detection region; it is then completely separate from the transmitted wave Φ_b except in the forward direction ($\theta < d/D$) where both of these waves are important and will interfere [1]). We thus find the qualitative results stated in Chapter X, § 4.

6. Calculation of Cross Sections

Before calculating the cross sections, let us clearly state what is involved in the operation of detection. Whatever the experimental arrangement used, it essentially consists in placing a diaphragm at a distance D in the direction $\Omega = (\theta, \varphi)$ with respect to the target. The opening of this diaphragm is arranged in such a way that it passes without distortion any wave emitted into a given solid

---

[1]) The interference term must not be neglected since it is this term which insures the conservation of the norm (Problem X.1).

angle $(\Omega, \Omega+\delta\Omega)$ to the exclusion of the rest. A detection system placed behind the diaphragm is triggered whenever a particle traverses the latter. The probability $P_{\boldsymbol{b}}(\Omega)\delta\Omega$ that the detector is triggered by a particle whose motion before detection is represented by the wave $\Psi_{\boldsymbol{b}}(\boldsymbol{r}, t)$, is equal to the integral of the flux of this wave across the diaphragm, evaluated over the entire duration of the collision [1]), or, what is equivalent, to the probability of presence of the scattered particle in the solid angle $(\Omega, \Omega+\delta\Omega)$ when the collision has occurred $(t=T \gg l/v)$.

Since the detector is located sufficiently far away in the transverse direction so that the transmitted wave can never reach it [condition (X.8)], only the scattered wave $\Psi_{\boldsymbol{b}}^{(d)}$ enters into this calculation of the probability of presence. From its expression (X.20), we obtain

$$\begin{aligned} P_{\boldsymbol{b}}(\Omega) &= \int_0^\infty |\Psi_{\boldsymbol{b}}^{(d)}(\boldsymbol{r}, T)|^2 r^2 \, \mathrm{d}r \\ &= |f_{\boldsymbol{k}}(\Omega)|^2 \int_0^\infty |\chi[\boldsymbol{u}(r-vT)+\boldsymbol{s}-\boldsymbol{b}]|^2 \, \mathrm{d}r. \end{aligned} \tag{X.21}$$

Since $vT \gg l$, one can make the change of variable

$$z = r - vT$$

and extend the limit of integration to $-\infty$, which yields

$$P_{\boldsymbol{b}}(\Omega) = |f_{\boldsymbol{k}}(\Omega)|^2 \int_{-\infty}^{+\infty} |\chi(\boldsymbol{u}z+\boldsymbol{s}-\boldsymbol{b})|^2 \, \mathrm{d}z.$$

Now consider a beam of particles of unit flux: there are $\mathrm{d}\boldsymbol{b}$ incident particles upon the surface element $(\boldsymbol{b}, \boldsymbol{b}+\mathrm{d}\boldsymbol{b})$ per unit time, and each of these has a probability $P_{\boldsymbol{b}}(\Omega)\delta(\Omega)$ of being scattered in the direction $(\Omega, \Omega+\delta\Omega)$. One obtains the probability of scattering into $(\Omega, \Omega+\delta\Omega)$ per unit time and unit flux by integrating this expression over $\boldsymbol{b}$:

$$\sigma(\Omega) = |f_{\boldsymbol{k}}(\Omega)|^2 \int_{-\infty}^{+\infty} \mathrm{d}z \int \mathrm{d}\boldsymbol{b} \, |\chi(\boldsymbol{u}z+\boldsymbol{s}-\boldsymbol{b})|^2.$$

The region of integration of $\boldsymbol{b}$ is the plane perpendicular to $\boldsymbol{u}$. By carrying out the change of variable $\boldsymbol{\rho}=\boldsymbol{u}z+\boldsymbol{s}-\boldsymbol{b}$, one reduces the

[1]) The instant at which the particle traverses the diaphragm is equal, on the average, to $[D + (\boldsymbol{u}\cdot\boldsymbol{s})]/v$. It can actually not be predicted with a precision greater than l/v, in accordance with the time-energy uncertainty relation. $(\boldsymbol{u}\cdot\boldsymbol{s})/v$ can be interpreted as a delay in the transmission of the scattered wave. However, the experimental arrangement considered here certainly does not permit such a delay to be observed since $s \ll l$ (cf. discussion on the delay upon reflection of Ch. III, § 6). The observation of delays of this kind presupposes a sufficiently poor energy definition (cf. Ch. X, § 16).

triple integral above to the normalization integral of the function χ [eq. (X.10)]. We thus arrive at the expression (X.2):

$$\sigma(\Omega) = |f_k(\Omega)|^2.$$

7. Collision of Two Particles. Laboratory System and Center-of-Mass System

The technique of separating the center-of-mass motion permits to extend the treatment of the scattering of a particle by a potential, to the collisions of two particles subjected to a potential $V(r)$ which depends only upon their relative position. As was shown in Sec. III of Chapter IX – we take over the notation of that section – the motion of the two particles is decomposed into two separate motions: that of their center of mass which travels like a free particle, and that of the "relative particle" of mass $m = m_1 m_2/(m_1 + m_2)$ subjected to the potential $V(r)$.

In a typical scattering experiment, one bombards a target composed of particles of type 2 with a monoërgic beam of particles of type 1, and one counts the number of particles of one type, particles 1 for instance, emitted in a given direction $\Omega_1 = (\theta_1, \varphi_1)$. Before collision, particle 2 is at rest, particle 1 travels with a given velocity $\mathbf{v}$, and the center of mass moves with the velocity

$$\mathbf{V} = \frac{m_1}{M}\mathbf{v} \qquad (M = m_1 + m_2).$$

The total energy of the system is the sum of the energies of the center-of-mass motion and the relative motion

$$E = E_R + E_r$$

$$\left(E = \tfrac{1}{2}m_1 v^2; \qquad E_R = \tfrac{1}{2}M V^2 = \frac{m_1}{m_1 + m_2}E; \qquad E_r = \tfrac{1}{2}m v^2 = \frac{m_2}{m_1 + m_2}E\right).$$

During the collision the center of mass continues its uniform rectilinear motion, the collision phenomenon affecting only the relative coordinate. It is clear that the scattering cross section $\sigma_1(\Omega)$ is related to the asymptotic behavior of the stationary states of energy E_r of the relative coordinate.

In order to establish that relation, it is convenient to change the system of reference and to study the same phenomenon in a frame of reference where the center of mass is at rest. One usually calls the

laboratory system the frame of reference where the target particle is at rest, and *center-of-mass system* the one where the center of mass is at rest. The first is the one considered above; the second is in uniform translational motion of velocity $\boldsymbol{V}$ with respect to the former. Passing from one frame of reference to the other affects only the center-of-mass motion; the motion of the relative particle remains unchanged.

The definition of cross sections given in Ch. X, § 2 in no way supposes that the target be initially at rest. We emphasize that the incident flux which enters into that definition is the *relative flux* of the projectile with respect to the scatterer; it is a quantity independent of the frame of reference. Once this is understood, one can define the differential cross section $\sigma(\Omega)$ in the center-of-mass system in the same way one defines the differential cross section $\sigma_1(\Omega)$ of the same process in the laboratory system. $\sigma(\Omega)$ is the number of particles 1 emitted per unit time and per unit solid angle in the direction Ω when one bombards particle 2 with an incident flux of particles 1 of relative flux unity, all observations being carried out in the center-of-mass system, the angles of emission also being measured in that system.

These definitions imply that

$$\sigma(\Omega)\,\mathrm{d}\Omega = \sigma_1(\Omega_1)\,\mathrm{d}\Omega_1, \tag{X.22}$$

an expression in which Ω_1 designates the direction of emission of particle 1 in the laboratory system when this particle travels in the direction Ω in the center-of-mass system. Note the equality of the total cross sections:

$$\sigma_{\text{tot.}} \equiv \int \sigma(\Omega)\,\mathrm{d}\Omega = \sigma_{1\,\text{tot.}}.$$

This in fact was evident *a priori* since the total cross section is the total number of particles scattered per unit incident flux, a quantity that is invariant under a change of reference frame.

$\sigma(\Omega)$ is related more directly than $\sigma_1(\Omega_1)$ to the (three-dimensional) scattering problem of the relative particle by the potential $V(\boldsymbol{r})$. Indeed, in the center-of-mass system the direction of propagation of particle 1 and that of the relative particle are identical (particle 2 travelling in the opposite direction). Moreover, since the incident flux of the relative particle with respect to the center of force ($\boldsymbol{r}=0$) is equal to the incident flux of our collision problem, $\sigma(\Omega)$ is the scattering cross section of the relative particle in the direction Ω, that is to say

the differential scattering cross section in the direction Ω of a particle of mass m and initial velocity $\mathbf{v}$, by the potential $V(\mathbf{r})$.

In particular, if $V(\mathbf{r})$ tends asymptotically to zero more rapidly than $1/r$, the Schrödinger equation of the relative particle

$$\left[-\frac{\hbar^2}{2m}\triangle + V(\mathbf{r})\right]\psi(\mathbf{r}) = E_r\psi(\mathbf{r})$$

possesses an eigensolution of energy E_r with asymptotic behavior

$$e^{i\mathbf{k}\cdot\mathbf{r}} + f(\Omega)\frac{e^{ikr}}{r}$$

($E_r = \hbar^2k^2/2m$, $\mathbf{k} = m\mathbf{v}/\hbar = m_2\mathbf{k}_1/M$, $\mathbf{k}_1$ being the incident wave vector

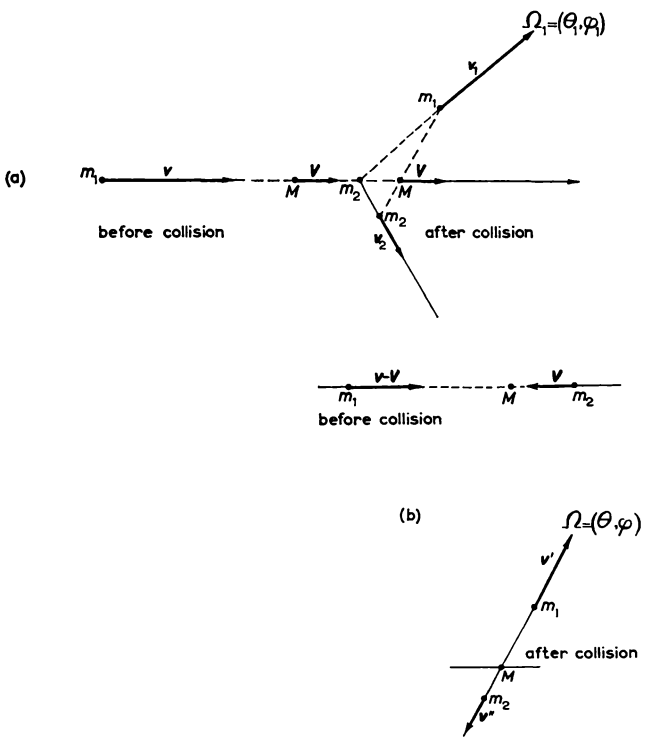


Fig. X.3. (*a*) Collision in the laboratory system ($\mathbf{V} = m_1\mathbf{v}/M$).
(*b*) The same collision in the center-of-mass system.

in the laboratory system) and, applying the results of the foregoing paragraphs

$$\sigma(\Omega) = |f(\Omega)|^2.$$

We have yet to establish the relation between Ω and Ω_1, in order to be able to derive from the above expressions [1]) the cross section in the laboratory system $\sigma_1(\Omega_1)$. To this effect we take as polar axis in the two systems the axis parallel to the direction of propagation. Fig. X.3*a* gives the schematic diagram of a collision in the laboratory system, while Fig. X.3*b* gives the same in the center-of-mass system. The initial and final velocities of the two particles are given in the following table, along with their spherical polar coordinates:

	Initial Velocities		Final Velocities	
	1	2	1	2
Laboratory System	$\boldsymbol{v}(v, 0, 0)$	0	$\boldsymbol{v}_1(v_1, \theta_1, \varphi_1)$	$\boldsymbol{v}_2(v_2, \theta_2, \varphi_2)$
Center-of-Mass System	$\boldsymbol{v} - \boldsymbol{V}(v - V, 0, 0)$	$-\boldsymbol{V}(V, \pi, 0)$	$\boldsymbol{v}'(v', \theta, \varphi)$	$\boldsymbol{v}''(v'', \pi - \theta, \varphi + \pi)$

$$\boldsymbol{V} = \frac{m_1}{m_1 + m_2}\boldsymbol{v}, \qquad v' = v - V = \frac{m_2}{m_1 + m_2}v, \qquad v'' = V = \frac{m_1}{m_1 + m_2}v$$

(θ_1, φ_1) is defined as function of (θ, φ) by the vector equality [2])

$$\boldsymbol{v}_1 = \boldsymbol{v}' + \boldsymbol{V} \tag{X.23}$$

with

$$\varphi_1 = \varphi, \qquad v_1 \sin\theta_1 = v' \sin\theta, \qquad v_1 \cos\theta_1 = v' \cos\theta + V,$$

[1]) All calculations made here are valid only in the non-relativistic approximation. However, the concept of center-of-mass system remains valid in relativistic mechanics: it is the frame of reference where the total momentum (momentum of particle 1 + momentum of particle 2) is zero. One passes from the center-of-mass system to the laboratory system by a Lorentz transformation. In the non-relativistic approximation considered here, this Lorentz transformation reduces to a Galilean transformation, namely

$$\boldsymbol{r}_i \to \boldsymbol{r}_i + \boldsymbol{V}t, \qquad \boldsymbol{p}_i \to \boldsymbol{p}_i + m_i\boldsymbol{V}.$$

[2]) Likewise the relation between (θ_2, φ_2) and (θ, φ) is defined by the equality $\boldsymbol{v}_2 = \boldsymbol{v}'' + \boldsymbol{V}$. Since $v'' = V$, it is easy to show that $\theta_2 = \frac{1}{2}(\pi - \theta)$, $\varphi_2 = \varphi + \pi$.

from which

$$\tan\theta_1 = \frac{\sin\theta}{\cos\theta + \tau},$$

or else

$$\cos\theta_1 = \frac{\cos\theta + \tau}{(1 + 2\tau\cos\theta + \tau^2)^{\frac{1}{2}}}, \tag{X.24}$$

expressions in which we have put

$$\tau = \frac{V}{v'} = \frac{m_1}{m_2}. \tag{X.25}$$

The vector sum (X.23) can actually be made graphically; relation (X.24) can be read directly from the figure and the discussion of the

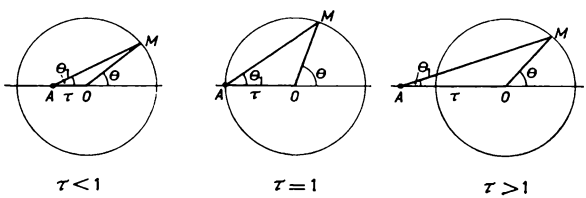


Fig. X.4. Geometrical construction of θ_1 as function of θ:
$OA = \tau$ $OM = 1$ $(\tau = V/v' = m_1/m_2)$.

results is easier in terms of the figure than in terms of formula (X.24) itself. This graphical construction is reproduced in Fig. X.4. Two cases may arise:

(*i*) $\tau < 1$ $(m_1 < m_2)$. — The angle θ_1 grows monotonically from 0 to π as θ grows from 0 to π. Note that $\frac{1}{2}\theta < \theta_1 < \theta$ for any θ. In the limit where $m_1 \ll m_2$, $\theta_1 \approx \theta$ (the center of mass remains practically identical with particle 2 during the collision, hence it is practically at rest in the laboratory system).

(*ii*) $\tau > 1$ $(m_1 > m_2)$. — As θ grows from 0 to π, θ_1 at first increases from 0 to some maximum value smaller than $\frac{1}{2}\pi$, namely $\theta_{1\,\max.} = \sin^{-1}(1/\tau)$, then it decreases from $\theta_{1\,\max.}$ to zero. To each value of θ_1 thus correspond two possible values of θ, $\theta_<$ and $\theta_>$, related by: $\theta_1 = \frac{1}{2}(\theta_< + \theta_> - \pi)$; to each of these correspond two distinct values of v_1, the larger one corresponding to the smaller value of θ.

[When $\tau = 1$ $(m_1 = m_2)$, one simply has $\theta_1 = \frac{1}{2}\theta$.]

From relation (X.24) we obtain

$$\frac{\mathrm{d}(\cos\theta_1)}{\mathrm{d}(\cos\theta)} = \frac{1+\tau\cos\theta}{(1+2\tau\cos\theta+\tau^2)^{3/2}}$$

and since

$$\frac{\mathrm{d}\Omega_1}{\mathrm{d}\Omega} = \left|\frac{\mathrm{d}(\cos\theta_1)}{\mathrm{d}(\cos\theta)}\right|,$$

we obtain, applying relation (X.22),

$$\sigma_1(\Omega_1) = \sigma(\Omega)\frac{\mathrm{d}\Omega}{\mathrm{d}\Omega_1} = \frac{(1+2\tau\cos\theta+\tau^2)^{3/2}}{|1+\tau\cos\theta|}\,|f(\Omega)|^2. \qquad \text{(X.26)}$$

II. SCATTERING BY A CENTRAL POTENTIAL. PHASE SHIFTS

8. Decomposition into Partial Waves. Phase-Shift Method

Consider the scattering of a particle by a *central potential* $V(r)$. To calculate the cross section, one needs the asymptotic form of the stationary scattering wave ψ. To this effect we solve the Schrödinger equation in spherical coordinates.

The direction of the incident wave vector $\boldsymbol{k}$ is an axis of rotational symmetry of the problem; if we take it as polar axis, the wave ψ and the scattering amplitude f are independent of φ. Expanding in a series of Legendre polynomials, we have

$$\psi(r,\theta) = \sum_l \frac{y_l(r)}{r} P_l(\cos\theta) \qquad \text{(X.27)}$$

$$f(\theta) = \sum_l f_l P_l(\cos\theta). \qquad \text{(X.28)}$$

Let us put

$$\varepsilon = k^2 = \frac{2m}{\hbar^2}E, \qquad U(r) = \frac{2m}{\hbar^2}V(r).$$

y_l is a regular solution of the radial equation

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}r^2} + \left(\varepsilon - U(r) - \frac{l(l+1)}{r^2}\right)\right] y_l = 0. \qquad \text{(X.29)}$$

Its asymptotic form reads

$$y_l \underset{r\to\infty}{\sim} a_l \sin(kr - \tfrac{1}{2}l\pi + \delta_l). \qquad \text{(X.30)}$$

All regular solutions of eq. (X.29) yield the same phase shift δ_l; they differ by the normalization constant a_l. Here, a_l must be adjusted in such a way that $\psi(r, \theta)$ has the desired asymptotic behavior. Using the expansions (IX.35) and (X.28), we can write the asymptotic form of ψ in the form of a series of Legendre polynomials:

$$e^{ik\cdot r} + f(\theta)\frac{e^{ikr}}{r} = \sum_l \left((2l+1)\, i^l\, j_l(kr) + f_l \frac{e^{ikr}}{r}\right) P_l(\cos\theta).$$

If one takes the asymptotic form of $j_l(kr)$ into account we can rewrite this expression, separating the incoming and outgoing waves

$$r\psi(r, \theta) \sim \sum_l \left[(-)^{l+1}\frac{2l+1}{2ik} e^{-ikr} + \left(\frac{2l+1}{2ik} + f_l\right) e^{ikr}\right] P_l(\cos\theta).$$

The asymptotic form of y_l must be equal to the quantity in brackets of the right-hand side. This condition fixes a_l uniquely and enables us to write f_l as function of the phase shift. We have successively

$$a_l = i^l \frac{2l+1}{k} e^{i\delta_l},$$

$$f_l = \frac{2l+1}{k} e^{i\delta_l} \sin\delta_l.$$

Substituting this expression in eq. (X.28) we obtain $f(\theta)$ as a function of the phase shifts δ_l:

$$\boxed{f(\theta) = \lambda\!\!\!^{-} \sum_{l=0}^{\infty} (2l+1)\, e^{i\delta_l} \sin\delta_l\, P_l(\cos\theta).} \qquad \text{(X.31)}$$

$\lambda\!\!\!^{-}$ here refers to the incident wavelength ($\lambda\!\!\!^{-} = 1/k$).

It is instructive to compare the asymptotic form of y_l/r, the component of angular momentum l of the stationary state of scattering, with the corresponding expression of the function $(2l+1)\, i^l\, j_l(kr)$, the component of the same angular momentum of the plane wave $\exp(i\mathbf{k}\cdot\mathbf{r})$:

$$\frac{y_l}{r} \sim \frac{(2l+1)}{2ikr}\left[(-)^{l+1} e^{-ikr} + e^{2i\delta_l} e^{ikr}\right],$$

$$(2l+1)\, i^l\, j_l(kr) \sim \frac{2l+1}{2ikr}\left[(-)^{l+1} e^{-ikr} + e^{ikr}\right].$$

They are both superpositions of an incoming and an outgoing wave of the same intensity. The incoming wave is obviously the same for the two functions. The outgoing wave of the stationary state of scattering differs from that of the plane wave by the presence of the phase factor $\exp(2i\delta_l)$: *the effect of the scattering potential is to shift the phase of each outgoing partial wave.*

The differential cross section is obtained by forming the square modulus of $f(\theta)$:

$$\sigma(\Omega) = \lambda\!\!\!^{-2} \sum_{ll'} (2l+1)(2l'+1)\, e^{i(\delta_l - \delta_{l'})} \sin\delta_l \sin\delta_{l'} P_l(\cos\theta) P_{l'}(\cos\theta). \quad \text{(X.32}$$

Integration over the angles (θ, φ) yields the total cross section $\sigma_{\text{tot.}}$. Taking into account the orthogonality relations of the Legendre polynomials, the latter is written in the form of a series:

$$\boxed{\sigma_{\text{tot.}} = 4\pi\lambda\!\!\!^{-2} \sum_{l=0}^{\infty} (2l+1) \sin^2\delta_l} \qquad \text{(X.33)}$$

each term of which,

$$\sigma_l = 4\pi(2l+1)\,\lambda\!\!\!^{-2} \sin^2\delta_l \qquad \text{(X.34)}$$

represents the contribution to the scattering of the corresponding angular momentum l. Note the inequality:

$$\sigma_l \leqslant 4\pi(2l+1)\lambda\!\!\!^{-2}. \qquad \text{(X.35)}$$

The maximum value of σ_l is reached when the phase shift is a "half-integral" multiple of π:

$$\delta_l = (n+\tfrac{1}{2})\pi \qquad (n \text{ integer}).$$

9. Semi-Classical Representation of the Collision. Impact Parameters

Consider the collision of a *classical* particle with a central force field. The energy of the incident particle being fixed at $E = p^2/2m$, every trajectory may be characterized by its impact parameter b, defined as the distance from the center of force C to the straight line containing the initial momentum $\mathbf{p}_0$ (Fig. X.5). In such a collision, the angular momentum L is a constant of the motion. b is directly proportional to L:

$$L = bp.$$

If the force field has a finite range r_0,

$$V(r) = 0 \quad \text{for} \quad r > r_0,$$

the incident particle does or does not suffer a deflection according to whether $b < r_0$ or $b > r_0$. The deflection is limited to particles whose angular momentum is sufficiently small.

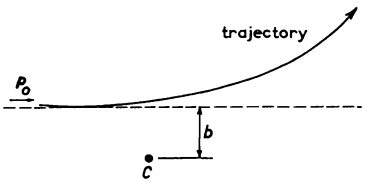


Fig. X.5. Scattering of a classical particle by a center of force C.
p_0 = initial momentum; b = impact parameter.

A collision in quantum theory is very different from a classical collision: it is essentially a wave phenomenon. However, when the scattering potential $V(r)$ is negligible – without necessarily being zero – beyond a certain distance r_0, the phenomenon is somewhat analogous to the scattering of a beam of classical particles by a potential of finite range r_0. *As a general rule, the contribution σ_l of the lth partial wave is negligible* [1] *if* $l\lambda\!\!\!^{-} \gtrsim r_0$; on the other hand, if $l\lambda\!\!\!^{-} < r_0$, it may take on all values between zero and its maximum value $4\pi(2l+1)\lambda\!\!\!^{-2}$. In the classical scattering problem the part of the beam which corresponds to the lth partial wave is given by the particles whose impact parameters range from $l\lambda\!\!\!^{-}$ to $(l+1)\lambda\!\!\!^{-}$, hence whose angular momenta range from $l\hbar$ to $(l+1)\hbar$. The contribution σ_l obviously vanishes if $l\lambda\!\!\!^{-} > r_0$, takes on the value $\pi(2l+1)\lambda\!\!\!^{-2}$ if $(l+1)\lambda\!\!\!^{-} < r_0$, and takes on some intermediate value if r_0 lies between $l\lambda\!\!\!^{-}$ and $(l+1)\lambda\!\!\!^{-}$.

The foregoing rule for quantum-mechanical scattering is based upon the following semi-quantitative argument. The incident wave is a superposition of spherical waves of given angular momentum. The radial portion of the term corresponding to the partial wave of angular momentum l is proportional to $j_l(r/\lambda\!\!\!^{-})$, hence a relative probability density of presence in the spherical shell $(r, r+\mathrm{d}r)$ equal to $r^2 j_l^2(r/\lambda\!\!\!^{-})$. This density is very low as long as $r < \sqrt{l(l+1)}\,\lambda\!\!\!^{-}$ and oscillates between 0 and approximately 1 when $r > \sqrt{l(l+1)}\,\lambda\!\!\!^{-}$ (cf. Fig. X.6). If $r_0 < l\lambda\!\!\!^{-}$, the

[1] This rule is not absolute; we shall encounter exceptions in § 14 in connection with resonance phenomena.

wave practically does not penetrate into the region of strong potential; there is therefore little chance that the wave be greatly affected by the presence of the potential.

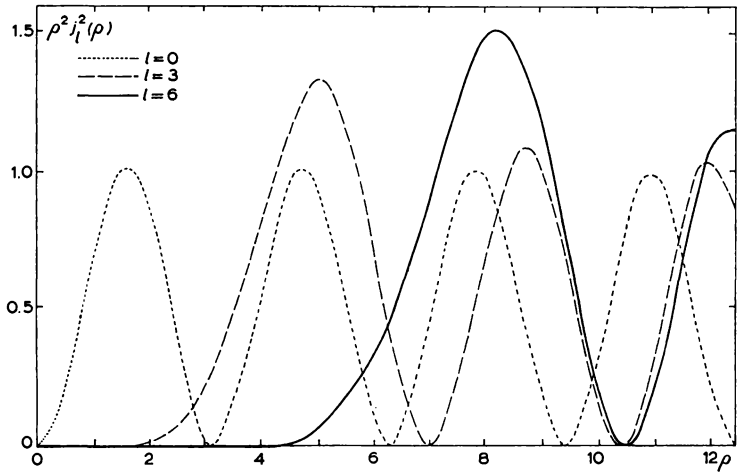


Fig. X.6. Diagram of the function $\varrho^2 j_l^2(\varrho)$ for $l = 0, 3, 6$.

The argument given here is not rigorous. More precise indications concerning the rapidity of convergence of the series (X.31) and (X.32) will be given in § 12. Be that as it may, *the phase-shift method is particularly well suited to the calculation of cross sections whenever the range of the scattering potential does not extend beyond several wavelengths.*

III. POTENTIAL OF FINITE RANGE [1])

10. Relation between Phase Shift and Logarithmic Derivative

Suppose that $V(r)$ is entirely concentrated in a finite region of space, so that one rigorously has

$$V(r) = 0 \quad \text{when} \quad r > r_0.$$

[1]) Range is used here in a somewhat restricted sense; it is the value of the radial distance beyond which the potential rigorously vanishes. More generally, the range of a potential is the value of the radial distance beyond which it takes on negligible values. The results obtained below hold essentially true for potentials of finite range in this broader sense; the effect of the tail of the potential may then be estimated with the help of the generalized Born formula. (See § 19 below.)

Denote by q_l the value taken at the point r_0 by the logarithmic derivative of the solution regular at the origin, of the radial equation (X.29)

$$q_l = \left.\frac{r(\mathrm{d}y_l/\mathrm{d}r)}{y_l}\right|_{r=r_0} \tag{X.36}$$

(the definition adopted here differs by a factor r from the usual definition). q_l is known to be a monotonically decreasing function of the energy (Ch. III, § 8); the exact form of this function evidently depends upon the potential $V(r)$ under consideration.

Since the latter vanishes in the region $r > r_0$, there exists between q_l and δ_l a relation which is independent of the particular form of $V(r)$; specifying q_l suffices to determine the asymptotic behavior of the solution.

In the following, we assume y_l to be normalized in such a way that

$$y_l \underset{r\to\infty}{\sim} \sin\,(kr - \tfrac{1}{2}l\pi + \delta_l). \tag{X.37}$$

We therefore have in the external region

$$y_l = kr[\cos\,\delta_l\, j_l(kr) + \sin\,\delta_l\, n_l(kr)] \qquad (r > r_0).$$

For purposes of later discussions, it is convenient to put

$$\xi = kr$$

and to introduce the outgoing and incoming waves

$$u_l^{(\pm)}(\xi) = \xi[n_l(\xi) \pm \mathrm{i} j_l(\xi)] = \xi h_l^{(\pm)}(\xi)$$

whose Wronskian, a quantity independent of ξ, is

$$u^{(-)}\frac{\mathrm{d}}{\mathrm{d}\xi}u^{(+)} - u^{(+)}\frac{\mathrm{d}}{\mathrm{d}\xi}u^{(-)} = 2\mathrm{i}. \tag{X.38}$$

In the external region,

$$\begin{aligned} y_l &= \frac{1}{2\mathrm{i}}\,(u_l^{(+)}\,\mathrm{e}^{\mathrm{i}\delta_l} - u_l^{(-)}\,\mathrm{e}^{-\mathrm{i}\delta_l}) \\ &= \mathrm{Im}\; u_l^{(+)}\,\mathrm{e}^{\mathrm{i}\delta_l}. \end{aligned} \qquad (r > r_0) \tag{X.39}$$

The continuity condition on the logarithmic derivative at $r = r_0$ yields the relation

$$q_l = \xi\left.\frac{\mathrm{Im}\;\mathrm{e}^{\mathrm{i}\delta_l}\,(\mathrm{d}u_l^{(+)}/\mathrm{d}\xi)}{\mathrm{Im}\;\mathrm{e}^{\mathrm{i}\delta_l}\,u_l^{(+)}}\right|_{\xi=kr_0}. \tag{X.40}$$

This is the relation between q_l and δ_l we were seeking.

To put this relation into a more tractable form, it is convenient to put

$$\begin{aligned} |u_l^{(\pm)}(kr_0)| &= v_l^{-\frac{1}{2}}, \\ u_l^{(\pm)}(kr_0) &= v_l^{-\frac{1}{2}}\, e^{\mp i\tau_l} \\ \frac{\xi\,(du_l^{(\pm)}/d\xi)}{u_l^{(\pm)}}\bigg|_{\xi=kr_0} &= q_l^{(\pm)}. \end{aligned} \tag{X.41}$$

From equation (X.38) we obtain the relation (Problem X.3)

$$\operatorname{Im} q_l^{(+)} = kr_0 v_l;$$

v_l is a positive quantity less than 1, which is smaller, the smaller kr_0 and the larger l; it is called the *penetration factor*.

With these notations the continuity relation (X.40) is written

$$e^{2i\delta_l} = e^{2i\tau_l}\,\frac{q_l - q_l^{(-)}}{q_l - q_l^{(+)}}$$

or

$$\delta_l = \tau_l + \varrho_l \tag{X.42}$$

with

$$\varrho_l = \arg\,(q_l - q_l^{(-)}) = \tan^{-1}\frac{kr_0\, v_l}{q_l - \operatorname{Re} q_l^{(+)}}. \tag{X.43}$$

In conclusion, δ_l can be put in the form of a sum of two terms the first of which, τ_l, is independent of the particular form of the scattering potential, and the second, ϱ_l, depends on it through q_l, in accordance with eq. (X.43).

Note that

$$y_l(r_0) = \frac{\sin \varrho_l}{\sqrt{v_l}}. \tag{X.44}$$

11. Behavior of the Phase Shift at Low Energies ($\lambda \to \infty$)

Knowing the behavior of the spherical Bessel functions for very small values of the argument [eq. (B.52)], one may deduce from eqs. (X.42) and (X.43) the behavior of δ_l when $kr_0 \ll l$, that is at very low energies or for very large values of the angular momentum. Indeed, if $kr_0 \ll l$,

$$\begin{aligned} \tau_l &\approx -\frac{(kr_0)^{2l+1}}{(2l+1)!!\,(2l-1)!!}, \\ v_l &\approx \frac{(kr_0)^{2l}}{[(2l-1)!!]^2} \qquad \operatorname{Re} q_l^{(+)} \approx -l + O(k^2 r_0^2) \end{aligned} \tag{X.45}$$

(these expressions are also valid for $l=0$ when $kr_0 \ll 1$).

In this section we examine the behavior of the cross sections as the energy approaches zero.

In that limit, the real quantity q_l increases toward a certain limit $\hat{q}_l$. In general $\hat{q}_l \neq -l$ and the phase shift δ_l tends to zero as k^{2l+1}. One finds

$$\delta_l \underset{k\to 0}{\sim} \frac{l+1-\hat{q}_l}{l+\hat{q}_l} \frac{(kr_0)^{2l+1}}{(2l+1)!!\,(2l-1)!!}. \tag{X.46}$$

The amplitude f_l, proportional to δ_l/k, thus tends toward zero as k^{2l}. Therefore, in the limit of very low energies, the cross section becomes isotropic, since all partial cross sections σ_l vanish as k^{4l} [eq. (X.34)], except for the s-wave cross section σ_0, which generally approaches a constant different from zero.

By definition, one calls *scattering length* the length

$$\begin{aligned} a &= -\lim_{k\to 0} f_0 = -\lim \frac{\delta_0}{k} \\ &= \left(1 - \frac{1}{\hat{q}_0}\right) r_0 . \end{aligned} \tag{X.47}$$

a is obtained by solving the radial equation for zero energy:

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}r^2} - U(r)\right] y_0 = 0.$$

It is the abscissa of the point where the asymptote of y_0 cuts the r axis. In the limit of very low energies

$$\sigma_{\text{tot.}} = \sigma_0 = 4\pi a^2.$$

If $\hat{q}_l = -l$ accidentally, f_l behaves as k^{2l-2} when the energy tends to zero (except for $l=0$ where $\hat{q}_0 = 0$ implies: $f_0 \sim \mathrm{i}/k$). There is said to be a resonance at zero energy in the state l. The foregoing conclusions must then be modified as follows. If it is an s resonance ($l=0$), a is infinite and $\sigma_{\text{tot.}}$ approaches infinity as $1/E$. If it is a p resonance ($l=1$), $f(\theta)$ approaches the form $-a+b\cos\theta$; the cross section remains finite but never becomes isotropic. If the resonance is of higher order, the zero-energy limit of the cross sections is not affected.

12. Partial Waves of Higher Order. Convergence of the Series ($l \to \infty$)

The asymptotic forms (X.45) likewise enable us to obtain asymptotic expressions for the phase shifts of higher order when one picks a

well-defined value of the energy. If l is sufficiently large, the expression

$$k^2 - U(r) - \frac{l(l+1)}{r^2}$$

always remains negative in the interval $(0, r_0)$; hence the solution y_l of eq. (X.29) has an "exponential" behavior in that interval, and q_l is certainly positive. Moreover, if $l \gg kr_0$, the expressions (X.45) can be used and since $q_l + l$ is certainly not zero, we find, for l sufficiently large, an asymptotic expression similar to expression (X.46):

$$\delta_l \underset{l\to\infty}{\sim} \frac{l+1-q_l}{l+q_l} \frac{(kr_0)^{2l+1}}{(2l-1)!!\,(2l+1)!!}. \tag{X.48}$$

This expression provides us with a measure of the rapidity of convergence of the expansion in a series of partial waves when the potential has a finite range. Thus the order of magnitude predictions stated in § 9 are confirmed.

13. Scattering by a Hard Sphere

If the potential of finite range is that of a hard sphere, that is to say, if $V(r)$ is infinitely repulsive in the region $r < r_0$, and zero in the region $r > r_0$,

$$V(r) = \begin{cases} +\infty & \text{if} \quad r < r_0 \\ 0 & \text{if} \quad r > r_0 \end{cases}$$

all formulae of § 10 simplify. The wave function must vanish on the surface of the sphere, hence $y_l(r_0) = 0$ for any $l (q_l = -\infty)$, which yields [eq. (X.44)] $\varrho_l = 0$, or

$$\delta_l = \tau_l = \arg u_l^{(-)}. \tag{X.49}$$

From this one deduces after some calculation that

$$\sigma_l = \frac{4\pi(2l+1)}{k^2} \frac{j_l^2(kr_0)}{j_l^2(kr_0) + n_l^2(kr_0)} \tag{X.50}$$

and in particlar,

$$\sigma_0 = 4\pi r_0^2 \left(\frac{\sin kr_0}{kr_0}\right)^2.$$

At very low energies, in accordance with the study of § 11, the

differential cross section becomes isotropic and the total cross section has as its limit

$$\lim_{k\to 0} \sigma_{\text{tot.}} = \lim_{k\to 0} \sigma_0 = 4\pi r_0^2, \tag{X.51}$$

corresponding to a scattering length $a = r_0$.

As the energy increases, the contribution of higher-order partial waves becomes more and more important, and the anisotropy of the scattering becomes more and more marked. At very high energies ($\lambda\!\!\!^{-} \ll r_0$), the differential and total cross sections can be evaluated by making use of the asymptotic properties of the Bessel functions for large values of the order l. One finds[1])

$$\sigma(\Omega) \underset{k\to\infty}{\sim} \tfrac{1}{4} r_0^2 \left(1 + \cot^2 \frac{\theta}{2} J_1^2(kr_0 \sin\theta)\right) \tag{X.52}$$

$$\lim_{k\to\infty} \sigma_{\text{tot.}} = 2\pi r_0^2. \tag{X.53}$$

Let us give here a simplified proof of relation (X.53). Knowing the general shape of the functions $j_l(\xi)$ and $n_l(\xi)$ we can deduce the behavior of the function

$$g_l(\xi) = \frac{j_l^2(\xi)}{j_l^2(\xi) + n_l^2(\xi)}.$$

It vanishes as $\xi^{4l+2}/[(2l+1)!!\,(2l-1)!!]^2$ for $\xi = 0$, increase regularly as ξ increases toward the vicinity of $\xi = l$, then oscillates indefinitely according to the law

$$g_l(\xi) \underset{\xi\to\infty}{\sim} \sin^2(\xi - \tfrac{1}{2}l)\pi. \tag{X.54}$$

Hence, in the summation

$$\sigma_{\text{tot.}} = \sum_{l=0}^{\infty} \sigma_l = \frac{4\pi}{k^2} \sum_{l=0}^{\infty} (2l+1)\, g_l(kr_0)$$

---

[1]) Cf. Morse and Feshbach, *loc. cit.*, (footnote in Ch. VI, p. 231), p. 1551. $J_1(x)$ is the first-order Bessel function. When x increases from 0 to $+\infty$, $J_1(x)$ increases from 0 [$J_1(x) \sim \frac{1}{2}x$] to a first maximum $J_1(1.84) \approx 0.58$, then decreases, and vanishes for the first time at $x \approx 3.83$; it then oscillates indefinitely according to the asymptotic expression

$$J_1(x) \underset{x\to\infty}{\sim} \left(\frac{2}{\pi x}\right)^{\frac{1}{2}} \cos(x - \tfrac{3}{4}\pi).$$

the contribution of the terms $l > kr_0$ can be neglected and that of the terms $l < kr_0$ roughly evaluated by replacing $g_l(kr_0)$ by its asymptotic form (X.54), whence

$$\sigma_{\text{tot.}} \underset{k\to\infty}{\sim} \frac{4\pi}{k^2} \sum_{l=0}^{kr} (2l+1) \sin^2 (kr_0 - \tfrac{1}{2}l\pi).$$

One evaluates the summation of the right-hand side by grouping successive terms two by two, which gives in the limit of very large values of k

$$\int_0^{kr_0} l \, dl = \tfrac{1}{2}k^2 r_0^2,$$

from which we obtain equation (X.53).

In contrast to what one might expect, in the limit of short wavelengths ($kr_0 \gg 1$) one does not find the scattering cross sections of a classical particle by a hard sphere of radius r_0. The classical total cross section,

$$\sigma_{\text{cl.}} = \pi r_0^2$$

is just half of the quantum-mechanical result in the limit of very short wavelengths. Likewise, the classical differential cross section is isotropic and equal to $\frac{1}{4}r_0^2$: it is equal to the first term of the asymptotic form (X.52) of $\sigma(\Omega)$.

In fact, the wave aspect of the phenomenon can never be neglected, because even at the very short wavelengths, the potential can never be considered as slowly varying because of the sudden discontinuity at the point $r = r_0$. The observed wave phenomenon is entirely analogous to the diffraction phenomena found in optics, as is revealed by an examination of the asymptotic form (X.52) of the differential cross section. The latter contains two terms. The first is an isotropic "*reflection*" term identical to the classical differential cross section. The second:

$$\tfrac{1}{4}r_0^2 \cot^2 (\tfrac{1}{2}\theta) J_1^2(kr_0 \sin \theta),$$

strongly anisotropic since its contribution is essentially limited to small angles of the order of $\lambda\!\!\!^{-}/r_0$, is the "*diffraction*" term (shadow scattering) resulting from the shadow cast by the perfectly reflecting sphere upon the path of the incident wave.

IV. SCATTERING RESONANCES

14. Scattering by a Deep Square Well

As another example of a potential of finite range, we once again consider the square well of Chapter IX, § 10. We put

$$E = \frac{\hbar^2 k^2}{2m}, \qquad V_0 = \frac{\hbar^2 K_0^2}{2m}, \qquad K^2 = K_0^2 + k^2$$

and we shall study the behavior of the various partial waves as function of energy when the well is very deep. More precisely, we assume:

$$Kr_0 \gg l \tag{X.55}$$

$$K \gg k. \tag{X.56}$$

In that case, the value of q_l to be substituted in the right-hand side of eq. (X.43) is to a good approximation

$$q_l \approx Kr_0 \cot (Kr_0 - \tfrac{1}{2}l\pi). \tag{X.57}$$

The general behavior of $\delta_l(E)$ at low energies — namely $E \ll V_0$ according to condition (X.56) — is easily obtained by inspection.

In accordance with formulae (X.42) and (X.43), δ_l depends on the energy through the quantities τ_l, $kr_0 v_l$, $\operatorname{Re} q_l^{(+)}$, and q_l. The first three are monotonic functions of kr_0 (τ_l a decreasing function, the two others increasing functions), whose behavior at the origin is given by the expressions (X.45) and in the asymptotic region ($kr_0 \gg l$) by the expressions:

$$\begin{gathered} \tau_l \underset{kr_0 \to \infty}{\sim} -(kr_0 - \tfrac{1}{2}l\pi), \\ \lim_{kr_0 \to \infty} v_l = 1 \qquad \lim_{kr_0 \to \infty} \operatorname{Re} q_l^{(+)} = 0, \end{gathered} \tag{X.58}$$

In the energy region of interest, these three functions vary rather slowly. On the other hand, the logarithmic derivative q_l, a monotonically decreasing function of the energy, varies very rapidly and possesses a succession of vertical asymptotes located at the values of the energy for which $Kr_0 = \frac{1}{2}l\pi + n\pi$ (n integer). The energy difference between two neighboring asymptotes is approximately

$$D \approx \pi \frac{\hbar^2 K}{m r_0} \approx \pi \frac{V_0}{K_0 r_0}. \tag{X.59}$$

When the energy varies by this amount, $|q_l|$ is almost everywhere of the order or larger than Kr_0, and one has over almost the entire interval

$$|q_l - \mathrm{Re}\, q_l^{(+)}| \gg kr_0 v_l.$$

The second term of the phase shift, ϱ_l, remains very small (to within $n\pi$):

$$\varrho_l \approx \frac{kr_0 v_l}{q_l - \mathrm{Re}\, q_l^{(+)}} \lesssim \frac{k}{K} v_l,$$

while there exists no *a priori* limitation on τ_l. One can therefore write

$$\delta_l \approx \tau_l.$$

The phase shift is practically the same as that which would be produced by a hard sphere of the same radius. Over the major portion of the range of variation of the energy, therefore, the potential scatters each partial wave in much the same way as a hard sphere: this is the so-called "potential scattering"; the incident wave practically does not penetrate into the internal region.

However, there exists a small energy domain surrounding the point E_r, where $q_l = \mathrm{Re}\, q_l^{(+)}$, for which

$$|q_l - \mathrm{Re}\, q_l^{(+)}| \lesssim kr_0 v_l.$$

Let us define the quantities

$$\left.\begin{aligned} \gamma &= -\left.\frac{\mathrm{d}E}{\mathrm{d}q_l}\right|_{E=E_r} \qquad (\gamma > 0) \\ \Gamma &= 2kr_0 v_l \gamma \end{aligned}\right\} \tag{X.60}$$

Γ is the width of the energy domain in question. One notes that [1])

$$\frac{\Gamma}{D} \approx \frac{2}{\pi}\frac{k}{K} v_l \ll 1.$$

---

[1]) Indeed, according to equation (X.57)

$$\frac{\mathrm{d}q_l}{\mathrm{d}(Kr_0)} = -Kr_0\left(1 + \frac{q_l(q_l-1)}{(Kr_0)^2}\right).$$

Since $Kr_0 \gg l$ and $|\mathrm{Re}\, q_l^{(+)}| \lesssim l$, $\mathrm{d}q_l/\mathrm{d}K \approx -Kr_0^2$ when $E = E_r$, hence

$$\gamma \approx \frac{\hbar^2}{mr_0^2} \quad \text{and} \quad \Gamma \approx 2v_l \frac{\hbar^2 k}{mr_0}.$$

For an energy variation of the order of several Γ on either side of E_r, $q_l - \mathrm{Re}\, q_l^{(+)}$ decreases from values very much larger than $kr_0 v_l$ to values very much smaller than $-kr_0 v_l$, and ϱ_l suddenly passes from values close to $n\pi$ to values in the neighborhood of $(n+1)\pi$. The partial cross section σ_l undergoes violent variations during which it reaches its maximum value $4\pi(2l+1)\lambda\!\!\!^{-2}$: there is said to be an *l resonance*. By definition, E_r is the resonance energy [1]), Γ is the width of the resonance. Γ is the product of a quantity depending upon the general shape of the potential in the internal region (practically independent of l) and the factor $kr_0 v_l$ depending upon the behavior of the wave in the external region (which is smaller, the smaller kr_0 and the larger l).

The resonance region is sufficiently narrow that one may replace the curve $q_l(E)$ by its slope at the point $E = E_r$ in the expression of ϱ_l, hence

$$\varrho_l \approx \tan^{-1} \frac{\Gamma}{2(E_r - E)}. \tag{X.61}$$

Using the continuity condition (X.44) to normalize the radial function in the internal region, one finds to this same approximation (valid over an energy interval surrounding the resonance such that $\Gamma \ll \Delta E \ll D$):

$$y_l = \frac{1}{\sqrt{v_l}} \frac{\Gamma}{\sqrt{4(E-E_r)^2 + \Gamma^2}} K_r j_l(Kr). \qquad (r < r_0) \tag{X.62}$$

The crossing of an l resonance is therefore accompanied by a sudden increase of the intensity of the partial wave of order l in the internal region.

The foregoing analysis simplifies considerably in the case of an s wave. One rigorously has

$$\tau_0 = -kr_0, \quad v_0 = 1, \quad \mathrm{Re}\, q_0^{(+)} = 0, \quad q_0 = Kr_0 \cot Kr_0$$

$$\delta_0 = -kr_0 + \tan^{-1}\left(\frac{k}{K} \tan Kr_0\right).$$

---

[1]) Because of the presence of the potential scattering term, the maximum value of σ_l is reached for an energy value that is somewhat different from the resonance energy, the latter being defined as the energy for which $\varrho_l = \frac{1}{2}\pi$ (to within $n\pi$). It may even happen that $\tau_l = \frac{1}{2}\pi$ at resonance and therefore that σ_l vanishes when $E = E_r$; the crossing of the resonance region manifests itself by a sudden drop to zero of the function $\sigma_l(E)$.

The solution y_0 reads explicitly

$$y_0 = \begin{cases} \sin(kr+\delta_0) & r>r_0 \\ \dfrac{k}{\sqrt{k^2+K_0^2\cos^2 Kr_0}}\sin Kr & r<r_0. \end{cases}$$

Except for some differences in notation ($L \to r_0$, $K \to K_0$, $\eta K \to k$, $\xi K \to K$, $\varphi_1 \to \varrho_0 - \frac{1}{2}\pi$) the problem of s-wave scattering is identical to the problem of wave reflection by the one-dimensional square well treated in Ch. III, § 6 (case b), and the discussion of the resonance effect can be carried over without modification.

15. Study of a Scattering Resonance. Metastable States

Resonance phenomena are frequently encountered in microscopic physics. Scattering resonances of the type we have discussed in the case of the square well occur more generally with potentials, which are strongly attractive in a limited region of space. Because of the importance of phenomena of this type, we shall make a detailed study of an l resonance of scattering. The particular study carried out here refers to the square well, but it can be generalized almost point for point to more general potentials, the form of the potential entering only through the law of variation of the logarithmic derivative q.

For simplicity, we assume the resonances to be sufficiently narrow and sufficiently well separated so that only a single partial wave exhibits a resonance in the energy region under consideration. Moreover, we assume the resonance energy to be so low that

$$kr_0 \ll 1 \tag{X.63}$$

and that, consequently, the contribution to the scattering of the potential scattering terms can be completely neglected [1]). In other words, all the phase shifts are practically zero with the exception of the phase shift δ_l, which varies as a function of the incident energy E according to the law

$$\delta_l \approx \varrho_l = \tan^{-1}\frac{\Gamma}{2(E_r - E)}.$$

---

[1]) The contribution of these terms to the cross section is of the order $4\pi r_0^2$, the contribution of an l resonance at the resonance energy is of the order of $4\pi(2l+1)\lambda\!\!\!^{-2} = 4\pi(2l+1)/k^2$.

Therefore

$$e^{i\delta_l} \sin \delta_l = \frac{\tan \delta_l}{1 - i \tan \delta_l} \approx - \frac{\Gamma}{2(E - E_r) + i\Gamma}$$

and the scattering amplitude may be simply written

$$f(\theta) \approx - \frac{2l+1}{k} P_l (\cos \theta) \frac{\Gamma}{2(E - E_r) + i\Gamma}. \tag{X.64}$$

Upon traversing the resonance, the modulus and the derivative of the argument of $f(\theta)$ exhibit very pronounced peaks. One has

$$\sigma(\Omega) = |f(\theta)|^2 = (2l+1)^2 P_l^2 (\cos \theta) \lambda\!\!\!^{-2} \times \frac{\Gamma^2}{4(E - E_r)^2 + \Gamma^2} \tag{X.65}$$

$$\frac{d}{dE} [\arg f(\theta)] = \frac{d\delta_l}{dE} = \frac{2}{\Gamma} \times \frac{\Gamma^2}{4(E - E_r)^2 + \Gamma^2}. \tag{X.66}$$

Equation (X.65) shows that in the vicinity of the resonance — to the extent that the effect of potential scattering is negligible — the angular distribution of scattering does not depend upon the energy, but only upon l, and that the total cross section varies as a function of the energy according to the "Lorentz law":

$$\sigma_{\text{tot.}} = 4\pi(2l+1) \lambda\!\!\!^{-2} \times \frac{\Gamma^2}{4(E - E_r)^2 + \Gamma^2}. \tag{X.67}$$

To interpret eq. (X.66), one must refer to the study of scattering of a wave packet made in Ch. X, §§ 4 to 6. Indeed, using the notations of these sections, one observes that

$$\frac{\boldsymbol{u} \cdot \boldsymbol{s}}{v} = \hbar \frac{d}{dE} [\arg f(\theta)]$$

and that, consequently, expression (X.66) gives the delay in the transmission of the scattered wave (cf. footnote on p. 379). This delay depends upon the energy according to the same Lorentz law as the total cross section, and reaches its maximum value, $2\hbar/\Gamma$, at the resonance energy.

One can therefore picture the resonance phenomenon as follows. Far from the resonance energy, the incident wave practically does not penetrate into the internal region [cf. eq. (X.62)]; everything proceeds as if the wave met a hard sphere. Only a relatively negligible

fraction of that wave is scattered, and this scattering takes place practically without delay (delay of the order $-r_0/v$). In the vicinity of the resonance energy, the incident wave penetrates deeply into the internal region; a large fraction of the incident wave packet thus remains in the internal region during a time of the order of $\hbar/\Gamma$ before being re-emitted in the form of a scattered wave. Thus the existence of a large resonance scattering cross section is explained. During the entire period preceding the re-emission, the probability of presence of the particle inside, or in the vicinity of the internal region is very large, as in a bound state. But whereas a bound state is a stationary state whose lifetime is infinite, the *metastable state* considered here has a lifetime of order $\hbar/\Gamma$. Consequently, instead of being a state of rigorously defined energy, it must be represented by a wave whose spread in energy is of order Γ according to the time-energy uncertainty relation. We are thus led to associate with each resonance a metastable state whose lifetime is $\hbar/\Gamma$ and whose energy is equal on the average to the resonance energy E_r with a spread equal to the width of the resonance Γ.

16. Observation of the Lifetime of Metastable States

Strictly speaking, this semi-classical picture of the resonant scattering phenomenon cannot be pushed too far without contradictions. In fact, under the normal conditions of observation of cross sections such as they were described in §§ 4 to 6, it is completely impossible to detect the metastable state which was just discussed. Indeed, to perform a cross-section measurement at a given energy, the energy spread ΔE of the incident wave packet must be so small that the scattering amplitude remains practically constant over the interval ΔE; in the resonance region this implies

$$\Delta E \ll \Gamma.$$

If this condition is realized, it is possible to detect the law of variation of the cross section as a function of the energy in the resonance region. On the other hand, the collision time $\hbar/\Delta E$, that is, the time necessary for the entire wave packet to penetrate into the scattering region, is much longer than the lifetime $\hbar/\Gamma$ of the metastable state. This latter is therefore completely unobservable (cf. footnote p. 379).

In order to detect the metastable state, one has to operate under complementary experimental conditions (in the sense of Bohr), namely

$$\Delta E \gg \Gamma. \tag{X.68}$$

To be more precise (cf. footnote on p. 372), consider a wave packet of the type contemplated in § 5 which simultaneously satisfies conditions (X.9) and condition (X.68). We further suppose that $E_r \gg \Delta E \gg \Gamma$ [1]). We adopt the notations of §§ 4 to 6, and suppose for simplicity that $b=0$ and $t_0=0$. Substituting expression (X.64) of $f(\theta)$ into (X.17), we obtain the following asymptotic form for the scattered wave packet:

$$\Psi^{(d)} \sim -(2l+1)\, P_l(\cos\theta)\, \frac{\Gamma}{2}\, \frac{e^{i[k_r r-(E_r t/\hbar)]}\, I}{r}$$

$$I = \int \frac{A(\boldsymbol{k}'-\boldsymbol{k})}{E'-E_r+\frac{1}{2}i\Gamma} \exp\left[i(k'-k_r)\, r - i(E'-E_r)t/\hbar\right] \frac{d\boldsymbol{k}'}{k'}.$$

(we use the notation $E_r = \frac{1}{2}\hbar^2 k_r^2/m = \frac{1}{2}m v_r^2$).
In the integral I, the main contribution comes from the region where $A(\boldsymbol{k}'-\boldsymbol{k})/(E'-E_r+\frac{1}{2}i\Gamma)$ is large. Going over to polar coordinates, we put

$$d\boldsymbol{k}' = k'^2\, d\Omega'\, dk' = \frac{mk'}{\hbar^2}\, d\Omega'\, dE'$$

and $\boldsymbol{k}' = k'\boldsymbol{u}'$. The angular integration involves only $A(\boldsymbol{k}'-\boldsymbol{k})$. As for the integration over dE', according to hypothesis (X.68), the only important region is the region $|E'-E_r| \lesssim \Gamma$, region in which $A(\boldsymbol{k}'-\boldsymbol{k})$ can be replaced by $A(k_r\boldsymbol{u}'-\boldsymbol{k})$, and k' by the first two terms of its Taylor expansion (we have assumed that $\Gamma \ll E_r$):

$$k' \approx k_r + \frac{1}{\hbar v_r}(E'-E_r).$$

We can therefore write

$$I \approx \frac{m}{\hbar^2} A_r F\left(t - \frac{r}{v_r}\right),$$

with the definitions:

$$A_r = \int A(k_r\boldsymbol{u}'-\boldsymbol{k})\, d\Omega'$$

$$F(\tau) = \int_0^\infty \frac{e^{-i(E'-E_r)\tau/\hbar}}{E'-E_r+\frac{1}{2}i\Gamma}\, dE'.$$

---

[1]) This restriction and condition (X.63) are not essential, but they permit to put the final result into a simpler form. In order that they may be realized at the same time as conditions (X.9) and (X.68), one must have

$$v_l \ll k r_0 \ll 1.$$

The reader may easily convince himself that this expression for I is justified only if $|\tau| \gg \hbar/\Delta E$. Note that

$$F(\tau) = \int_{-(2E_r/\Gamma)}^{\infty} \frac{\exp[-i\Gamma\tau z/2\hbar]}{z+i}\,dz.$$

Since $|\tau| \gg \hbar/E_r$, one can replace the lower limit of the integral by $-\infty$, and the integral $F(\tau)$ is easily calculated by the method of residues, namely

$$F(\tau) = \begin{cases} 0 & \text{if } \tau \ll -\hbar/\Delta E \quad <0 \\ -2\pi i\, e^{-\Gamma\tau/2\hbar} & \text{if } \tau \gg \hbar/\Delta E \quad >0. \end{cases} \tag{X.69}$$

One finally has

$$\Psi^{(d)} \sim -(2l+1)\,P_l(\cos\theta)\,\frac{mA_r}{2\hbar^2}\,\Gamma \times F\left(t-\frac{r}{v_r}\right)\frac{e^{i[k_r r-(E_r t/\hbar)]}}{r}. \tag{X.70}$$

The general behavior of this wave derives from the properties of the function $F(\tau)$ given by expression (X.69). It is an outgoing wave bounded by a wave front moving according to the law $r = v_r t$; at a given point, the intensity of the wave is zero at first, then passes suddenly from 0 to some positive value. This transition corresponds to the passage of the wave front; it lasts for a time of order $\hbar/\Delta E$, i.e. a time interval very short compared to $\hbar/\Gamma$; afterwards, the intensity decreases according to the law $\exp(-\Gamma t/\hbar)$.

In a typical experiment designed to observe this law of exponential decrease, one turns on, during a very short time, a beam of wave packets satisfying the foregoing conditions, and one counts the particles scattered into the solid angle $(\Omega, \Omega+\delta\Omega)$ by placing a detector in that direction at a certain distance D from the scattering center. Since the energy spread [1]) of the incident wave packets is very large, $(\Delta E \gg \Gamma)$, the instant $t=0$ at which the collisions take place is very well defined: $\Delta t \ll \hbar/\Gamma$. The counting rate is given by the value assumed by $|\Psi^{(d)}|^2$ at the place where the counter is located. According to eq.(X.70), this is proportional to $F^2(t-D/v_r)$. No particle is detected up to the instant D/v_r at which the wave front reaches the counter. D/v_r is the time needed for a particle emitted by the scattering center at the "resonance velocity" v_r to reach the counter. Afterwards, the counting rate is given by the law $\exp(-\Gamma t/\hbar)$ corresponding to the formation at the instant $t=0$ of a metastable state of mean life $\hbar/\Gamma$.

The foregoing experimental conditions are usually realized in the disintegration of radioactive nuclei (α and β radioactivity, γ radioactivity of isomeric nuclei).

---

[1]) We are dealing here with the energy spread of *each wave packet taken individually*.

V. VARIOUS FORMULAE AND PROPERTIES

17. Integral Representations of Phase Shifts

Certain properties or methods of calculation of phase shifts may be obtained by starting from suitable integral representations of the phase shifts. There are many such integral representations. Most of them are obtained simply by applying the Wronskian theorem to suitably defined solutions of the corresponding radial equations. We shall give one of them in this section. Another one will be given in § 20.

The expression we are seeking serves to compare the phase shifts δ_l and $\hat{\delta}_l$ corresponding respectively to the potentials $V(r)$ and $\hat{V}(r)$ at the same energy. We use the notation of § 8 and put $\hat{U} = 2m\hat{V}/\hbar^2$. y_l is the regular solution of eq. (X.29) whose asymptotic form is given by expression (X.37). Likewise, we designate by $\hat{y}_l$ the regular solution of the radial equation

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}r^2} + \left(\varepsilon - \hat{U} - \frac{l(l+1)}{r^2}\right)\right]\hat{y}_l = 0, \tag{X.71}$$

having for its asymptotic expression

$$\hat{y}_l \underset{r\to\infty}{\sim} \sin\left(kr - \tfrac{1}{2}l\pi + \hat{\delta}_l\right).$$

The Wronskian $W(y_l, \hat{y}_l)$ is zero at the origin and asymptotically approaches the limit

$$\lim_{r\to\infty} W(y_l, \hat{y}_l) = k \sin(\delta_l - \hat{\delta}_l).$$

According to the Wronskian theorem:

$$W(y_l, \hat{y}_l)\Big|_a^b = -\int_a^b \hat{y}_l(U - \hat{U})y_l \,\mathrm{d}r.$$

Letting a and b approach 0 and ∞, respectively, one has

$$\sin(\delta_l - \hat{\delta}_l) = -\frac{2m}{\hbar^2 k}\int_0^\infty \hat{y}_l(V - \hat{V})y_l \,\mathrm{d}r. \tag{X.72}$$

This important relation is valid for any form of the potentials V and $\hat{V}$, provided that they vanish at infinity more rapidly than $1/r$ and that they have no singularity as strong as $1/r^2$ at the origin.

If $\hat{V}=0$, one has $\hat{\delta}_l=0$, $\hat{y}_l=krj_l(kr)$; relation (X.72) in that special case is written:

$$\sin \delta_l = -\frac{2m}{\hbar^2}\int_0^\infty j_l(kr)\, V y_l r \, dr. \tag{X.73}$$

18. Dependence upon the Potential. Sign of the Phase Shifts

Equation (X.72) allows to draw certain conclusions concerning the effects on the phase shifts when one changes the potential. Indeed, if $\Delta V \equiv V-\hat{V}$ is infinitesimal, $\Delta\delta_l=\delta_l-\hat{\delta}_l$ is also infinitesimal; on the other hand, the difference between the solutions y_l and $\hat{y}_l$ may be neglected in the integral of the right-hand member of eq. (X.72), whence

$$\Delta\delta_l = -\frac{2m}{\hbar^2 k}\int_0^\infty y_l^2 \Delta V \, dr. \tag{X.74}$$

If the variation of the potential $\Delta V(r)$ has the same sign over the entire interval $(0, \infty)$ the variation of the phase shift $\Delta\delta_l$ has the opposite sign. Hence any increase of the potential (greater repulsion) decreases the phase shift, any decrease of the potential (greater attraction) increases the phase shift.

Heretofore the phase shift δ_l was defined to within $2n\pi$. In order to remove this ambiguity, we imagine a continuous modification of the potential from 0 up to $V(r)$; in this operation, the phase shift varies in a continuous manner from 0 up to a certain value δ_l; one can show that this value is independent of the path followed to go from zero potential to potential $V(r)$. It is this value we adopt henceforth as the definition of the phase shift.

Having established this, if the potential $V(r)$ is repulsive everywhere, one can go over from potential zero to potential $V(r)$ by adding up all the repulsive infinitesimal contributions. According to eq. (X.74), each of these contributions diminishes the phase shift; consequently δ_l is negative. Likewise, if $V(r)$ is attractive everywhere, δ_l is positive.

More generally:

$$\text{if } V(r) > \hat{V}(r) \text{ for any } r, \ \delta_l < \hat{\delta}_l;$$

$$\text{if } V(r) < \hat{V}(r) \text{ for any } r, \ \delta_l > \hat{\delta}_l.$$

19. The Born Approximation

In order to know the phase shift δ_l exactly, one must in principle integrate equation (X.29). However, if $V(r)$ is sufficiently small, the

regular solution y_l of this equation differs very little from the free spherical wave $krj_l(kr)$, and the phase shift δ_l is close to zero. One can then without much error replace y_l by the free wave relation in (X.73), which yields

$$\delta_l \approx -\frac{2m}{\hbar^2} k \int_0^\infty j_l^2(kr)\ V(r)\ r^2\ \mathrm{d}r. \tag{X.75}$$

This is the expression for the phase shift in "*Born approximation*".

The error is small if $V(r)$ is sufficiently small compared to $E-[l(l+1)\hbar^2/2mr^2]$ over most of the region of variation of r. One expects therefore that the Born approximation is good at high energies or, provided that $V(r)$ decreases sufficiently rapidly at infinity, for large values of l. In fact, expression (X.75) is the first term of an expansion in powers of V; it is therefore possible to estimate the error by roughly evaluating the term of immediately higher order. This question will be treated in Chapter XIX.

Starting from relation (X.72), one likewise obtains an approximate expression for $\delta_l-\hat{\delta}_l$, namely

$$\delta_l-\hat{\delta}_l \approx -\frac{2m}{\hbar^2 k}\int \hat{y}_l^2(V-\hat{V})\,\mathrm{d}r. \tag{X.76}$$

This "generalized Born formula" is useful when one knows the regular solution $\hat{y}_l$ of the radial equation for a potential $\hat{V}$ which differs only slightly from the potential V. It then enables us to obtain δ_l to a good approximation, without the necessity of solving exactly the radial equation for the potential V [1]).

20. Effective Range Theory. The Bethe Formula

The formulae of § 17 allow us to study the variation suffered by the phase shift when one modifies the potential while keeping the

---

[1]) One can also make use of this formula to study the effect of the tail of the potential V. It suffices to take for the potential $\hat{V}$ the potential

$$\hat{V} = \begin{cases} V(r) & \text{if } r < r_0 \\ 0 & \text{if } r > r_0 \end{cases}$$

r_0 having a suitably chosen value. $\hat{V}$ is a potential of finite range and possesses all the properties of potentials of this type (Sections III and IV). $V-\hat{V}$ is the tail of potential V; its effect can be evaluated by means of formula (X.76), if it is small.

energy constant. The formula we establish in this section concerns the variation of the phase shift as a function of the energy. It is particularly useful in the limit of low energies for a short range potential.

Denote by u one of the regular solutions of eq. (X.29). We do not specify its normalization for the moment. Let $\hat{u}$ be the (irregular) solution of eq. (X.71) corresponding to the same value of the energy and having the same asymptotic form as u, with the same normalization. Now consider two different energies E_1 and E_2; we label by the index 1 the functions and quantities related to the energy E_1, and by the index 2 those related to energy E_2. According to the Wronskian theorem [eq. (III.27)],

$$W(u_1, u_2)\Big|_a^b = (\varepsilon_1 - \varepsilon_2) \int_a^b u_1 u_2 \, dr$$

and an analogous relation for the $\hat{u}$, whence

$$W(\hat{u}_1, \hat{u}_2) - W(u_1, u_2)\Big|_a^b = (\varepsilon_1 - \varepsilon_2) \int_a^b (\hat{u}_1\hat{u}_2 - u_1 u_2) \, dr.$$

As $b \to \infty$, since the u and $\hat{u}$ have the same asymptotic form, the integral of the right-hand side converges, and the difference of the Wronskians evaluated at b tends to zero. Since furthermore $\lim_{a\to 0} W(u_1, u_2) = 0$, the foregoing relation may be written in the limit where $b \to \infty$ and $a \to 0$,

$$\lim_{a\to 0} [W(\hat{u}_1, \hat{u}_2) + (\varepsilon_1 - \varepsilon_2) \int_a^\infty (\hat{u}_1\hat{u}_2 - u_1 u_2) \, dr] = 0. \qquad \text{(X.77)}$$

With an appropriate choice of normalization for u, one can put this fundamental relation into a form in which the value of the difference $\delta - \hat{\delta}$ at energies E_1 and E_2 enters explicitly.

We shall restrict this study to the case of s waves ($l = 0$) [1]. Moreover, we suppose that $\hat{V} = 0$ and denote by v_1, v_2 the functions $\hat{u}_1$, $\hat{u}_2$ in this special case. Let us fix the normalization of the u by the condition $v(0) = 1$:

$$v = \cos kr + \cot \delta \sin kr.$$

---

[1] When $l \neq 0$, u_1 and u_2 have a singularity in $(1/r)^l$ at the origin. In formula (X.77), the terms $W(\hat{u}_1, \hat{u}_2)$ and $(\varepsilon_1 - \varepsilon_2) \int \hat{u}_1 \hat{u}_2 dr$ approach infinity as $(1/a)^{2l-1}$; however, their sum tends toward a finite limit.

Formula (X.77) is then written

$$W(v_2, v_1)\Big|_{a=0} \equiv k_1 \cot \delta_1 - k_2 \cot \delta_2 = (\varepsilon_1 - \varepsilon_2) \int (v_1 v_2 - u_1 u_2)\, dr. \qquad (X.78)$$

Provided that $V(r)$ tends to zero sufficiently rapidly as $r \to \infty$ so that the integral of the right-hand side converges, (X.78) holds in the limit where $\varepsilon_2 \to 0$. Let us designate by u_0, v_0 the functions u, v corresponding to zero energy. Note that

$$v_0 = 1 - \frac{r}{a} \quad \text{and} \quad \lim_{\varepsilon \to 0} k \cot \delta = -\frac{1}{a},$$

where a is the scattering length defined by equation (X.47). Upon setting $\varepsilon_1 = \varepsilon$ and $\varepsilon_2 = 0$ in relation (X.78) (cf. Fig. X.7), we have (*Bethe formula*)

$$k \cot \delta = -\frac{1}{a} + \varepsilon \int (v v_0 - u u_0)\, dr. \qquad (X.79)$$

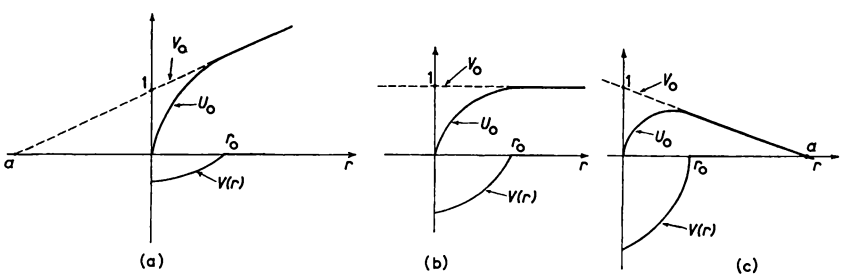


Fig. X.7. s-wave functions of zero energy occurring in effective range theory for increasing depths of the potential of limited range $V(r) = \varpi W(r)$ [ϖ = depth parameter: $\varpi = 1$ yields the depth which is just sufficient for a bound state to occur]

(*a*) $\varpi < 1$ ($a < 0$) (*b*) $\varpi = 1$ ($a = \infty$) (*c*) $\varpi > 1$ ($a > 0$).

N.B. a is a decreasing function of ϖ having a vertical asymptote at each value of ϖ for which there exists a bound state of energy 0.

This relation is exact. It is useful when the integral of the right-hand side varies slowly as a function of energy. This happens when $V(r)$ is a short range potential of the type encountered in nuclear physics, that is to say, when one can divide configuration space into an internal region ($r < r_0$, $kr_0 \ll 1$) for which $|V| \gg E$, and an external region ($r > r_0$) where the potential V is negligible. The contribution to the integral then comes essentially from the internal region, where one can replace, without committing a large error, u by u_0 and v by v_0,

since one strictly has $u = u_0 = 0$ and $v = v_0 = 1$ at the origin; also, the functions u and u_0 have practically the same relative curvature ($u''/u \approx 2mV/\hbar^2$) in this whole region (Fig. X.7). One thus has to a very good approximation:

$$k \cot \delta \approx -\frac{1}{a} + \varepsilon \int_0^\infty (v_0^2 - u_0^2)\, dr. \tag{X.80}$$

The quantity $r_{\text{eff.}} \equiv 2\int_0^\infty (v_0^2 - u_0^2)\, dr$ is a characteristic parameter of the potential $V(r)$, the so-called *effective range*.

The two terms of the right-hand side of (X.80) are the first two terms of the expansion of $k \cot \delta$ in a series of powers of the energy. To write down the terms of higher order, it is necessary to expand u and v in power series of ε and to substitute the expressions thus obtained in the integral of the right-hand side of (X.79) [1]. According to the argument given above, these expansions are rapidly converging in the internal region; consequently, the convergence of the expansion of $k \cot \delta$ as a function of the energy is likewise very rapid.

EXERCISES AND PROBLEMS

1. Consider the scattering of a particle of wavelength $\lambda\!\!\!^{-}$ by a potential $V(r)$ tending to 0 more rapidly than $1/r$ as $r \to \infty$. Let $f(\Omega)$ be the scattering amplitude in the direction $\Omega \equiv (\theta, \varphi)$. Show that

$$\sigma_{\text{tot.}} \equiv \int |f(\Omega)|^2\, d\Omega = 4\pi\lambda\!\!\!^{-}\, \text{Im}\, f(0),$$

where $f(0)$ stands for the forward scattering amplitude ($\theta = 0$) [*Bohr–Peierls–Placzek relation*].

2. By bombarding nucleus A by nucleus a one can form nuclei b and B: $a + A \to b + B$. In the laboratory system the target A is at rest. Let m_a, m_A, m_b, m_B be the masses of the particles involved in this reaction. We make the non-relativistic approximation: $m_a + m_A = m_b + m_B$. Let E_i and E_f be the respective total kinetic energies of the initial state ($a + A$) and of the final state ($b + B$) in the center-of-mass system, and let θ and θ_1 be the emission angles of particle b in the center-of-mass system and the laboratory system, respectively. Show that θ_1 is given as a function of θ by relation (X.24), provided that τ represents the ratio of the velocity V of the center of mass, to the velocity v_b of particle b in the center-of-mass system, namely:

$$\tau = \frac{V}{v_b} = \left[\frac{m_a m_b}{m_A m_B}\frac{E_i}{E_f}\right]^{\frac{1}{2}}$$

---

[1] Cf. G. Chew and M. L. Goldberger, Phys. Rev. **75** (1949) 1637; H. A. Bethe, Phys. Rev. **76** (1949) 38.

3. Prove the following relations between the quantities τ_l, v_l, and $q_l^{(+)}$ introduced in § 10 [definition (X.41)]:

$$v_l = -\frac{\mathrm{d}\tau_l}{\mathrm{d}\xi}, \qquad \operatorname{Im} q_l^{(+)} = \xi v_l, \qquad \operatorname{Re} q_l^{(+)} = -\frac{1}{2}\frac{\xi}{v_l}\frac{\mathrm{d}v_l}{\mathrm{d}\xi}.$$

4. Show that in the WKB approximation the phase shift δ_l is given by the formula

$$\delta_l = \lim_{R\to\infty}\left[\int_a^R \sqrt{k^2 - U(r) - \frac{l(l+1)}{r^2}}\,\mathrm{d}r - \int_{a_0}^R \sqrt{k^2 - \frac{l(l+1)}{r^2}}\,\mathrm{d}r\right].$$

Take the same definition of δ_l, k and $U(r)$ as in § 8. The lower limits of the integrals, a and a_0 are the zeros of the corresponding integrands [if $k^2 - U(r) - l(l+1)/r^2$ has several zeros, a is the largest amongst them]. Discuss the conditions of validity of this approximation [in order that the method apply to small values of l, one must, following the recommendation of Langer (cf. Problem IX.6), replace $l(l+1)$ by $(l+\frac{1}{2})^2$ in the two integrands].

5. Apply the theory of the effective range to p-wave scattering. Show that it gives the expansion of $k^3 \cot \delta_1$ as a function of the energy, and give the expression for the first two terms of this expansion ($\delta_1 = p$-wave phase shift).

CHAPTER XI

THE COULOMB INTERACTION

1. Introduction

Let r be the distance between two particles of electric charge $Z_1 e$ and $Z_2 e$, respectively; the electrostatic interaction potential of these two particles is the Coulomb potential

$$V(r) = \frac{Z_1 Z_2 e^2}{r}.$$

Let $\mathbf{p}_1$, $\mathbf{p}_2$ be the momenta, m_1, m_2 the masses, and $\mathbf{r}_1$, $\mathbf{r}_2$ the respective positions of these two particles ($\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$). If their interaction is purely of the Coulomb type, their motion is governed by the Hamiltonian

$$\frac{\mathbf{p}_1^2}{2m_1} + \frac{\mathbf{p}_2^2}{2m_2} + \frac{Z_1 Z_2 e^2}{r}.$$

The center-of-mass motion can be separated from the motion of the relative particle, using the method outlined in Chapter IX. The motion of the relative particle is governed by the Hamiltonian

$$H \equiv \frac{\mathbf{p}^2}{2m} + \frac{Z_1 Z_2 e^2}{r} \tag{XI.1}$$

where m is the reduced mass:

$$m = \frac{m_1 m_2}{m_1 + m_2}. \tag{XI.2}$$

The treatment of a quantum system of two particles with Coulomb interaction therefore reduces to that of a particle in the central potential $Z_1 Z_2 e^2/r$.

Because of its slow decrease for large values of r, a certain number of the properties of central potentials obtained in Chapters IX and X do not apply to the Coulomb potential. In scattering problems especially, the asymptotic behavior of the stationary solutions is less simple than in the case of potentials with shorter range, and the definition of phase shifts must be changed accordingly; actually, the

usefulness of the treatment of scattering by separation of angular and radial variables is not as obvious since the expansion of the scattering amplitude in spherical harmonics converges very slowly.

On the other hand, the solution of the Schrödinger equation of a particle in a Coulomb potential can always be reduced to the solution of a Laplace differential equation, an equation that is well known in the theory of functions. Therefore, the most interesting quantities concerning the Coulomb interaction, such as the energy spectrum of bound states and the scattering cross section, can be calculated exactly.

This chapter consists of two sections. The first is devoted to the study of bound states of the hydrogen atom, a study that is easily extended to hydrogen-like atoms, and more generally to systems of two particles with mutual attraction of the form $1/r$. Section II deals with problems of Coulomb scattering.

I. THE HYDROGEN ATOM

2. Schrödinger Equation of the Hydrogen Atom

The simplest system of two bodies with a Coulomb interaction is the hydrogen atom. The two particles, electron and proton, are subject to the attractive potential $-e^2/r$. The reduced mass m of the electron-proton system is slightly smaller than the electron mass: $(m_e - m)/m_e \approx 5 \times 10^{-4}$. Let E be the energy of the electron-proton system in the center-of-mass system. The wave function $\psi(\boldsymbol{r})$ of the relative particle is a solution of the Schrödinger equation

$$\left[-\frac{\hbar^2}{2m} \triangle - \frac{e^2}{r} \right] \psi(\boldsymbol{r}) = E\psi(\boldsymbol{r}). \tag{XI.3}$$

The nature of the regular solutions of this equation is easily revealed if one carries out the separation of angular and radial variables. Thus the eigensolution of energy E and angular momentum (lm) is the function

$$Y_l{}^m(\theta, \varphi) \frac{y_l(r)}{r},$$

where y_l is the solution of the radial equation [cf. eq. (IX.2)] which vanishes at the origin

$$y_l'' + \left[\varepsilon + \frac{2m}{\hbar^2} \frac{e^2}{r} - \frac{l(l+1)}{r^2} \right] y_l = 0. \tag{XI.4}$$

We have put

$$\varepsilon = \frac{2mE}{\hbar^2}. \tag{XI.5}$$

If $E>0$, this solution has an oscillatory behavior at infinity and can be accepted as eigensolution for any E. It represents an unbound state and enters into the construction of stationary states of electron-proton collisions at energy E.

If $E<0$, the asymptotic form of the solution which is regular at the origin is a linear combination of exponentials $\exp(\varkappa r)$ and $\exp(-\varkappa r)$, where

$$\varkappa = (-\varepsilon)^{\frac{1}{2}} = \frac{(-2mE)^{\frac{1}{2}}}{\hbar}. \tag{XI.6}$$

For this solution to be an acceptable eigensolution, the coefficient in front of $\exp(\varkappa r)$ must vanish: this happens only for certain discrete values of E. These values are the energies of the discrete spectrum of the hydrogen atom, and the corresponding wave function represents one of the possible bound states of this atom.

The purpose of this section is to study the bound states of the hydrogen atom. The results of this study may easily be extended to hydrogen-like atoms (He^+, Li^{++}, etc.) in which the proton is replaced by a heavier nucleus. Let M_A be the mass of this nucleus, M_p the mass of the proton. The reduced mass of the hydrogen-like atom:

$$m' = \frac{m_e M_A}{m_e + M_A},$$

differs slightly from that of the hydrogen atom

$$m = \frac{m_e M_p}{m_e + M_p}.$$

On the other hand, if Ze is the charge of the nucleus, the Coulomb attraction is Ze^2/r and not e^2/r. All formulae relating to the hydrogen atom apply to the hydrogen-like atom provided that one replaces m by m', and e^2 by Ze^2.

3. Order of Magnitude of the Binding Energy of the Ground State

Denote by r_0 the "radius" of the atom in its ground state. We understand thereby that the wave function is practically entirely

concentrated inside a sphere of radius r_0, in other words that the probability of presence of the electron at a distance r from the proton is very small when $r > r_0$, whereas it has a non-negligible value if $r < r_0$. A very crude model consists in taking for the approximate form of the probability density a uniform distribution over a sphere of radius r_0.

Clearly, the average value of the potential energy is smaller (in algebraic value) the smaller r_0: it is of order $(-e^2/r_0)$. The mean value of the kinetic energy, on the other hand, is larger the smaller r_0. Indeed, if the electron is localized inside a sphere of radius r_0, the uncertainty relations impose a lower limit upon the value of its momentum; the root-mean-square deviation of the momentum cannot be smaller than $\hbar/r_0$, hence there is a minimum mean kinetic energy $\hbar^2/2mr_0^2$. The total energy is thus at least equal to the sum of these two quantities:

$$\frac{\hbar^2}{2mr_0^2} - \frac{e^2}{r_0},$$

whose minimum is attained when $r_0 = \hbar^2/me^2$. One expects the value of this minimum

$$E_1 = -\frac{1}{2}\frac{me^4}{\hbar^2} \qquad (= -13.5 \text{ eV}) \tag{XI.7}$$

to be of the order of the energy of the ground state, and the value of the corresponding radius:

$$a = \frac{\hbar^2}{me^2} \qquad (= 0.529 \times 10^{-8} \text{ cm}) \tag{XI.8}$$

to give the order of magnitude of the extension of the wave function in the ground state.

In fact — and this is a mere coincidence — the energy of the ground state is rigorously given by eq. (XI.7). The length a is called the *Bohr radius* or "radius of the hydrogen atom".

4. Solution of the Schrödinger Equation in Spherical Coordinates

To solve the Schrödinger equation, we go over into spherical polar coordinates. As is the case for any central potential, the angular and radial variables separate in this coordinate system and the problem reduces to a search for regular solutions of the radial equation (XI.4).

It is also possible to solve the Schrödinger equation in parabolic coordinates, because the variables likewise separate in that coordinate system. We merely mention this characteristic property of the Coulomb potential for future reference, and limit ourselves to a treatment of the problem in spherical polar coordinates.

If one makes the change of variable

$$x = 2\varkappa r \tag{XI.9}$$

equation (XI.4) depends only upon the dimensionless parameter

$$\nu = \frac{1}{\varkappa a} = \frac{e^2}{\hbar c}\sqrt{\frac{mc^2}{-2E}}; \tag{XI.10}$$

$\varkappa$ and a were defined by equations (XI.6) and (XI.8), respectively. Equation (XI.4) is equivalent to the equation

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}x^2} - \frac{l(l+1)}{x^2} + \frac{\nu}{x} - \frac{1}{4}\right] y_l = 0 \tag{XI.11}$$

y_l is the solution which goes as x^{l+1} at the origin. For x very large, it increases exponentially except for certain particular values of ν where it behaves as $\exp(-\frac{1}{2}x)$. We intend to determine these special values of ν and their corresponding eigensolutions.

To this effect, we perform the change of function

$$y_l = x^{l+1}\,\mathrm{e}^{-\frac{1}{2}x}\,v_l(x),$$

which yields

$$\left[x\frac{\mathrm{d}^2}{\mathrm{d}x^2} + (2l+2-x)\frac{\mathrm{d}}{\mathrm{d}x} - (l+1-\nu)\right] v_l = 0. \tag{XI.12}$$

This differential equation is a Laplace equation (cf. Appendix B, § 1). To within a constant, only one solution is finite at the origin; all others have a singularity in $(1/x)^{2l+1}$ there. This regular solution is the confluent hypergeometric series:

$$F(l+1-\nu|2l+2|x) \equiv \sum_{p=0}^{\infty} \frac{\Gamma(l+1+p-\nu)}{\Gamma(l+1-\nu)} \frac{(2l+1)!}{(2l+1+p)!} \frac{x^p}{p!}. \tag{XI.13}$$

To show this, it is actually sufficient to look for the solution of equation (XI.12) represented by the Taylor series expansion at the origin,

$$v_l(x) = 1 + a_1 x + a_2 x^2 + \ldots + a_p x^p + \ldots$$

Substituting this expansion in eq. (XI.12), one can write the left-hand side in the form of an expansion in a power series of x. All the coefficients of this expansion must vanish; hence

$$\begin{aligned} (2l+2)a_1 &= (l+1-\nu) \\ 2(2l+3)a_2 &= (l+2-\nu)a_1 \\ &\vdots \\ p(2l+1+p)a_p &= (l+p-\nu)a_{p-1}, \end{aligned}$$

whence

$$a_p = \frac{(p+l-\nu)\,(p-1+l-\nu)\,\ldots\,(1+l-\nu)}{(p+2l+1)\,(p-1+2l+1)\,\ldots\,(1+2l+1)} \times \frac{1}{p!}.$$

a_p is actually the coefficient of x^p in the hypergeometric series (XI.13).

In general, the series (XI.13) is an infinite series and behaves as $x^{-l-1-\nu} \exp(x)$ [eqs. (B.9) to (B.11)], for large x. Then y_l behaves in the asymptotic region as $x^{-\nu} \exp(\frac{1}{2}x)$; it cannot be an eigensolution.

However, for certain privileged values of ν the coefficients all vanish from a certain order on, and the hypergeometric series reduces to a polynomial. For this to happen, $l+1-\nu$ must be a negative integer, or zero, namely

$$\nu = n = l+1+n' \qquad (n' = 0, 1, 2, \ldots, \infty). \tag{XI.14}$$

In that case the hypergeometric series reduces to a polynomial of degree n', the radial function behaves as $x^n \exp(-\frac{1}{2}x)$ for $x \to \infty$, and the regular solution of the Schrödinger equation is acceptable as eigensolution.

The quantum condition (XI.14) thus gives the energy levels of the bound states of angular momentum (lm). Each of them is defined by a particular value of the integer n'. The wave function of the corresponding bound state is the (unnormalized) wave function constructed with the radial solution

$$y_l = x^{l+1}\, e^{-\frac{1}{2}x} \sum_{p=0}^{n'} (-)^p \frac{n'!\,(2l+1)!}{(n'-p)!\,(2l+1+p)!\,p!}\, x^p. \tag{XI.15}$$

The polynomial of degree n' represented by the summation of the right-hand side is, to within a constant, the associated Laguerre polynomial $L_{n'}^{2l+1}(x)$, the definition and main properties of which are given in Appendix B (§ 2).

5. Energy Spectrum. Degeneracy

Replacing in eq. (XI.14) the parameter ν by its expression (XI.10) as a function of energy, one obtains the energy spectrum of the states of angular momentum l,

$$E_{ln'} = -\left(\frac{e^2}{\hbar c}\right)^2 \frac{mc^2}{2(l+1+n')^2}. \tag{XI.16}$$

n', the radial quantum number, is equal to the number of nodes of the radial part of the wave function. The spectrum contains a denumerably infinite set of levels since n' may take on all integral values from 0 to $+\infty$. When $n' \to \infty$, these levels become more and more closely spaced and tend to $E=0$ in the limit, at which point the continuous spectrum begins.

This circumstance is characteristic of long-range potentials. Short-range potentials such as square wells, on the other hand, give a finite (and sometimes zero) number of bound states. One can show very generally that the set of energy levels is denumerably infinite (with a point of accumulation at the value 0) when the potential approaches zero asymptotically through negative values, and less rapidly than $1/r^2$:

$$r^2 V \underset{r\to\infty}{\to} -\infty;$$

one can also show that their number is finite – and possibly zero – if the potential asymptotically approaches zero more rapidly than $1/r^2$.

The ensemble of the spectra belonging to different possible values of l ($l=0, 1, 2, \ldots, \infty$) represents the complete spectrum of the hydrogen atom according to the Schrödinger theory. This complete spectrum is thus formed of the set of numbers $E_{ln'}$ defined by eq. (XI.16), l and n' taking all possible integral, non-negative values. Note that these quantities depend only upon the sum $l+n'$ or, what is equivalent, upon the "principal quantum number"

$$n = l+n'+1.$$

One has

$$E_n = -\left(\frac{e^2}{\hbar c}\right)^2 \frac{mc^2}{2n^2} \qquad (n=1, 2, \ldots, \infty). \tag{XI.17}$$

For each energy E_n, that is to say for each value of the integer n,

the angular momentum may take on all integral values from 0 up to $n-1$. The order of degeneracy of the level E_n is therefore

$$\sum_{l=0}^{n-1}(2l+1) = n(n-1)+n = n^2.$$

The n^2-dimensional subspace of the eigenfunctions is spanned by n^2 functions, each corresponding to a state of given angular momentum (lm), the "azimuthal quantum number" taking on the n values

$$l = 0, 1, 2, \ldots, n-1,$$

and the "magnetic quantum number" taking on the $(2l+1)$ values

$$m = -l, -l+1, \ldots, +l.$$

According to the tradition of spectroscopy, the various eigenstates thus defined are designated by the positive integer n followed by a letter (s, p, d, f, g, ...) indicating the value of l in conformity with the definition stated in the preceding chapter. The quantum number m which indicates the orientation of the system is simply not mentioned. Thus the ground state is a 1s state, the first-excited state is four-fold degenerate and contains one 2s state and three 2p states; the second-excited state is nine-fold degenerate and contains one 3s state, three 3p states and five 3d states; and so forth (Fig. XI.1).

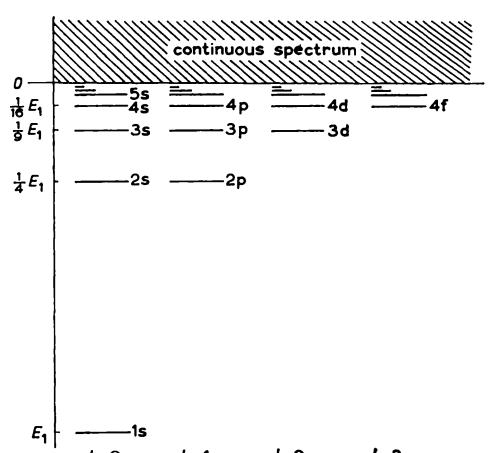


Fig. XI.1. Spectrum of the hydrogen atom.

This spectrum is just the one predicted by the Old Quantum Theory; its excellent agreement with the experimental spectrum was already pointed out. To be more precise, the theory correctly accounts for the position of the spectral lines but not for their fine structure. Its essential shortcoming is to be a non-relativistic theory. The relativistic effects in the positions of the levels are of the order v^2/c^2, or about E_n/mc^2. The relativistic corrections are thus of the order 10^{-4} to 10^{-5}. On the other hand, the Schrödinger theory does not take the electron spin into account: the spin is an internal degree of freedom without classical analogue; we shall have occasion to discuss it in Ch. XIII. The analysis of the fine structure of the hydrogen atom will be taken up again in Ch. XX during the discussion of the Relativistic Quantum Mechanics of the electron.

6. The Eigenfunctions of the Bound States

The eigenfunctions belonging to the energy level E_n are linear combinations of n^2 linearly independent functions. The investigation of § 4 furnishes us n^2 orthogonal eigenfunctions, namely those which correspond to a well-defined value of the angular momentum. Thus the wave function of the quantum state (nlm) is written

$$\psi_{nlm} = a^{-3/2} N_{nl} F_{nl}\left(\frac{2r}{na}\right) Y_l^m(\theta, \varphi) \tag{XI.18}$$

with

$$F_{nl}(x) = x^l e^{-\frac{1}{2}x} L_{n-l-1}^{2l+1}(x). \tag{XI.18a}$$

N_{nl} is a normalization constant. One calculates the norm of ψ_{nlm} using the generating function of the Laguerre polynomials [eq. (B.15)]. This norm is equal to unity if one takes

$$N_{nl} = \frac{2}{n^2}\sqrt{\frac{(n-l-1)!}{[(n+l)!]^3}}. \tag{XI.18b}$$

The mean values of successive powers of r in the quantum state (nlm) are of interest. Their calculation will not be undertaken here (Problem XI.1). The results are given in Appendix B (§ 3). One has, in particular

$$\langle r\rangle_{nl} = \tfrac{1}{2}a[3n^2 - l(l+1)]. \tag{XI.19}$$

The electron is thus further removed – on the average – from the proton, the larger n. For the ground state, one finds $\langle r\rangle_{1s} = \frac{3}{2}a$ in accord with the rough predictions of § 3.

When l takes its maximum value $(n-1)$, the wave function assumes an especially simple form: it is the product of $Y_l^m(\theta, \varphi)$ and the radial function

$$[(2n)!]^{-\frac{1}{2}} \left(\frac{2}{na}\right)^{\frac{1}{2}} \left(\frac{2r}{na}\right)^{n-1} e^{-r/na}.$$

The mean value of r in that state is

$$\langle r \rangle = \frac{1}{(2n)!} \left(\frac{2}{na}\right)^3 \int_0^\infty \left(\frac{2r}{na}\right)^{2n-2} e^{-2r/na} r^3 \, dr$$
$$= n(n+\tfrac{1}{2})\, a,$$

in accord with the more general formula indicated above. An analogous calculation yields

$$\langle r^2 \rangle = n^2(n+\tfrac{1}{2})\,(n+1)\, a^2,$$

whence one obtains the expression for the root-mean-square radial deviation

$$\Delta r = \sqrt{\langle r^2 \rangle - \langle r \rangle^2} = \tfrac{1}{2} na \sqrt{2n+1} = \frac{\langle r \rangle}{\sqrt{2n+1}}.$$

For very large values of n, $\Delta r / \langle r \rangle$ becomes very small and the electron remains practically localized in the vicinity of a sphere of radius n^2a, whereas the energy of the level, $-\frac{1}{2}e^2/n^2a$, is the same as that of a classical electron describing a circular orbit of radius n^2a.

We verify in this particular example the general correspondence rule according to which one returns to the classical laws of motion in the limit of very large quantum numbers. In order to compare the quantum theory and the classical theory in detail, one would have to study the motion of wave packets. We shall not do so here. We merely note that the states of maximum l $(l=n-1)$ correspond to the classical circular orbits, a fact that is to be compared with the result of the Old Quantum Theory according to which the eccentricity of the quantized orbits is equal to $\sqrt{1-(l^2/n^2)}$ and vanishes when l takes on its maximum value (cf. Ch. I, § 15).

II. COULOMB SCATTERING

7. The Coulomb Scattering Wave

After reduction to the center of mass, the Schrödinger equation of the collision problem of two particles with Coulomb interaction is written, using the notation of § 1:

$$\left[-\frac{\hbar^2}{2m}\triangle+\frac{Z_1 Z_2 e^2}{r}\right]\psi(\boldsymbol{r})=E\psi(\boldsymbol{r}) \tag{XI.20}$$

where E is the energy in the center-of-mass system. The scattering cross section is related to the asymptotic behavior of the eigensolutions of positive energy of eq. (XI.20). Let us put [1])

$$E=\frac{\hbar^2 k^2}{2m}=\tfrac{1}{2}mv^2 \tag{XI.21}$$

$$\gamma=\frac{Z_1 Z_2 e^2}{\hbar v}; \tag{XI.22}$$

equation (XI.20) may also be written

$$\left(\triangle+k^2-\frac{2\gamma k}{r}\right)\psi(\boldsymbol{r})=0. \tag{XI.23}$$

This equation has a regular solution of the form

$$\mathrm{e}^{\mathrm{i}kz}f(r-z). \tag{XI.24}$$

Indeed, if we substitute this expression in eq. (XI.23) and put $u=r-z$, we obtain the differential equation

$$\left[u\frac{\mathrm{d}^2}{\mathrm{d}u^2}+(1-\mathrm{i}ku)\frac{\mathrm{d}}{\mathrm{d}u}-\gamma k\right]f(u)=0,$$

or else, putting

$$v=\mathrm{i}ku=\mathrm{i}k(r-z),$$

$$\left[v\frac{\mathrm{d}^2}{\mathrm{d}v^2}+(1-v)\frac{\mathrm{d}}{\mathrm{d}v}+\mathrm{i}\gamma\right]f(v)=0.$$

---

[1]) γ is the analogue of the parameter ν introduced in the problem of the hydrogen atom. If one sets $a=\hbar^2/Z_1Z_2me^2$, one has $\gamma=1/ka$ [cf. eq. (XI.10)].

This is an equation of the Laplace type whose regular solution at the origin is the confluent hypergeometric series $F(-\mathrm{i}\gamma|1|v)$. The Schrödinger equation thus actually has a regular solution of the form (XI.24), the function

$$\psi_c = A\, e^{ikz} F[-\mathrm{i}\gamma\,|1|\,\mathrm{i}k(r-z)]. \tag{XI.25}$$

A is an adjustable constant.

In keeping with the study made in Appendix B (§ 1), the hypergeometric series occurring in equation (XI.25) is the sum of two functions whose asymptotic forms for large values of $|v| = 2kr\sin^2\frac{1}{2}\theta$ are given by equations (B.10) and (B.11), respectively. Let us adopt the notation of Appendix B and put

$$\psi_i = A\, e^{ikz} W_1(-\mathrm{i}\gamma\,|1|\,\mathrm{i}ku) \tag{XI.26}$$

$$\psi_s = A\, e^{ikz} W_2(-\mathrm{i}\gamma\,|1|\,\mathrm{i}ku). \tag{XI.27}$$

We have

$$\psi_c = \psi_i + \psi_s. \tag{XI.28}$$

The functions ψ_i and ψ_s are (irregular) solutions of eq. (XI.20).

Upon taking

$$A = \Gamma(1+\mathrm{i}\gamma)\, e^{-\frac{1}{2}\pi\gamma} \tag{XI.29}$$

one obtains the following asymptotic forms for ψ_i and ψ_s:

$$\psi_i \underset{|r-z|\to\infty}{\sim} e^{\mathrm{i}[kz+\gamma \ln k(r-z)]}\left[1 + \frac{\gamma^2}{\mathrm{i}k(r-z)} + \ldots\right] \tag{XI.30}$$

$$\psi_s \underset{|r-z|\to\infty}{\sim} -\frac{\gamma}{k(r-z)}\frac{\Gamma(1+\mathrm{i}\gamma)}{\Gamma(1-\mathrm{i}\gamma)}\, e^{\mathrm{i}[kr-\gamma \ln k(r-z)]}\left[1 + \frac{(1+\mathrm{i}\gamma)^2}{\mathrm{i}k(r-z)} + \ldots\right]. \tag{XI.31}$$

Since $z = r\cos\theta$, the first term of the asymptotic expansion of ψ_s can also be written

$$\psi_s \underset{|r-z|\to\infty}{\sim} \frac{1}{r}\exp\left[\mathrm{i}(kr - \gamma \ln 2kr)\right] f_c(\theta) \tag{XI.32}$$

with

$$f_c(\theta) = -\frac{\gamma}{2k\sin^2\frac{1}{2}\theta}\exp\left[-\mathrm{i}\gamma \ln(\sin^2\tfrac{1}{2}\theta) + 2\mathrm{i}\sigma_0\right] \tag{XI.33}$$

$$\sigma_0 = \arg\Gamma(1+\mathrm{i}\gamma). \tag{XI.34}$$

8. The Rutherford Formula

The wave ψ_c represents the stationary state of collision of a particle of incident momentum $\hbar\boldsymbol{k}$ directed along the z axis. In the case of a potential tending to zero at least as fast as $1/r^2$ for $r \to \infty$ we know that the same stationary state of collision is represented by the wave function which has the asymptotic form $\exp(ikz) + f(\theta)\exp(ikr)/r$; the latter can be interpreted as the sum of the incident plane wave and the outgoing scattered wave. Likewise, the wave ψ_c can be put in the form of a sum of two terms ψ_i, ψ_s whose respective asymptotic forms resemble a plane wave and an outgoing wave.

However, even at infinitely large distances from the origin, ψ_i cannot be likened to a plane wave by reason of the presence of the factor $\exp[i\gamma \ln k(r-z)]$; in other words, the Coulomb field has such a long range that it affects the incident wave even in the asymptotic region. In spite of this fact, for negative and very large values of z, ψ_i is a wave of unit density whose current density

$$\boldsymbol{j}_i = \frac{\hbar}{2im}\left[\psi_i^*(\nabla\psi_i) - \psi_i(\nabla\psi_i)^*\right]$$

is directed along the axis Oz and is equal to $v \equiv \hbar k/m$ (the logarithmic term introduces corrections of the order $1/r$ which one can neglect). This justifies the interpretation of ψ_i as incident wave.

Likewise, the radial dependence of the wave ψ_s never (not even for very large values of r) approaches the form $\exp(ikr)/r$ characteristic of outgoing waves, but the more complex expression $\exp[i(kr - \gamma \ln 2kr)]/r$. Nevertheless ψ_s behaves in the asymptotic region (except along the positive z axis where the separation into incident wave and scattered wave is meaningless) like a scattered wave, since the current density $\boldsymbol{j}_s$ calculated with this wave is actually directed radially and in the direction of increasing r; the effect of the factor $\exp(-i\gamma \ln 2kr)$ can be neglected to lowest order in $1/r$; in this approximation, the wave ψ_s is a wave of density $|f_c(\theta)|^2/r^2$ and of current density $v|f_c(\theta)|^2/r^2$.

Forming the ratio of the current density scattered into the solid angle $(\Omega, \Omega + d\Omega)$, to the incident current density, one obtains the differential scattering cross section

$$\sigma_c(\Omega) = |f_c(\theta)|^2. \tag{XI.35}$$

This formula is analogous to formula (X.2) for collisions with a potential of shorter range. The proof which was just given is open

to the same criticism as the one of Ch. X, § 3. However, a more rigorous proof modeled on the one of Ch. X, §§ 4 to 6 can be worked out easily.

$f_c(\theta)$ is called the Coulomb scattering amplitude. Its explicit form is given by eq. (XI.33). One derives the expression for the Coulomb scattering cross section:

$$\begin{aligned}\sigma_c(\Omega) &= \frac{\gamma^2}{4k^2 \sin^4 \frac{1}{2}\theta} \\ &= \left(\frac{Z_1 Z_2 e^2}{4E}\right)^2 \sin^{-4} \tfrac{1}{2}\theta.\end{aligned} \tag{XI.36}$$

The foregoing expression turns out to be identical to the classical Coulomb scattering cross section calculated in Chapter VI [eq. (VI.29)]. The classical Rutherford formula thus remains true even when the classical approximation ceases to be justified. This fact must be considered as an accident.

In expression (XI.36) of the cross section one notes the following noteworthy properties of Coulomb scattering:

(*i*) it depends solely upon the absolute value of the potential and not upon its sign;

(*ii*) the angular distribution is independent of the energy;

(*iii*) at a given angle, the cross section falls off like $1/E^2$ as the energy increases;

(*iv*) the total cross section is infinite: $\int \sigma_c(\Omega)\, d\Omega$ diverges at small angles.

This divergence is characteristic of the pure Coulomb field. In nature, such a field is never encountered; thus, in the scattering of a charged particle by an atomic nucleus, the Coulomb field due to the nuclei is gradually compensated at increasing distances by the field of the electron cloud, and the potential vanishes at distances sufficiently large compared to the radius of the atom. This screening modifies the scattered wave at small angles in such a way that the differential cross section no longer diverges. One can show that the modification due to screening is negligible at angles large compared to both $2\gamma/ka$ and $1/ka$ (a is the atomic radius). At the energies ordinarily used in nuclear physics, these limiting angles are so small that the screening effect can be completely neglected.

9. Decomposition into Partial Waves

The Schrödinger equation (XI.20) may be solved by separating angular and radial variables. This method is of no great interest for the treatment of pure Coulomb scattering since one has a more direct method at one's disposal. Moreover, since the potential has a long range, the expansion of the scattering amplitude $f_c(\theta)$ in spherical harmonics is known to converge very slowly. However, the decomposition into partial waves is useful in all problems where a short-range interaction is added to the Coulomb interaction proper. The presence of this additional interaction affects only the first few terms of the expansion of the amplitude $f(\theta)$ in spherical harmonics and, consequently, the expansion $f(\theta)-f_c(\theta)$ converges rapidly.

The separation of angular and radial variables has already been performed in connection with the hydrogen atom. With the present notation, equation (XI.4) is written

$$y_l'' + \left[k^2 - \frac{2\gamma k}{r} - \frac{l(l+1)}{r^2}\right] y_l = 0. \tag{XI.37}$$

To construct the solutions of this equation, one proceeds as in the problem of the hydrogen atom, by a change of function and of variable:

$$\begin{aligned} y_l &= e^{ikr}\,(kr)^{l+1}\,v_l \\ \xi &= -\,2ikr; \end{aligned} \tag{XI.38}$$

v_l is a solution of Laplace's equation [cf. eq. (XI.12)]

$$\left[\xi \frac{d^2}{d\xi^2} + (2l+2-\xi)\frac{d}{d\xi} - (l+1+i\gamma)\right] v_l = 0. \tag{XI.39}$$

From the asymptotic expansion [eqs. (B.10)–(B.11)] of the two irregular solutions $W_{1,2}(l+1+i\gamma|2l+2|\xi)$ one deduces the asymptotic form of the general solution of (XI.39) and, after a brief calculation, the asymptotic form of the general solution of (XI.37): it is a linear combination of the two exponentials

$$e^{\pm i(kr-\gamma \ln 2kr)}.$$

The solution regular at the origin of (XI.39) is the hypergeometric series $F(l+1+i\gamma|2l+2|\xi)$; it is the sum of two functions W_1 and W_2

[eq. (B.9)]. The corresponding solution of eq. (XI.37) asymptotically approaches a multiple of $\sin(kr-\gamma \ln 2kr-\frac{1}{2}l\pi+\sigma_l)$, with

$$\sigma_l = \arg \Gamma(l+1+i\gamma); \tag{XI.40}$$

σ_l is called the *Coulomb phase shift*. By definition, the regular Coulomb wave function $F_l(\gamma; kr)$ is the regular solution of (XI.37) whose asymptotic form is

$$F_l \underset{r\to\infty}{\sim} \sin(kr-\gamma \ln 2kr-\tfrac{1}{2}l\pi+\sigma_l). \tag{XI.41}$$

From the above we obtain

$$F_l(\gamma; kr) = c_l(\gamma)\, e^{ikr}\, (kr)^{l+1}\, F(l+1+i\gamma\,|\,2l+2\,|-2ikr), \tag{XI.42}$$

the constant $c_l(\gamma)$ being adjusted so that F_l satisfies (XI.41), namely

$$c_l = \frac{2^l\, e^{-\frac{1}{2}\pi\gamma}\, |\Gamma(l+1+i\gamma)|}{(2l+1)!}; \tag{XI.43}$$

F is a real function. One often denotes it under the name of regular spherical Coulomb function. It is a function of kr depending on the parameter γ.

One also defines "irregular spherical Coulomb functions". They are solutions irregular at the origin of eq. (XI.37). Those most commonly used are defined in Appendix B, § 5. Let us merely mention here the outgoing and incoming waves $u_l^{(+)}$ and $u_l^{(-)}$, of asymptotic form

$$u_l^{(\pm)} \underset{r\to\infty}{\sim} e^{\pm i(kr-\gamma \ln 2kr-\frac{1}{2}l\pi)}.$$

These functions are complex conjugates of each other, and one has

$$F_l = \operatorname{Im} e^{+i\sigma_l}\, u_l^{(+)}. \tag{XI.44}$$

10. Expansion of the Wave ψ_c in Spherical Harmonics

The Coulomb wave function ψ_c defined in Chapter XI, § 7

$$\psi_c \equiv e^{-\frac{1}{2}\pi\gamma}\, \Gamma(1+i\gamma)\, e^{ikz}\, F[-i\gamma\,|\,1\,|\,ik(r-z)] \tag{XI.45}$$

can be represented by the expansion in a series of Legendre polynomials

$$\psi_c = \frac{1}{kr} \sum_{l=0}^{\infty} (2l+1)\, i^l\, e^{i\sigma_l}\, F_l(\gamma; kr)\, P_l(\cos\theta). \tag{XI.46}$$

The expansion is analogous to the expansion of a plane wave

$$e^{ikz} = \sum_{l=0}^{\infty} (2l+1)\, i^l\, j_l(kr)\, P_l(\cos\theta) \tag{XI.47}$$

to which it reduces in the limit where $\gamma \to 0$.

To prove the relation (XI.46), we use the integral representation (B.6) of the hypergeometric function occurring in the definition of ψ_c, which yields

$$\begin{aligned}\psi_c &= \frac{e^{-\frac{1}{2}\pi\gamma}}{(1-e^{-2\pi\gamma})\,\Gamma(-i\gamma)}\, e^{ikz} \int_{\Gamma_0} e^{ik(r-z)t}\, t^{-i\gamma-1}\,(1-t)^{i\gamma}\,dt \\ &= \frac{e^{-\frac{1}{2}\pi\gamma}}{(1-e^{-2\pi\gamma})\,\Gamma(-i\gamma)} \int_{\Gamma_0} e^{ikrt}\, e^{ikz(1-t)}\, t^{-i\gamma-1}\,(1-t)^{i\gamma}\,dt.\end{aligned}$$

Expanding the exponential $\exp[ikz(1-t)]$ in the integrand in a series of Legendre polynomials, in accordance with formula (XI.47), and interchanging summation and integration, we obtain

$$\psi_c = \sum_{l=0}^{\infty} (2l+1)\, i^l\, \varphi_l(r)\, P_l(\cos\theta) \tag{XI.48}$$

with

$$\varphi_l(r) = \frac{e^{-\frac{1}{2}\pi\gamma}}{(1-e^{-2\pi\gamma})\,\Gamma(-i\gamma)} \int_{\Gamma_0} e^{ikrt}\, j_l[kr(1-t)]\, t^{-i\gamma-1}\,(1-t)^{i\gamma}\,dt. \tag{XI.49}$$

On the other hand, one knows that

$$\xi j_l(\xi) = F_l(0;\xi) = \frac{2^l\, l!}{(2l+1)!}\, \xi^{l+1}\, e^{i\xi}\, F(l+1\,|\,2l+2\,|\,-2i\xi),$$

or

$$j_l(\xi) = 2^l\, \xi^l\, e^{i\xi} \sum_{p=0}^{\infty} \frac{(l+p)!}{(2l+1+p)!}\, \frac{(-2i\xi)^p}{p!}.$$

Substituting this expression of j_l in the integral of the right-hand side of (XI.49) and once again interchanging summation and integration, one obtains

$$\varphi_l(r) = \frac{2^l\, e^{-\frac{1}{2}\pi\gamma}}{(1-e^{-2\pi\gamma})\,\Gamma(-i\gamma)}\, (kr)^l\, e^{ikr} \sum_{p=0}^{\infty} \frac{(l+p)!}{(2l+1+p)!} \left[\int_{\Gamma_0} t^{-i\gamma-1}\,(1-t)^{l+p+i\gamma}\,dt\right] \frac{(-2ikr)^p}{p!};$$

and since, according to relation (B.5)

$$\int_{\Gamma_s} t^{-i\gamma-1}(1-t)^{l+p+i\gamma}\,dt = (1-e^{-2\pi\gamma})\,\frac{\Gamma(-i\gamma)\,\Gamma(l+p+1+i\gamma)}{(l+p)!},$$

$$\varphi_l(r) = 2^l e^{-\frac{1}{2}\pi\gamma}(kr)^l e^{ikr}\sum_{p=0}^{\infty}\frac{\Gamma(l+p+1+i\gamma)}{(2l+1+p)!}\,\frac{(-2ikr)^p}{p!}$$

$$= 2^l e^{-\frac{1}{2}\pi\gamma}\frac{\Gamma(l+1+i\gamma)}{(2l+1)!}(kr)^l e^{ikr}F(l+1+i\gamma\,|\,2l+2\,|-2ikr).$$

Taking into account definitions (XI.40), (XI.42) and (XI.43), this yields

$$kr\varphi_l(r) = e^{i\sigma_l}F_l(\gamma;\,kr).$$

One obtains the desired expansion by substituting this expression into equation (XI.48).

Actually, one might have expected a result of this form. Indeed, since ψ_c is a regular solution of the Schrödinger equation (XI.20), $r\varphi_l(r)$ is necessarily a regular solution of the radial equation (XI.37) and therefore proportional to $F_l(\gamma;\,kr)$. The main purpose of the above calculation was to find that constant of proportionality.

It is instructive to express the expansion (XI.46) by means of outgoing and incoming Coulomb waves. Substituting for F_l its expression as function of $u_l^{(+)}$ and $u_l^{(-)}$ [eq. (XI.44)], one finds

$$\psi_c = \frac{1}{2ikr}\sum_{l=0}^{\infty}(2l+1)\,i^{l+1}\,[u_l^{(-)} - e^{2i\sigma_l}u_l^{(+)}]\,P_l(\cos\theta). \qquad \text{(XI.50)}$$

11. Modifications of the Coulomb Potential by a Short-Range Interaction

When a short-range interaction $V'(r)$ is added to the Coulomb field $V_c(r)$, the stationary state of collision is no longer represented by the pure Coulomb wave, but by a wave ψ whose expansion in a series of Legendre polynomials is of the form

$$\psi = \frac{1}{kr}\sum_{l=0}(2l+1)\,i^l\,\chi_l(r)\,P_l(\cos\theta). \qquad \text{(XI.51)}$$

The phase-shift method which allows us to treat the scattering of a particle by the potential $V'(r)$ may be extended almost step by step to the scattering by the potential $V_c(r)+V'(r)$; it suffices in practice

to replace the free waves by the corresponding Coulomb waves at all stages.

The function $\chi_l(r)$ is a solution of the radial equation

$$\left[\frac{d^2}{dr^2} + k^2 - \frac{2\gamma k}{r} - \frac{2m}{\hbar^2} V' - \frac{l(l+1)}{r^2}\right] \chi_l(r) = 0. \qquad \text{(XI.52)}$$

One can show (Problem XI.2) that if $V'(r)$ tends to zero asymptotically at least as fast as $1/r^2$, the solutions of this radial equation asymptotically approach linear combinations of the exponentials

$$e^{\pm i(kr - \gamma \ln 2kr)}$$

or, what is equivalent, linear combinations of outgoing and incoming Coulomb functions $u_l^{(+)}$ and $u_l^{(-)}$. In particular, the regular solution of this equation tends asymptotically toward a certain linear combination of these two functions. Let

$$A_l[u_l^{(-)} - e^{2i\delta_l} e^{2i\sigma_l} u_l^{(+)}]$$

be this linear combination; A_l is an arbitrary constant. The phase shift δ_l is characteristic of the potential $V'(r)$ we have added to the Coulomb potential. It is zero if $V'(r) = 0$ and hereafter plays a role analogous to that of the phase shifts in the theory of scattering by short-range potentials.

The regular solution χ_l of the radial equation must be chosen in such a way that the wave ψ represents a stationary state of collision. Therefore, $\psi - \psi_c$ must behave asymptotically like a purely outgoing wave of the type $\exp[i(kr - \gamma \ln 2kr)]/r$. This condition is fulfilled if $A = \frac{1}{2}i$ for all values of l, as one can readily see by comparing (XI.50) and (XI.51). In the asymptotic region — that is, for values of r sufficiently large so that $V'(r)$ can be neglected — ψ can thus be expanded as follows:

$$\begin{aligned} \psi &\underset{r\to\infty}{\sim} \frac{1}{2kr} \sum_l (2l+1) i^{l+1} [u_l^{(-)} - e^{2i\delta_l} e^{2i\sigma_l} u_l^{(+)}] P_l(\cos\theta) \\ &\underset{r\to\infty}{\sim} \psi_c - \frac{1}{2kr} \sum_l (2l+1) i^{l+1} e^{2i\sigma_l} (e^{2i\delta_l} - 1) u_l^{(+)} P_l(\cos\theta). \end{aligned} \qquad \text{(XI.53)}$$

As in § 7, one can put ψ in the form of a sum:

$$\psi = \psi_i + \psi_s \qquad \text{(XI.54)}$$

where ψ_l is the function defined by eq. (XI.26) and represents the incident wave. ψ_s on the other hand differs from the function defined by eq. (XI.27); its asymptotic form, after some calculation, is written

$$\psi_s \underset{r\to\infty}{\sim} \frac{1}{r}\exp\left[\mathrm{i}(kr-\gamma \ln 2kr)\right]\cdot f(\theta)$$

with

$$f(\theta) = f_c(\theta)+f'(\theta) \tag{XI.55}$$

$$f_c(\theta) = -\frac{\gamma}{2k\sin^2\frac{1}{2}\theta}\exp\left[-\mathrm{i}\gamma\ln(\sin^2\tfrac{1}{2}\theta)+2\mathrm{i}\sigma_0\right] \tag{XI.55a}$$

$$f'(\theta) = \frac{1}{2\mathrm{i}k}\sum(2l+1)\,\mathrm{e}^{2\mathrm{i}\sigma_l}(\mathrm{e}^{2\mathrm{i}\delta_l}-1)\,P_l(\cos\theta). \tag{XI.55b}$$

It is easily shown, by an argument analogous to the one given in § 8, that the scattering cross section is

$$\sigma(\Omega) = |f(\theta)|^2. \tag{XI.56}$$

One can write it in the form of a sum of three terms if one replaces $f(\theta)$ by its expression (XI.55), namely:

$$\sigma(\Omega) = \sigma_c(\Omega)+2\mathrm{Re}\, f_c^* f'+|f'(\theta)|^2.$$

Numerous properties of ordinary phase shifts apply without much change to the phase shifts introduced here. In particular, the series (XI.55*b*) converges more rapidly the shorter the range of the additional potential $V'(r)$. Formulae (X.39) through (X.44) of Ch. X, § 10 remain rigorously valid, it being understood that the functions $u_l^{(\pm)}$ denote Coulomb waves, and not free waves (Problem XI.3). However, the numerical values of the quantities τ_l, v_l, $q_l^{(+)}$ may be very different from the same quantities for free waves, so that the discussions of the behavior at low energies and of the convergence of the series must be rather thoroughly revised. In particular, if the Coulomb potential is repulsive, the penetration factor is smaller the smaller the incident energy, and one has $v_l \ll 1$ for any l as soon as $E \lesssim Z_1Z_2e^2/r_0$ (energy below the Coulomb barrier at the point $r=r_0$). Apart from this, the entire treatment of scattering resonances can be taken over without change. As long as we make a few changes in the definition of the quantities involved (Problem XI.4), the formulae (X.72) and (X.73), starting point of the Born approximation, and formula (X.77), starting point of the "effective range" approximation, remain valid.

EXERCISES AND PROBLEMS

1. Making use of the radial equation of the hydrogen atom, derive the recurrence relation (*Kramers relation*)

$$\frac{s+1}{n^2}\langle r^s\rangle - (2s+1)a\langle r^{s-1}\rangle + \frac{s}{4}[(2l+1)^2 - s^2]\, a^2\langle r^{s-2}\rangle = 0$$

in which $\langle r^s\rangle$ stands for the mean value of r^s when the atom is in the quantum state (nlm) $(a > -2l-3)$. Derive the expressions for $\langle r^{-1}\rangle$, $\langle r\rangle$, $\langle r^2\rangle$ given in Appendix B, § 3. [This relation does not enable us to determine $\langle r^{-2}\rangle$].

Show that for any stationary state of a hydrogen-like atom, the mean value of the kinetic energy is equal and opposite to the eigenvalue of the energy:

$$E_n = -\langle p^2/2m\rangle_{nlm}.$$

2. Consider the scattering of a particle of mass m by the central potential $V(r) = Ze^2/r + V'(r)$, where $V'(r)$ tends to zero at least as fast as $1/r^2$ when $r \to \infty$. Show that the solutions of the radial equation tend asymptotically toward linear combinations of the exponentials

$$\exp\,[\pm i(kr - \gamma \ln 2kr)]$$

$(k = \sqrt{2mE}/\hbar,\ \gamma = Ze^2/\hbar v$; E and v are the incident energy and velocity, respectively).

3. Show that the formulae (XI.39) through (XI.44) of Chapter X (§ 10) remain valid if one adds a Coulomb-interaction term $Z_1Z_2e^2/r$ to the finite-range potential, provided that one uses the appropriate definitions for the functions $u_l^{(+)}$ and $u_l^{(-)}$.

4. How must the integral representations (X.72) and (X.73) be modified when the scattering potential is the sum of a short-range term and a Coulomb-interaction term? Answer the same question for formula (X.77). Examine the effective-range theory for s-wave scattering when the short-range term has the properties stated in Chapter X, § 20 (Cf. H. A. Bethe, *loc. cit.*, footnote p. 409).

CHAPTER XII

THE HARMONIC OSCILLATOR

1. Introduction

In Classical Mechanics, a harmonic oscillator is a particle constrained to move along an axis and subject to a restoring force proportional to a point located on that axis. It is a typical problem whose solution is well known. Let q be the position coordinate of the particle on the axis, taking the center of force as the origin, and let p be its momentum, m its mass, and $-m\omega^2 q$ the restoring force. The equations of motion of the particle are derived from the Hamiltonian $(p^2+m^2\omega^2q^2)/2m$; one easily shows that the motion is an oscillatory, sinusoidal motion of (angular) frequency ω about the origin.

The corresponding quantum-mechanical problem is that of a particle of mass m in one dimension with the Hamiltonian

$$\mathscr{H} = \frac{1}{2m}(p^2+m^2\omega^2q^2), \tag{XII.1}$$

the position variable q and the momentum p being connected by the commutation relation

$$[q, p] = i\hbar. \tag{XII.2}$$

One is dealing with a particularly simple quantum system whose Schrödinger equation one knows how to solve rigorously. Moreover, it possesses a certain number of noteworthy properties.

The study of the harmonic oscillator is of great importance in Quantum Theory since the Hamiltonian of the type (XII.1) enters in all problems involving quantized oscillations: one encounters it in Quantum Electrodynamics, and more generally in Quantum Field Theory; one likewise encounters it in the theory of molecular and crystalline vibrations. On the other hand, problems related to the harmonic oscillator furnish an excellent illustration of the general principles and of the formalism of Quantum Theory. All these reasons justify the detailed study which is made of it in this chapter.

The first two sections are devoted to the one-dimensional oscillator. The general solution of the eigenvalue problem of the Hamiltonian is

given in Section I. Section II is devoted to various applications: determination of the generating function of the stationary states, solution of Heisenberg's equations of motion, comparison of the classical and quantum-mechanical oscillator; a study of the motion of a wave packet – which constitutes a good illustration both of the correspondence principle as well as of the uncertainty relations – and finally, an examination of some properties of ensembles of harmonic oscillators in thermodynamic equilibrium.

Section III deals with the isotropic harmonic oscillator in several dimensions. The essential characteristic of this problem is the existence of degenerate eigenvalues. The consequences of degeneracy are examined in detail in the two particular cases of the isotropic oscillator in two and three dimensions.

I. EIGENSTATES AND EIGENVECTORS OF THE HAMILTONIAN

2. The Eigenvalue Problem

In order to avoid cluttering the calculations by useless constants, we put

$$\mathscr{H} = H\hbar\omega \tag{XII.3}$$

$$q = \left(\frac{\hbar}{m\omega}\right)^{\frac{1}{2}} Q \tag{XII.4}$$

$$p = (m\hbar\omega)^{\frac{1}{2}} P. \tag{XII.5}$$

Our problem is to find the eigenvalues, and to construct the eigenvectors of the operator

$$H = \tfrac{1}{2}(P^2 + Q^2), \tag{XII.6}$$

the Hermitean operators P and Q satisfying the commutation relations

$$[Q, P] = \mathrm{i}. \tag{XII.7}$$

To solve this problem, one can choose a particular representation, the $\{Q\}$ representation for instance, and solve the Schrödinger equation in that representation. Since P is then represented by the differential operator $(-\mathrm{i}\,\mathrm{d}/\mathrm{d}Q)$, we have the one-dimensional Schrödinger equation

$$\frac{1}{2}\left[-\frac{\mathrm{d}^2}{\mathrm{d}Q^2} + Q^2\right] u(Q) = \varepsilon u(Q). \tag{XII.8}$$

The method, due to Dirac, which we follow here, is more direct; it consists in constructing the eigenvectors of H by the application of suitable operators to one of them. One thus manages to solve the eigenvalue problem without referring to a particular representation, basing oneself solely upon the fundamental axioms of Hilbert space and the commutation relation (XII.7). The method described here can in fact be regarded as a method of constructing the vector space $\mathscr{E}$ of the dynamical states of the system, and presents strong analogies with the one described in Chapter VIII, § 6.

3. Introduction of the Operators a, $a^\dagger$ and N

Let us put:

$$a = \tfrac{1}{2}\sqrt{2}(Q+iP) \tag{XII.9a}$$

$$a^\dagger = \tfrac{1}{2}\sqrt{2}(Q-iP); \tag{XII.9b}$$

a and $a^\dagger$ are Hermitean conjugates of each other. The commutation relation (XII.7) is equivalent to

$$[a, a^\dagger] = 1. \tag{XII.10}$$

If one replaces Q and P by their expressions as functions of a and $a^\dagger$ in equation (XII.6), one finds

$$H = \tfrac{1}{2}(a\,a^\dagger + a^\dagger\,a). \tag{XII.11}$$

We put

$$N = a^\dagger\,a. \tag{XII.12}$$

From (XII.10) and (XII.11) one deduces

$$H = N + \tfrac{1}{2}. \tag{XII.13}$$

From (XII.10) and (XII.12) we extract the important relations

$$Na = a(N-1) \tag{XII.14a}$$

$$Na^\dagger = a^\dagger(N+1). \tag{XII.14b}$$

The eigenvalue problem we seek to solve is equivalent to the problem of constructing the eigenvectors of the operator N defined by expression (XII.12), in which the operators a and $a^\dagger$ are two Hermitean conjugate operators satisfying the relation (XII.10).

For this purpose, we shall prove an important theorem.

THEOREM: *If $|v\rangle$ is an eigenvector of N, and v the corresponding eigenvalue, then*:

(*i*) *necessarily* $\nu \geqslant 0$;

(*ii*) *if* $\nu = 0$, $a|\nu\rangle = 0$; *if not,* $a|\nu\rangle$ *is a non-zero vector of norm*

$$\nu\langle\nu|\nu\rangle,$$

and it is an eigenvector of N *belonging to the eigenvalue* $\nu - 1$;

(*iii*) $a^\dagger|\nu\rangle$ *is certainly not zero; its norm is*

$$(\nu+1)\langle\nu|\nu\rangle,$$

and it is an eigenvector of N *corresponding to the eigenvalue* $\nu+1$.

By hypothesis,

$$N|\nu\rangle = \nu|\nu\rangle, \quad \langle\nu|\nu\rangle > 0.$$

Making use of the definition (XII.12) and the relation (XII.10) we deduce the respective norms of $a|\nu\rangle$ and of $a^\dagger|\nu\rangle$

$$\langle\nu|a^\dagger a|\nu\rangle = \langle\nu|N|\nu\rangle = \nu\langle\nu|\nu\rangle \qquad \text{(XII.15}a\text{)}$$

$$\langle\nu|a\, a^\dagger|\nu\rangle = \langle\nu|(N+1)|\nu\rangle = (\nu+1)\langle\nu|\nu\rangle. \qquad \text{(XII.15}b\text{)}$$

Now, the norm of a vector of Hilbert space is non-negative, and the vanishing of the norm is a necessary and sufficient condition for the vanishing of the vector. In order that this fundamental axiom be fulfilled here, it is necessary and sufficient that $\nu \geqslant 0$ [property (*i*)] [1]. The condition for the vanishing of $a|\nu\rangle$ is a special case of eq. (XII.15*a*). On the other hand, $a|\nu\rangle$ and $a^\dagger|\nu\rangle$ actually satisfy the stated eigenvalue equations since, according to (XII.14*a*) and (XII.14*b*),

$$N\, a|\nu\rangle = a(N-1)|\nu\rangle = (\nu-1)\, a|\nu\rangle$$

$$N\, a^\dagger|\nu\rangle = a^\dagger(N+1)|\nu\rangle = (\nu+1)\, a^\dagger|\nu\rangle.$$

Q.E.D.

4. Spectrum and Basis of N

If $\nu > 0$, the foregoing theorem applies equally well to the vector $a|\nu\rangle$ whose eigenvalue is $\nu - 1$; consequently $\nu \geqslant 1$. If $\nu > 1$, the theorem also applies to the vector $a^2|\nu\rangle$. One thus successively forms the set of eigenvectors

$$a|\nu\rangle, a^2|\nu\rangle, \ldots, a^p|\nu\rangle, \ldots$$

belonging respectively to the eigenvalues

$$\nu-1, \nu-2, \ldots, \nu-p, \ldots .$$

---

[1] Cf. Problem VII.9.

This set is certainly limited since the eigenvalues of N have a lower limit of zero. In other words, the vectors of this set all vanish from a certain rank $n+1$ on: the action of a on the non-zero eigenvector $a^n|\nu\rangle$ belonging to the eigenvalue $\nu-n$ yields zero; according to (*ii*), this requires that $\nu=n$.

In the same way we can apply the theorem to the vector $a^\dagger|\nu\rangle$ which is certainly not zero and belongs to the eigenvalue $\nu+1$, then to the vector $a^{\dagger 2}|\nu\rangle$, and so forth. One thus successively forms an unlimited set of non-zero vectors

$$a^\dagger|\nu\rangle,\ a^{\dagger 2}|\nu\rangle,\ \ldots,\ a^{\dagger p}|\nu\rangle,\ \ldots\ ,$$

eigenvectors of N belonging respectively to the eigenvalues

$$\nu+1,\ \nu+2,\ \ldots,\ \nu+p,\ \ldots\ .$$

In conclusion, *the spectrum of eigenvalues of N is formed by the set of non-negative integers*. Moreover, by repeated action of a or $a^\dagger$ on one of them, one obtains a set of eigenvectors, each belonging to one of the eigenvalues of the spectrum. The ratio of the norms of each vector to that of the succeeding one is given by one or the other of the relations (XII.15*a*) or (XII.15*b*). This ensemble of vectors forms a complete set. Indeed, one can show that any function of a and of $a^\dagger$ which commutes with N is a function of N (Problem XII.1). Consequently N by itself forms a complete set of commuting observables, and none of its eigenvalues is degenerate.

The vectors thus constructed are not normalized to unity. To form an orthonormal basis of the observable N, it suffices to multiply each of them by a suitable constant, which is easily deduced from relations (XII.15*a*) and (XII.15*b*). This constant is defined to within an arbitrary phase, which we shall fix in such a way as to obtain formulae that are as simple as possible. We thus form the set of orthonormal vectors

$$|0\rangle,\ |1\rangle,\ \ldots,\ |n\rangle,\ \ldots \qquad \text{(XII.16)}$$

corresponding, respectively, to the following eigenvalues of N:

$$0,\ 1,\ \ldots,\ n,\ \ldots\ .$$

They can be deduced from each other by the recursion relations

$$a^\dagger|n\rangle = (n+1)^{\frac{1}{2}}|n+1\rangle \qquad \text{(XII.17)}$$

$$a|n\rangle = n^{\frac{1}{2}}|n-1\rangle \qquad (n\neq 0) \qquad \text{(XII.18)}$$

$$a|0\rangle = 0. \qquad \text{(XII.19)}$$

One verifies easily that they can all be deduced from the vector $|0\rangle$ by the relation

$$|n\rangle = (n!)^{-\frac{1}{2}} a^{\dagger n}|0\rangle, \tag{XII.20}$$

that they actually obey the eigenvalue equation

$$N|n\rangle = n|n\rangle, \tag{XII.21}$$

and that their norm is actually equal to 1, hence that they satisfy the relations

$$\langle n|n'\rangle = \delta_{nn'}. \tag{XII.22}$$

Since N forms a complete set by itself, the sequence of vectors (XII.16) forms a complete set of vectors orthonormal in the space $\mathscr{E}$, namely the space of the dynamical states of the quantum system under study. The internal consistency of this construction of $\mathscr{E}$ remains to be verified, that is to say one has to make sure that the vectors of $\mathscr{E}$ all satisfy the characteristic axioms of vectors of Hilbert space, and that the physical quantities associated with the system are observables satisfying suitable algebraic relations. We shall not insist here upon these points of mathematical rigor (Problem XII.3).

5. The $\{N\}$-Representation

The vectors of the sequence (XII.16) form the basis of a certain representation which we shall call the $\{N\}$ representation. From equations (XII.17), (XII.18), (XII.19) and (XII.20) one easily derives the representative matrices of the operators N, a and $a^\dagger$ in this representation. If one adopts the convention of arranging the rows and columns of these matrices in the order of increasing quantum numbers n (the uppermost row corresponds to $n=0$, the following row to $n=1$, etc.; the extreme left column corresponds to $n=0$, the following column to $n=1$, etc.), one finds for N the diagonal matrix

$$N = \begin{pmatrix} 0 & 0 & \cdots & & \\ 0 & 1 & 0 & & \\ \vdots & 0 & 2 & 0 & \\ & & 0 & 3 & 0 \\ & & & 0 & \ddots \end{pmatrix},$$

for a the real matrix

$$a = \begin{pmatrix} 0 & \sqrt{1} & 0 & 0 & \cdots \\ 0 & 0 & \sqrt{2} & & \\ \vdots & 0 & 0 & \sqrt{3} & \\ & & 0 & 0 & \sqrt{4} \\ & & & 0 & 0 \end{pmatrix}$$

whose only non-zero elements are those of the diagonal located immediately above the main diagonal, and for $a^\dagger$ the Hermitean conjugate matrix

$$a^\dagger = \begin{pmatrix} 0 & 0 & 0 & \cdots \\ \sqrt{1} & 0 & 0 & \cdots \\ 0 & \sqrt{2} & 0 & \\ \vdots & & \sqrt{3} & 0 \\ & & & \sqrt{4} \quad 0 \\ & & & \ddots \end{pmatrix}$$

whose only non-zero elements are those of the diagonal located immediately below the main diagonal. Since the observables of the quantum system are all functions of a and of $a^\dagger$, it is easy to form the matrices representing them in the $\{N\}$ representation. In particular, one has

$$\mathscr{H} = (N+\tfrac{1}{2})\hbar\omega \tag{XII.23}$$

$$q = \left(\frac{\hbar}{2m\omega}\right)^{\frac{1}{2}}(a^\dagger + a) \tag{XII.24}$$

$$p = \mathrm{i}\left(\frac{m\hbar\omega}{2}\right)^{\frac{1}{2}}(a^\dagger - a). \tag{XII.25}$$

$\mathscr{H}$ is diagonal in this representation and its eigenvalues are

$$(n+\tfrac{1}{2})\hbar\omega \qquad (n = 0, 1, 2, \ldots, \infty).$$

As q and p are linear functions of a and $a^\dagger$, their only non-zero matrix elements are located on the two diagonals adjacent to the main diagonal. We leave it to the reader to set up these matrices.

6. Creation and Destruction Operators

The operators N, $a^\dagger$ and a were introduced to facilitate the solution of the eigenvalue problem. If $\mathscr{H}$ is the Hamiltonian of a one-dimensional quantized particle, these operators have no immediate physical significance.

But the eigenvalue problem of $\mathscr{H}$ is susceptible of another interpretation. Indeed, since the energy levels are equidistant by $\hbar\omega$, one may consider $\mathscr{H}$ as the Hamiltonian of a system of indistinguishable corpuscles, all being in one and the same dynamical state, whose energy is $\hbar\omega$. Their number N can vary, each eigenstate of $\mathscr{H}$ corresponding to a well-defined value of N and therefore to a well-defined

value of the energy of the total system. Thus the vector $|n\rangle$ represents a state constituted of n corpuscles; the vector $|0\rangle$ is the vacuum state, for which the number of corpuscles is zero. When one goes over from the state $|n\rangle$ to the state $|n+1\rangle$, the number of corpuscles increases by one unit and the total energy of the system increases by the amount $\hbar\omega$. One notes that the energy of the vacuum is not zero but equal to $\frac{1}{2}\hbar\omega$; this anomaly may be avoided if one takes $\mathscr{H} - \frac{1}{2}\hbar\omega$, and not $\mathscr{H}$ as the operator defining the energy of the system.

According to this interpretation, the operator N represents the number of corpuscles and can take on all integral values between 0 and $+\infty$. The operator $a^\dagger$ transforms a state constituted of n corpuscles into a state of $(n+1)$ corpuscles: $a^\dagger$ is a *creation operator*. The operator a, on the contrary, diminishes the number of corpuscles present by one unit: a is a *destruction operator*.

This interpretation of the harmonic oscillator is widely used in Quantum Field Theory and in the theory of crystalline and molecular vibrations. The electromagnetic field, for instance, can be put in the form of a superposition of plane waves characterized by their polarization $\boldsymbol{\epsilon}$ and their wave vector $\boldsymbol{k}$; their frequency is $\omega = k/c$. Classically, the intensity of each component can vary in continuous fashion; actually, it varies by light quanta or photons of energy $\hbar\omega$. The Hamiltonian of the quantized electromagnetic field is a superposition of terms, each referring to a particular type of photon characterized by $\boldsymbol{\epsilon}$ and $\boldsymbol{k}$ [we use the subscript s to denote the ensemble $(\boldsymbol{\epsilon}, \boldsymbol{k})$]:

$$\boldsymbol{H} = \sum_s \mathscr{H}_s.$$

Each partial Hamiltonian can be written in the form

$$\mathscr{H}_s = \hbar\omega_s a_s^\dagger a_s.$$

The operators a_s and $a_s^\dagger$ are Hermitean conjugates of each other and satisfy the commutation relations

$$[a_s, a_{s'}^\dagger] = \delta_{ss'},$$

a simple generalization of relation (XII.10). The operators $a_s^\dagger$ and a_s are respectively interpreted as creation and destruction operators of photons of the type s (cf. Ch. XXI).

7. $\{Q\}$ Representation. Hermite Polynomials

In the language of Wave Mechanics the eigenvalue problem of $\mathscr{H}$ consists in determining the values E for which the equation

$$\mathscr{H}\psi(q) \equiv \left(-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}q^2} + \tfrac{1}{2}m\omega^2 q^2 \right) \psi(q) = E\psi(q)$$

possesses a regular solution at the two limits of the interval $(-\infty, +\infty)$. If one applies the discussion of Ch. III (§ 10) to this problem, one finds that the values of E which fulfill this condition form a discrete spectrum, and that to each of them there corresponds one and only one solution (defined to within a constant); moreover, this solution has a finite norm. This is quite in keeping with the foregoing study, according to which the spectrum of $\mathscr{H}$ is entirely discrete and non-degenerate. Upon solving the eigenvalue problem thus stated, we would again find the sequence of eigenvalues of $\mathscr{H}$,

$$\tfrac{1}{2}\hbar\omega, \tfrac{3}{2}\hbar\omega, \ldots, (n+\tfrac{1}{2})\hbar\omega, \ldots .$$

The corresponding eigenfunctions $\psi_n(q) \equiv \langle q|n\rangle$ are the functions representing the eigenstates $|n\rangle$ in the $\{q\}$ representation.

In what follows, we adopt the $\{Q\}$ representation deduced from the $\{q\}$ representation by the change of variable (XII.4). The eigenfunctions $u_n(Q)$ and $\psi_n(q)$ which represent the same eigenstate $|n\rangle$ in the $\{Q\}$ and $\{q\}$ representations, respectively, are evidently connected by the relation

$$\langle Q|n\rangle \equiv u_n(Q) = \left(\frac{\hbar}{m\omega}\right)^{\frac{1}{4}} \psi_n(q).$$

Equation (XII.8) is (to within the constant $\hbar\omega$) the Schrödinger equation in the $\{Q\}$ representation.

The eigenfunctions $u_0(Q), u_1(Q), \ldots, u_n(Q), \ldots$ are easily obtained by means of the relations (XII.17) to (XII.19). The eigenfunction of the ground state satisfies eq. (XII.19)

$$\left[\frac{\mathrm{d}}{\mathrm{d}Q} + Q\right] u_0(Q) = 0,$$

whose solution, normalized to unity, is

$$u_0(Q) = \pi^{-\frac{1}{4}}\, \mathrm{e}^{-\frac{1}{2}Q^2}. \qquad \text{(XII.26)}$$

From (XII.17) and (XII.18) we deduce relations between normalized eigenfunctions belonging to neighboring eigenvalues (cf. Appendix B, Sec. III); in particular, the repeated application of (XII.17) allows one to build up all eigenfunctions starting from the function u_0. Rather than use (XII.17), we can equally well make use of relation (XII.20) which is its equivalent; this yields

$$u_n(Q) = [\pi^{\frac{1}{2}} 2^n (n!)]^{-\frac{1}{2}} \left(Q - \frac{\mathrm{d}}{\mathrm{d}Q}\right)^n \mathrm{e}^{-\frac{1}{2}Q^2}. \qquad \text{(XII.27)}$$

Using the operator identity

$$\left(Q - \frac{\mathrm{d}}{\mathrm{d}Q}\right) \equiv \left(-\mathrm{e}^{\frac{1}{2}Q^2} \frac{\mathrm{d}}{\mathrm{d}Q} \mathrm{e}^{-\frac{1}{2}Q^2}\right)$$

one can write equation (XII.27) in the form (B.70) in which $H_n(Q)$ is the Hermite polynomial of order n in accordance with the definition (B.59). Thus $u_n(Q)$ is the product of $\exp(-\frac{1}{2}Q^2)$ and an nth order polynomial of parity $(-)^n$. The main properties of these polynomials are listed in Appendix B (§ 7).

II. APPLICATIONS AND VARIOUS PROPERTIES

8. Generating Function for the Eigenfunctions $u_n(Q)$.

As an application, we shall determine a *generating function* of the functions $u_n(Q)$, that is a function $F(t, Q)$ such that

$$F(t, Q) = \sum_{n=0}^{\infty} c_n u_n(Q) t^n,$$

where the c_n are suitable normalization constants. Since $u_n(Q)$ represents the vector $(n!)^{-\frac{1}{2}} a^{\dagger n}|0\rangle$ [eq. (XII.20)], the function $F(t, Q)$ considered as a function of Q represents the vector

$$\sum_n \frac{c_n}{(n!)^{\frac{1}{2}}} (a^\dagger t)^n |0\rangle.$$

With the choice

$$c_n = \frac{1}{(n!)^{\frac{1}{2}}},$$

$F(t, Q)$ represents the vector $\exp(a^\dagger t)|0\rangle$:

$$F(t, Q) = \langle Q| \exp(a^\dagger t)|0\rangle. \qquad \text{(XII.28)}$$

To calculate this last expression, we shall make use of the following lemma:

LEMMA. – *If the commutator of two operators A, B commutes with each of them:*

$$[A, [A, B]] = [B, [A, B]] = 0,$$

one has the identity

$$e^{A+B} = e^{A}\, e^{B}\, e^{-\frac{1}{2}[A,B]}. \tag{XII.29}$$

The following proof is due to Glauber.

Let us consider the operator depending on the parameter x:

$$f(x) = e^{Ax}\, e^{Bx}.$$

One has

$$\frac{df}{dx} = A\, e^{Ax}\, e^{Bx} + e^{Ax}\, B\, e^{Bx}$$

$$= (A + e^{Ax}\, B\, e^{-Ax}) f(x).$$

But since $[B, A]$ commutes with A,

$$[B, A^n] = n\, A^{n-1}[B, A]$$

$$[B, e^{-Ax}] = \sum_n (-)^n \frac{x^n}{n!} [B, A^n]$$

$$= \sum_n (-)^n \frac{x^n}{(n-1)!} A^{n-1}[B, A]$$

$$= -e^{-Ax}[B, A]x,$$

hence (cf. Problem VIII.4)

$$e^{Ax}\, B\, e^{-Ax} = B - [B, A]x$$

and therefore

$$\frac{df}{dx} = (A + B + [A, B]x)\, f(x).$$

$f(x)$ is the solution of this differential equation for which $f(0) = 1$. Since the operators $(A+B)$ and $[B, A]$ commute, they can be considered here as quantities of ordinary algebra. The differential equation can be easily integrated and gives

$$f(x) = \exp\,[(A+B)x]\, \exp\,(\tfrac{1}{2}[A, B]x^2).$$

The identity (XII.29) results if one sets $x = 1$.

Q.E.D.

Taking $A = Qt/\sqrt{2}$, $B = -\mathrm{i}Pt/\sqrt{2}$, $[A, B] = \frac{1}{2}t^2$, we apply the identity XII.29) to the operator $\exp(a^\dagger t)$, namely

$$\exp(a^\dagger t) = \exp[Qt/\sqrt{2}] \exp[-\mathrm{i}Pt/\sqrt{2}] \exp(-\tfrac{1}{4}t^2).$$

Inserting this expression in eq. (XII.28),

$$F(t, Q) = \exp(-\tfrac{1}{4}t^2) \exp[Qt/\sqrt{2}] \langle Q| \exp[-\mathrm{i}Pt/\sqrt{2}]|0\rangle.$$

But

$$\langle Q| \exp[-\mathrm{i}Pt/\sqrt{2}]|0\rangle = \exp\left(-\frac{t}{\sqrt{2}}\frac{\mathrm{d}}{\mathrm{d}Q}\right) u_0(Q)$$

$$= u_0\left(Q - \frac{t}{\sqrt{2}}\right).$$

With expression (XII.26) for u_0 we obtain after some calculation

$$F(t, Q) \equiv \sum_{n=0}^{\infty} \frac{u_n(Q)}{(n!)^{\frac{1}{2}}} t^n = \pi^{-\frac{1}{4}} \exp(-\tfrac{1}{2}Q^2 + tQ\sqrt{2} - \tfrac{1}{2}t^2). \tag{XII.30}$$

9. Integration of the Heisenberg Equations

Let us consider the harmonic oscillator *in the Heisenberg "representation"*. Since all operators occurring in this paragraph are operators in the Heisenberg representation, we shall omit the subscript H which permitted to distinguish them from corresponding operators in the Schrödinger representation in the discussions of Ch. VIII. These operators evolve in time. We attach the subscript 0 to the values they take at the initial instant $t = 0$.

Taking into account eq. (XII.23) and relations (XII.14), the Heisenberg equations of the operators a and $a^\dagger$ are respectively written

$$\mathrm{i}\hbar \frac{\mathrm{d}a}{\mathrm{d}t} = [a, \mathscr{H}] = \hbar\omega a,$$

$$\mathrm{i}\hbar \frac{\mathrm{d}a^\dagger}{\mathrm{d}t} = [a^\dagger, \mathscr{H}] = -\hbar\omega a^\dagger.$$

These equations may be easily integrated and yield

$$a(t) = a_0 \, \mathrm{e}^{-\mathrm{i}\omega t} \tag{XII.31a}$$

$$a^\dagger(t) = a_0^\dagger \, \mathrm{e}^{+\mathrm{i}\omega t}. \tag{XII.31b}$$

Making use of relations (XII.24) and (XII.25) which give q and p as functions of a and $a^\dagger$, one has

$$q(t) = \left(\frac{\hbar}{2m\omega}\right)^{\frac{1}{2}}(a_0^\dagger\, e^{i\omega t} + a_0\, e^{-i\omega t}) \tag{XII.32}$$

$$p(t) = i\left(\frac{m\hbar\omega}{2}\right)^{\frac{1}{2}}(a_0^\dagger\, \epsilon^{i\omega t} - a_0\, e^{-i\omega t}). \tag{XII.33}$$

If one replaces in these equations a_0 and $a_0^\dagger$ by their expressions as functions of the initial position q_0 and of the initial momentum p_0, one has

$$q(t) = q_0 \cos \omega t + \frac{1}{m\omega} p_0 \sin \omega t \tag{XII.34}$$

$$p(t) = p_0 \cos \omega t - m\omega q_0 \sin \omega t. \tag{XII.35}$$

One finds the same sinusoidal functions as for the classical harmonic oscillator. In particular, the mean values $\langle q\rangle_t$, $\langle p\rangle_t$ follow the classical laws of motion

$$\langle q\rangle_t = \langle q\rangle_0 \cos \omega t + \frac{1}{m\omega}\langle p\rangle_0 \sin \omega t \tag{XII.36}$$

$$\langle p\rangle_t = \langle p\rangle_0 \cos \omega t - m\omega\langle q\rangle_0 \sin \omega t. \tag{XII.37}$$

This property of the harmonic oscillator was already pointed out in Chapter VI.

10. Classical and Quantized Oscillator

In order to illustrate the correspondence between Classical Mechanics and Quantum Mechanics, we compare in this and the following section the motion of the classical oscillator with that of the corresponding quantum-mechanical oscillator.

The general solution of the equations of motion of the classical harmonic oscillator can be written

$$q_{\text{cl.}} = A \sin(\omega t + \varphi),$$
$$p_{\text{cl.}} = m\omega A \cos(\omega t + \varphi).$$

It is a sinusoidal, oscillatory motion of (angular) frequency ω. It depends upon two parameters A and φ. The energy of the oscillator is connected with the amplitude of oscillation A by the relation

$$E_{\text{cl.}} = \frac{m\omega^2 A^2}{2}. \tag{XII.38}$$

If one fixes the energy $E_{\text{cl.}}$, the various possible motions differ from each other in the phase constant φ.

Let $F_{\text{cl.}}$ be a dynamical variable of the system. Since it is a function of $q_{\text{cl.}}$ and of $p_{\text{cl.}}$, $F[q_{\text{cl.}}(t), p_{\text{cl.}}(t)]$, it varies periodically (but not necessarily sinusoidally) with time with frequency ω. The time dependence of $F_{\text{cl.}}$ for two motions of equal energy is the same except for the phase. The average $\bar{F}_{\text{cl.}}$ of $F_{\text{cl.}}$ taken over all motions of the same energy (microcanonical ensemble) is obtained by performing the average over the phase constants; $\bar{F}_{\text{cl.}}$ is independent of time and equal to the average over a period $2\pi/\omega$ of the values assumed by $F_{\text{cl.}}$ during any one of these motions. One finds in particular:

$$\bar{q}_{\text{cl.}} = \bar{p}_{\text{cl.}} = 0 \tag{XII.39}$$

$$\overline{q_{\text{cl.}}^2} = \frac{A^2}{2} = \frac{E_{\text{cl.}}}{m\omega^2} \tag{XII.40}$$

$$\overline{p_{\text{cl.}}^2} = m^2\omega^2\,\overline{q_{\text{cl.}}^2} = mE_{\text{cl.}} \tag{XII.41}$$

(the mean kinetic energy and the mean potential energy of an oscillator are equal).

Let us see how this compares with the behavior of the quantized oscillator in a stationary state. In the state $|n\rangle$, the quantized oscillator has a well-defined energy which is constant in time: $(n+\frac{1}{2})\hbar\omega = E_n$. On the other hand, the observables of position q and of momentum p do not have precise values. One may only define the statistical distribution of the results of measurement of one or the other of these quantities in the eventuality that such a measurement is performed. Since the state is stationary, these statistical distributions are constant in time. In particular, the mean values of q and p are respectively equal to the diagonal element of rank n of the observables q and p in the $\{N\}$ representation:

$$\langle n|q|n\rangle = \langle n|p|n\rangle = 0. \tag{XII.42}$$

The mean values of q^2 and p^2 are easily calculated by expressing these operators as functions of the a and $a^\dagger$ [eqs. (XII.24) and (XII.25)] and using relations (XII.17) to (XII.19). We have

$$\langle n|q^2|n\rangle = \frac{1}{2}\frac{\hbar}{m\omega}\langle n|(a^\dagger a + aa^\dagger)|n\rangle = \frac{E_n}{m\omega^2} \tag{XII.43}$$

$$\langle n|p^2|n\rangle = \tfrac{1}{2}m\hbar\omega\langle n|(a^\dagger a + aa^\dagger)|n\rangle = mE_n. \tag{XII.44}$$

The correspondence principle demands (cf. Problem XII.4) that in the limit where $n \to \infty$, the expressions for the average values (XII.42), (XII.43) and (XII.44) become respectively identical to the classical expressions (XII.39), (XII.40) and (XII.41) for the same value of the energy ($E_n = E_{\text{cl.}}$). The fact that this identity is rigorously true for all values of n, even small ones, is a characteristic property of the harmonic oscillator.

We note in passing that in the state $|n\rangle$

$$\Delta p \cdot \Delta q = \frac{E_n}{\omega} = (n+\tfrac{1}{2})\hbar, \tag{XII.45}$$

in accordance with the position-momentum uncertainty relations.

11. Motion of the Minimum Wave Packet and Classical Limit

Consider the one-dimensional wave packet

$$f(q) = \left(\frac{m\omega}{\pi\hbar}\right)^{\frac{1}{4}} \exp\left[\frac{\mathrm{i}}{\hbar}\langle p\rangle q - \frac{m\omega}{2\hbar}(q-\langle q\rangle)^2\right]. \tag{XII.46}$$

This is a minimum wave packet (Problem IV.4): it represents a particle localized in configuration space about its mean position $\langle q\rangle$ with a root-mean-square deviation $\Delta q = (\hbar/2m\omega)^{\frac{1}{2}}$ and localized in momentum space about its mean position $\langle p\rangle$ with a root-mean-square deviation $\Delta p = (\hbar m\omega/2)^{\frac{1}{2}}$.

If this particle is subject to the Hamiltonian $\mathscr{H}$, one can show (Problem XII.6) that such a packet retains minimum size in the course of time and that it oscillates with the frequency ω. In more precise fashion, the statistical distribution of q, $\varrho(q, t)$, varies according to the law

$$\varrho(q, t) \equiv |f(q, t)|^2 = \left(\frac{m\omega}{\pi\hbar}\right)^{\frac{1}{2}} \exp\left[-\frac{m\omega}{\hbar}(q-\langle q\rangle_t)^2\right].$$

It oscillates without distortion, its center $\langle q\rangle_t$ carrying out the sinusoidal motion predicted by classical theory. The statistical distribution of p exhibits analogous behavior.

The statistical distribution of $\mathscr{H}$, on the other hand, is constant in time. The probability of finding the system in the state of energy $(n+\frac{1}{2})\hbar\omega$ is equal at every instant to [Problem (XII.6)]

$$\mathrm{e}^{-E_{\text{cl.}}/\hbar\omega}\frac{(E_{\text{cl.}}/\hbar\omega)^n}{n!}.$$

We make use of the notation

$$E_{\text{cl.}} = \frac{1}{2m}(\langle p\rangle^2 + m^2\omega^2\langle q\rangle^2).$$

From this probability law, one easily deduces the average value of the energy

$$\begin{aligned}\langle E\rangle &= \mathrm{e}^{-E_{\text{cl.}}/\hbar\omega}\sum_{n=0}^{\infty}(n+\tfrac{1}{2})\hbar\omega\frac{(E_{\text{cl.}}/\hbar\omega)^n}{n!} \\ &= E_{\text{cl.}} + \tfrac{1}{2}\hbar\omega\end{aligned} \qquad \text{(XII.47)}$$

and its root-mean-square deviation

$$\Delta E = \sqrt{\langle \mathscr{H}^2\rangle - \langle \mathscr{H}\rangle^2} = \sqrt{\hbar\omega E_{\text{cl.}}}. \qquad \text{(XII.48)}$$

This wave packet provides us with a good illustration of the uncertainty relations.

It was chosen in such a way that the product of the deviations $\Delta p \cdot \Delta q$ remains constant and equal to the minimum value $\frac{1}{2}\hbar$.

As far as the time-energy relation is concerned, we can compare ΔE to the characteristic time τ_q of the rate of evolution of the statistical distribution of q. τ_q is the time required for the center $\langle q\rangle_t$ of the distribution to travel by an amount equal to its extension Δq; since the velocity of the center is $\langle p\rangle_t/m$, one has

$$\tau_q = \frac{m}{\langle p\rangle_t}\Delta q = \frac{1}{\langle p\rangle_t}(\hbar m/2\omega)^{\frac{1}{2}}.$$

τ_q passes periodically through a minimum when $\langle p\rangle_t$ attains its largest value $(2mE_{\text{cl.}})^{\frac{1}{2}}$. One then has

$$\tau_{q\,\text{min.}} = \tfrac{1}{2}\hbar\left(\frac{1}{\hbar\omega E_{\text{cl.}}}\right)^{\frac{1}{2}}.$$

According to (XII.48),

$$\tau_{q\,\text{min.}}\,\Delta E = \tfrac{1}{2}\hbar \qquad \text{(XII.49)}$$

in accordance with the time-energy uncertainty relation (VIII.47).

The amplitude of oscillation A of the center of the wave packet is given by the classical relation (XII.38):

$$A = (2E_{\text{cl.}}/m\omega^2)^{\frac{1}{2}}.$$

In the limit where this amplitude is large compared to the extension $(\hbar/2m\omega)^{\frac{1}{2}}$ of the packet, and to the extent one can consider lengths of the order of $(\hbar/2m\omega)^{\frac{1}{2}}$ negligible, the *classical picture* of a point particle oscillating according to the law $\langle q\rangle_t$ provides a statisfactory description of the phenomenon. This limit is just the one of very large quantum numbers as required by the correspondence principle. Indeed, it is realized when $E_{\text{cl.}} \gg \hbar\omega$; now, the number of states of quantized energy contributing appreciably to the composition of the wave packet is of the order of the ratio of the spread ΔE to the level spacing, namely

$$\frac{\Delta E}{\hbar\omega} = \left(\frac{E_{\text{cl.}}}{\hbar\omega}\right)^{\frac{1}{2}} \gg 1.$$

Of course, this classical picture likewise supposes that one treats the dispersion in momentum, $\Delta p = (\hbar m\omega/2)^{\frac{1}{2}}$, and the dispersion in energy, $\Delta E = (\hbar\omega E_{\text{cl.}})^{\frac{1}{2}}$, as negligible quantities. As far as the energy is concerned, one has at this level of precision $\langle E\rangle \approx E_{\text{cl.}}$. Indeed,

$$\langle E\rangle - E_{\text{cl.}} = \tfrac{1}{2}\hbar\omega \ll \Delta E.$$

One is thus well justified in attributing to the system the energy $E_{\text{cl.}}$ of the corresponding classical particle.

12. Harmonic Oscillators in Thermodynamic Equilibrium

Consider a harmonic oscillator in thermodynamic equilibrium with a heat reservoir at temperature T. Its dynamical state is not a pure state but a statistical mixture represented by the density operator

$$\varrho = \frac{\mathrm{e}^{-\mathscr{H}/kT}}{\operatorname{Tr} \mathrm{e}^{-\mathscr{H}/kT}} \tag{XII.50}$$

in conformity with the Boltzmann law. We shall examine some of the properties of this mixture.

Let us first of all calculate the *partition function*

$$Z(\mu) = \operatorname{Tr} \mathrm{e}^{-\mu\mathscr{H}}.$$

This calculation of the trace is easily performed in the representation where $\mathscr{H}$ is diagonal:

$$\begin{aligned} Z(\mu) &= \sum_{n=0}^{\infty} \exp\left[-\mu(n+\tfrac{1}{2})\hbar\omega\right] \\ &= \mathrm{e}^{-\frac{1}{2}\mu\hbar\omega} \sum_{n=0}^{\infty} (\mathrm{e}^{-\mu\hbar\omega})^n, \end{aligned}$$

from which, after summation of the geometric series of the right-hand member:

$$Z(\mu) = \frac{e^{-\frac{1}{2}\mu\hbar\omega}}{1-e^{-\mu\hbar\omega}}. \tag{XII.51}$$

The mean energy

$$\langle E \rangle \equiv \mathrm{Tr}\, \varrho \mathscr{H}$$

is deduced from the partition function by applying eq. (VIII.84). One has

$$\ln Z = -\tfrac{1}{2}\mu\hbar\omega - \ln(1-e^{-\mu\hbar\omega}),$$

from which

$$\begin{aligned} \langle E \rangle &= -\left.\frac{\partial(\ln Z)}{\partial\mu}\right|_{\mu=1/kT} \\ &= \tfrac{1}{2}\hbar\omega \coth\frac{\hbar\omega}{2kT} \\ &= \tfrac{1}{2}\hbar\omega + \frac{\hbar\omega}{e^{\hbar\omega/kT}-1}. \end{aligned} \tag{XII.52}$$

This is just Planck's formula (to within the constant $\frac{1}{2}\hbar\omega$) for the average energy of a quantized oscillator.

At very low temperatures ($kT \ll \hbar\omega$), the oscillator remains with near certainty in its ground state

$$\langle E \rangle \approx \tfrac{1}{2}\hbar\omega.$$

At very high temperatures ($kT \gg \hbar\omega$), the average energy tends toward the one given by classical Maxwell–Boltzmann statistics:

$$\langle E \rangle \approx kT.$$

As a last property of the quantized oscillator in thermodynamic equilibrium, let us mention the following theorem, due to F. Bloch:

THEOREM. – *The probability law of a given combination $\alpha q + \beta p$ of the momentum and the position is a Gaussian.*

To prove this theorem we calculate the characteristic function $\varphi(\xi)$ of this probability law. $\varphi(\xi)$ is by definition the mean value of $\exp[i\xi(\alpha q+\beta p)]$:

$$\varphi(\xi) = \mathrm{Tr}\, \varrho\, e^{i\xi(\alpha q+\beta p)}. \tag{XII.53}$$

We shall calculate this trace in the $\{N\}$ representation, where ϱ is diagonal.

We first calculate the quantities

$$g_n(\xi) = \langle n|\ \mathrm{e}^{\mathrm{i}\xi(\alpha q+\beta p)}\ |n\rangle. \tag{XII.54}$$

One has [eq. (XII.24) and (XII.25)]

$$\alpha q+\beta p = \gamma a+\gamma^* a^\dagger,$$

with

$$\gamma = \left(\frac{\hbar}{2m\omega}\right)^{\frac{1}{2}}(a-\mathrm{i}m\omega\beta).$$

According to the identity (XII.29),

$$\exp\,[\mathrm{i}\xi(\alpha q+\beta p)] = \exp\,[\mathrm{i}\xi(\gamma a+\gamma^* a^\dagger)] = \exp\,(\tfrac{1}{2}\xi^2\gamma\gamma^*)\exp\,(\mathrm{i}\xi\gamma a)\exp\,(\mathrm{i}\xi\gamma^* a^\dagger),$$

whence

$$g_n(\xi) = \exp\,(\tfrac{1}{2}\xi^2\gamma\gamma^*)\,\langle n|\exp\,(\mathrm{i}\xi\gamma a)\exp\,(\mathrm{i}\xi\gamma^* a^\dagger)\,|n\rangle.$$

Expanding the exponentials and taking into account relation (XII.20), one obtains

$$\mathrm{e}^{\mathrm{i}\xi\gamma^* a^\dagger}|n\rangle = \sum_{t=0}^{\infty}\left(\frac{(n+t)!}{n!}\right)^{\frac{1}{2}}\frac{(\mathrm{i}\xi\gamma^*)^t}{t!}\,|n+t\rangle$$

$$\langle n|\ \mathrm{e}^{\mathrm{i}\xi\gamma a} = \sum_{s=0}^{\infty}\left(\frac{(n+s)!}{n!}\right)^{\frac{1}{2}}\frac{(\mathrm{i}\xi\gamma)^s}{s!}\,\langle n+s|.$$

In the scalar product of these two vectors, the cross terms of the double summation are all zero by virtue of the orthogonality relations; hence finally

$$g_n(\xi) = \mathrm{e}^{\frac{1}{2}\xi^2\gamma\gamma^*}\sum_{s=0}^{\infty}\frac{(-\xi^2\gamma\gamma^*)^s}{(s!)^2}\,\frac{(n+s)!}{n!}. \tag{XII.55}$$

Let us put

$$x = -\xi^2\gamma\gamma^*, \qquad y = \mathrm{e}^{-\hbar\omega/kT}. \tag{XII.56}$$

In the $\{N\}$ representation, ϱ is diagonal and its nth diagonal element is equal to

$$\varrho_n \equiv \langle n|\varrho|n\rangle = \frac{\mathrm{e}^{-(n+\frac{1}{2})\hbar\omega/kT}}{Z(1/kT)} = (1-y)y^n. \tag{XII.57}$$

From equations (XII.53) to (XII.57) we extract

$$\varphi(\xi) = \sum_{n=0}^{\infty} \varrho_n g_n(\xi)$$
$$= (1-y)\, e^{-\frac{1}{2}x} \sum_{n=0}^{\infty} \sum_{s=0}^{\infty} \frac{(n+s)!}{n!(s!)^2}\, x^s y^n.$$

This double series can be summed exactly. The summation over n is carried out first by means of the series expansion

$$\frac{1}{(1-y)^{s+1}} = 1 + (s+1)\, y + \frac{(s+1)(s+2)}{2!}\, y^2 + \ldots$$
$$= \sum_{n=0}^{\infty} \frac{(s+n)!}{s!\, n!}\, y^n.$$

Hence

$$\varphi(\xi) = e^{-\frac{1}{2}x} \sum_s \frac{1}{s!} \left(\frac{x}{1-y}\right)^s$$
$$= \exp\left[x\left(\frac{1}{1-y} - \frac{1}{2}\right)\right].$$

Taking into account the definitions of x and y, this is written

$$\varphi(\xi) = e^{-\frac{1}{2}\sigma\xi^2} \tag{XII.58}$$

with

$$\sigma = \gamma\gamma^* \coth\left(\frac{\hbar\omega}{2kT}\right). \tag{XII.59}$$

Since the characteristic function of the distribution is a Gaussian, the probability law is Gaussian as well: its mean square deviation is σ. Q.E.D.

III. ISOTROPIC HARMONIC OSCILLATORS IN SEVERAL DIMENSIONS

13. General Treatment of the Isotropic Oscillator in p Dimensions

The isotropic harmonic oscillator in p dimensions is the p-dimensional system with Hamiltonian

$$\mathcal{H} = \sum_{i=1}^{p} \mathcal{H}_i \tag{XII.60}$$

$$\mathcal{H}_i = \frac{1}{2m}(p_i^2 + m^2\omega^2 q_i^2). \tag{XII.61}$$

Let $\mathscr{E}_1$ be the space of the dynamical states relating to the pair of variables (p_1, q_1), $\mathscr{E}_2$ the space of the dynamical states relating to the pair of variables (p_2, q_2), etc. The space $\mathscr{E}$ of the dynamical states of the system under consideration is the tensor product of the spaces $\mathscr{E}_1, \mathscr{E}_2, \ldots, \mathscr{E}_p$:

$$\mathscr{E} = \mathscr{E}_1 \otimes \mathscr{E}_2 \otimes \ldots \otimes \mathscr{E}_p. \tag{XII.62}$$

Let us denote by $|n_i\rangle$ (i fixed, $n_i = 0, 1, \ldots, \infty$) the eigenvectors of the Hamiltonian $\mathscr{H}_i$ considered as an operator of the space $\mathscr{E}_i$; they form a complete orthonormal set in $\mathscr{E}_i$. In the following we suppose that their relative phases are chosen in such a way as to satisfy the relations (XII.17) to (XII.20) written in terms of the destruction and creation operators relating to the variables of the type i. The vectors

$$\begin{gathered} |n_1\, n_2 \ldots n_p\rangle \equiv |n_1\rangle\, |n_2\rangle \ldots |n_p\rangle \\ (n_1 = 0, 1, \ldots, \infty; n_2 = 0, 1, \ldots, \infty; \ldots; n_p = 0, 1, \ldots, \infty) \end{gathered} \tag{XII.63}$$

formed by the tensor product of p vectors belonging respectively to the spaces $\mathscr{E}_1, \mathscr{E}_2, \ldots, \mathscr{E}_p$, form a complete orthonormal set in $\mathscr{E}$ [1]). Clearly, these vectors are eigenvectors of $\mathscr{H}$. Moreover, since

$$\begin{gathered} \mathscr{H}_1|n_1\rangle = (n_1 + \tfrac{1}{2})\, \hbar\omega |n_1\rangle \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ \mathscr{H}_p|n_p\rangle \equiv (n_p + \tfrac{1}{2})\, \hbar\omega |n_p\rangle, \end{gathered}$$

one has

$$\begin{aligned} \mathscr{H}|n_1 \ldots n_p\rangle &\equiv (\mathscr{H}_1 + \ldots + \mathscr{H}_p)|n_1 \ldots n_p\rangle \\ &= (n_1 + \ldots + n_p + \tfrac{1}{2}p)\, \hbar\omega |n_1 \ldots n_p\rangle. \end{aligned}$$

The vectors of the basis of $\mathscr{H}$ which we formed are labelled by means of p quantum numbers $n_1, n_2, \ldots, n_p$, which can take on all integral values from 0 to $+\infty$. However, the corresponding eigenvalue of the energy,

$$(n_1 + \ldots + n_p + \tfrac{1}{2}p)\hbar\omega,$$

depends only upon the sum

$$n = n_1 + n_2 + \ldots + n_p$$

---

[1]) In the representation $\{q\} \equiv \{q_1 q_2 \ldots q_p\}$, the vector $|n_1 n_2 \ldots n_p\rangle$ is represented by the product

$$\langle q_1|n_1\rangle\, \langle q_2|n_2\rangle \ldots \langle q_p|n_p\rangle \equiv \psi_{n_1}(q_1)\, \psi_{n_2}(q_2) \ldots \psi_{n_p}(q_p).$$

of these p numbers. For a given integral value of $n(>0)$ there exist

$$C^n_{n+p-1} \equiv \frac{(n+p-1)!}{n!(p-1)!} \tag{XII.64}$$

distinct possible values for the set of numbers $n_1, n_2, \ldots, n_p$. The eigenvalue $(n+\frac{1}{2}p)\hbar\omega$ is thus C^n_{n+p-1}-fold degenerate [1]).

Let us introduce the destruction and creation operators of quanta of the type i:

$$\begin{aligned} a_i &= \left(\frac{m\omega}{2\hbar}\right)^{\frac{1}{2}} q_i + \mathrm{i}(2m\hbar\omega)^{-\frac{1}{2}} p_i \\ a_i^\dagger &= \left(\frac{m\omega}{2\hbar}\right)^{\frac{1}{2}} q_i - \mathrm{i}(2m\hbar\omega)^{-\frac{1}{2}} p_i. \end{aligned} \tag{XII.65}$$

They satisfy the commutation relations [cf. relation (XII.10)]

$$\begin{aligned} &[a_i, a_j] = [a_i^\dagger, a_j^\dagger] = 0 \\ &[a_i, a_j^\dagger] = \delta_{ij} \end{aligned} \qquad (i, j, = 1, \ldots, p). \tag{XII.66}$$

In accordance with the definition of the vectors $|n_i\rangle$ given above, the vectors $|n_1 \ldots n_p\rangle$ satisfy the relations generalizing the relations (XII.17) to (XII.20). In particular, if we designate by $|0\rangle$ the eigenvector of the ground state:

$$|0\rangle \equiv |\overbrace{0 \ldots\ldots\ldots 0}^{p \text{ times}}\rangle$$

we can write

$$a_1|0\rangle = a_2|0\rangle = \ldots = a_p|0\rangle = 0 \tag{XII.67}$$

$$|n_1 \ldots n_p\rangle = (n_1! \ldots n_p!)^{-\frac{1}{2}} a_1^{\dagger n_1} \ldots a_p^{\dagger n_p}|0\rangle. \tag{XII.68}$$

The observables

$$N_i \equiv a_i^\dagger a_i \qquad (i = 1, 2, \ldots, p) \tag{XII.69}$$

each have a spectrum consisting of the sequence of non-negative integers; they are interpreted respectively as the number of quanta of the type 1, 2, ..., p. Their sum

$$N \equiv \sum_{i=1}^{p} N_i,$$

---

[1]) In fact, there are as many distinct sets $(n_1, n_2, \ldots, n_p)$ of p integral numbers $\geqslant 0$ with sum equal to n as there are *distinguishable arrangements of putting n indistinguishable objects into p cells.*

is the total number of quanta. One has

$$\mathscr{H} = (N + \tfrac{1}{2}p)\hbar\omega.$$

It is clear that $N_1, N_2, \ldots, N_p$ form a complete set of commuting observables and that their basis is precisely the basis of $\mathscr{H}$ which we have just formed.

The N_i are evidently not the only constants of the motion forming a complete set. Any operator of the form $a_i a_j^\dagger$ commutes with $\mathscr{H}$; by linear combination of operators of this type, and of their adjoints, one can form p^2 independent Hermitean operators in all. Among the functions of these p^2 constants of the motion there exist several complete sets of commuting observables. We shall illustrate this point in the two special cases $p=2$ and $p=3$.

14. Two-Dimensional Isotropic Oscillator

Here we have the two-dimensional system with Hamiltonian

$$\mathscr{H} = \frac{1}{2m}(p_1^2 + m^2\omega^2 q_1^2) + \frac{1}{2m}(p_2^2 + m^2\omega^2 q_2^2).$$

The study of the preceding section applies to this special case. The following table gives the eigenvalues of $\mathscr{H}$ (first column) and the set of eigenvectors common to N_1 and N_2 which span their respective subspaces (third column):

$$\begin{array}{lll}
\hbar\omega & 1 & |00\rangle \\
2\hbar\omega & 2 & |10\rangle, |01\rangle \\
3\hbar\omega & 3 & |20\rangle, |11\rangle, |02\rangle \\
\cdots & \cdots & \cdots \\
(n+1)\hbar\omega & n+1 & |n\,0\rangle, |n-1\;1\rangle, \ldots, |n-s\;s\rangle, \ldots |0\,n\rangle. \\
\cdots & \cdots & \cdots
\end{array} \tag{XII.70}$$

The angular momentum operator L, defined by

$$L \equiv \frac{1}{\hbar}(q_1 p_2 - q_2 p_1) = \mathrm{i}(a_1 a_2^\dagger - a_1^\dagger a_2), \tag{XII.71}$$

is a constant of the motion. We shall show that N and L form another complete set of commuting observables. To this effect we introduce the operators

$$\begin{aligned} A_\pm &= \tfrac{1}{2}\sqrt{2}\,(a_1 \mp \mathrm{i}a_2) \\ A_\pm^\dagger &= \tfrac{1}{2}\sqrt{2}\,(a_1^\dagger \pm \mathrm{i}a_2^\dagger). \end{aligned} \tag{XII.72}$$

These operators satisfy commutation relations identical to the relations (XII.66) between the a and the $a^\dagger$:

$$\begin{aligned}[A_r, A_s] &= [A_r^\dagger, A_s^\dagger] = 0 \\ [A_r, A_s^\dagger] &= \delta_{rs}\end{aligned} \qquad (r = + \text{ or } -\,;\ s = + \text{ or } -). \qquad \text{(XII.73)}$$

A_+ and $A_+^\dagger$ can thus be interpreted as destruction and creation operators of quanta of type $+$, A_- and $A_-^\dagger$ as destruction and creation, operators of quanta of type $-$; following this interpretation, the operators

$$N_+ \equiv A_+^\dagger A_+ \quad \text{and} \quad N_- \equiv A_-^\dagger A_- \qquad \text{(XII.74)}$$

represent the numbers of "$+$ quanta" and of "$-$ quanta", respectively. Since the commutation relations (XII.73) are identical with relations (XII.66), the problem of forming the eigenvectors common to N_+ and N_- is mathematically identical to the problem of forming the eigenvectors common to N_1 and N_2. Therefore, N_+ and N_- each have as their spectrum the sequence of non-negative integers

$$n_+ = 0, 1, 2, \ldots \qquad n_- = 0, 1, 2, \ldots$$

and these two observables form a complete set of commuting observables: to each pair of quantum numbers (n_+, n_-) corresponds a single common eigenvector (to within a constant). In fact, the relations (XII.67) imply

$$A_+|00\rangle = A_-|00\rangle = 0. \qquad \text{(XII.75)}$$

Thus the vector $|00\rangle$ of Table (XII.70) is an eigenvector of the ground state $(n_+ = n_- = 0)$. The vectors

$$|n_+ n_-\rangle \equiv (n_+! n_-!)^{-\frac{1}{2}} A_+^{\dagger n_+} A_-^{\dagger n_-} |00\rangle \qquad \text{(XII.76)}$$

orm a complete orthonormal eigenset common to N_+ and N_-:

$$\begin{aligned}N_+|n_+ n_-\rangle &= n_+|n_+ n_-\rangle, \\ N_-|n_+ n_-\rangle &= n_-|n_+ n_-\rangle.\end{aligned}$$

Now, if one expresses N and L as functions of the A and $A^\dagger$, one has after some calculation

$$\begin{aligned}N &= N_+ + N_-, \\ L &= N_+ - N_-.\end{aligned}$$

Since the observables N_+ and N_- form a complete set of commuting

observables, their sum N and their difference L has the same property. This is what we wanted to show.

In conclusion, we have at our disposal another complete orthonormal set of eigenvectors of $\mathscr{H}$, namely the set of vectors $|n_+n_-\rangle$. They satisfy the eigenvalue equations

$$\mathscr{H}|n_+n_-\rangle = (n_+ + n_- + 1)\hbar\omega|n_+n_-\rangle \tag{XII.77}$$

$$L|n_+n_-\rangle = (n_+ - n_-)|n_+n_-\rangle. \tag{XII.78}$$

Let us examine the commutation relations of L with the A and $A^\dagger$. A simple calculation yields

$$[L, A_\pm^\dagger] = \pm A_\pm^\dagger \tag{XII.79}$$

$$[L, A_\pm] = \mp A_\pm. \tag{XII.80}$$

Consequently, when they act upon an eigenvector of L, $A_+^\dagger$ and A_- increase L by one unit, $A_-^\dagger$ and A_+ decrease it by one unit. This may be interpreted in various ways. In the quantum theory of charged fields where the field appears as a set of two-dimensional, isotropic oscillators, N_+ is the number of particles with positive charge, N_- that of particles of negative charge, and L the total charge (to within a constant). Following this interpretation $A_+^\dagger$ creates a positive charge, A_- destroys a negative charge; both therefore increase the charge by one unit; in similar fashion $A_-^\dagger$ and A_+ decrease the charge by one unit.

In the theory of crystalline vibrations, the motions of the lattice are likewise represented by a set of isotropic two-dimensional oscillators; the quanta of oscillation are called *phonons*. The representation in terms of phonons of type 1 and 2 corresponds to standing waves; the representation in terms of phonons of type + and − corresponds to traveling waves "propagating" in one or the other direction. In the problems of scattering (neutrons, X rays, etc.) by a crystal lattice, the representation by traveling waves is the most suitable for purposes of calculation.

15. Three-Dimensional Isotropic Oscillator

The three-dimensional isotropic harmonic oscillator is a particle located in a central potential proportional to the square of the distance from the center. Its Hamiltonian is

$$\mathscr{H} = \frac{\mathbf{p}^2}{2m} + \tfrac{1}{2}m\omega^2\mathbf{r}^2. \tag{XII.81}$$

$\mathscr{H}$ is the sum of three terms:

$$\mathscr{H} = \mathscr{H}_x + \mathscr{H}_y + \mathscr{H}_z$$
$$\mathscr{H}_i = \frac{1}{2m}(p_i^2 + m^2\omega^2 r_i^2). \qquad (i = x, y \text{ or } z) \qquad \text{(XII.82)}$$

According to our study of Ch. XII, § 13, the eigenvalues of $\mathscr{H}$ are given by the formula

$$(n+\tfrac{3}{2})\hbar\omega \qquad (n = 0, 1, 2, \ldots, \infty) \qquad \text{(XII.83}$$

and are $\frac{1}{2}(n+1)(n+2)$-fold degenerate. The observables N_x, N_y, N_z form a complete set of constants of the motion and the eigenvectors $|n_x n_y n_z\rangle$ of their basis are labelled by the three corresponding eigenvalues n_x, n_y, and n_z. These vectors are deduced by the formula

$$|n_x n_y n_z\rangle = (n_x! n_y! n_z!)^{-\frac{1}{2}} a_x^{\dagger n_x} a_y^{\dagger n_y} a_z^{\dagger n_z} |000\rangle \qquad \text{(XII.84)}$$

from the ground-state vector $|000\rangle$, itself defined (to within a constant) by the three equations

$$a_x|000\rangle = a_y|000\rangle = a_z|000\rangle = 0. \qquad \text{(XII.85)}$$

Let us introduce the angular momentum

$$\boldsymbol{l} \equiv \boldsymbol{r} \times \boldsymbol{p}.$$

According to the well-known property of the Hamiltonian for a central potential (Ch. IX), $\mathscr{H}$, $\boldsymbol{l}^2$ and l_z likewise constitute a complete set of commuting observables. The eigenvectors $|nlm\rangle$ common to these three observables are labelled by the three quantum numbers n, l, m, and the corresponding eigenvalues of $\mathscr{H}$, $\boldsymbol{l}^2$ and l_z are respectively $(n+\frac{3}{2})\hbar\omega$, $l(l+1)\hbar^2$ and $m\hbar$. The vectors $|nlm\rangle$ form a complete orthonormal set of eigenvectors of $\mathscr{H}$. They are derived from the vectors $|n_x n_y n_z\rangle$ by a unitary transformation. The explicit construction of these vectors will not be given here [1]). We shall merely determine

[1]) In the $\{r\}$ representation, $|n_x n_y n_z\rangle$ is represented by the wave function $\psi_{n_x}(x)\,\psi_{n_y}(y)\,\psi_{n_z}(z)$; $|nlm\rangle$ is represented by the function

$$\Psi_{nlm}(\boldsymbol{r}) \equiv \frac{y_{nl}(r)}{r} Y_l^m(\theta, \varphi),$$

$y_{nl}(r)$ is the solution which vanishes at the origin — and is regular at infinity — of the differential equation

$$\left[-\frac{\hbar^2}{2m^2}\frac{\mathrm{d}^2}{\mathrm{d}r^2} + \frac{l(l+1)\hbar^2}{2mr^2} + \tfrac{1}{2}m\omega^2 r^2\right] y_{nl} = (n+\tfrac{1}{2})\hbar\omega y_{nl}.$$

the values which the quantum numbers l and m may take when one fixes n; in other words, we try to find the different possible states of the angular momentum corresponding to each energy level.

The great similarity between the operator l_z considered here and the operator L of the preceding section suggests an analogous change of variable. Let us introduce the operators A_m ($m=1, 0, -1$) defined by

$$\begin{aligned} A_1 &= \tfrac{1}{2}\sqrt{2}[a_x - ia_y] \\ A_0 &= a_z \\ A_{-1} &= \tfrac{1}{2}\sqrt{2}[a_x + ia_y] \end{aligned} \tag{XII.86}$$

and the Hermitean conjugate operators $A_m{}^\dagger$. The A_m and $A_m{}^\dagger$ satisfy commutation relations analogous to relations (XII.73) and can be interpreted respectively as destruction and creation operators of quanta of the type m. The number of quanta of type m is represented by the operator $N_m = A_m{}^\dagger A_m$. Clearly, N_1, N_0, and N_{-1} form a complete set of commuting observables, and

$$\begin{aligned} \mathscr{H} &= (N_1 + N_0 + N_{-1} + \tfrac{3}{2})\hbar\omega, \\ N &= N_1 + N_0 + N_{-1}. \end{aligned}$$

To each triplet of eigenvalues (n_1, n_0, n_{-1}) there corresponds an eigenvector common to these three observables, namely the vector

$$|n_1 n_0 n_{-1}\rangle = (n_1!\, n_0!\, n_{-1}!)^{-\frac{1}{2}} A_1{}^{\dagger n_1} A_0{}^{\dagger n_0} A_{-1}{}^{\dagger n_{-1}} |000\rangle.$$

The ensemble of these vectors forms a complete set of eigenvectors of $\mathscr{H}$. Indeed,

$$\begin{aligned} \mathscr{H}|n_1 n_0 n_{-1}\rangle &= (n + \tfrac{3}{2})\hbar\omega |n_1 n_0 n_{-1}\rangle, \\ n &= n_1 + n_0 + n_{-1}. \end{aligned}$$

The vectors we have just formed are not, in general, eigenvectors of $\boldsymbol{l}^2$, but they are eigenvectors of l_z, because

$$l_z = (N_1 - N_{-1})\hbar, \tag{XII.87}$$

and consequently

$$m = n_1 - n_{-1}. \tag{XII.88}$$

Consider the subspace of the eigenvectors of $\mathscr{H}$ corresponding to

the eigenvalue $(n+\frac{3}{2})\hbar\omega$. The $\frac{1}{2}(n+1)(n+2)$ vectors $|n_1 n_0 n_{-1}\rangle$ which span it $(n_1+n_0+n_{-1}=n)$ form a complete orthonormal set of eigenvectors of l_z. According to equation (XII.88), the quantum number m may take on all integral values from $-n$ to $+n$. It is easy to determine the number c_m of linearly independent vectors corresponding to each of these values of m; the result is given by the following table:

$$\begin{array}{lccccccc} |m|= & n & n-1 & n-2 \ldots & n-2s & n-(2s+1) & n-(2s+2) \ldots \\ c_m= & 1 & 1 & 2 \quad \ldots & s+1 & s+1 & s+2 \ldots \end{array} \qquad \text{(XII.89)}$$

Now according to the properties of angular momentum, to each eigenvalue of $\boldsymbol{l}^2$, i.e. to each value of l, corresponds a certain number of sets of $(2l+1)$ vectors of well-defined angular momentum (lm), m in each set taking the $(2l+1)$ integral values contained between $-l$ and $+l$. Let d_l be that number. It is obvious that

$$c_m = \sum_{l \geqslant m} d_l,$$

hence that

$$d_l = c_l - c_{l+1}.$$

Referring to Table XII.89, we see that $d_l = 1$ for $l = n, n-2, \ldots, n-2s, \ldots$, that is to say, for all integral values of l of parity $(-)^n$ contained between 0 and n (limits included), and that $d_l = 0$ for all other values of l.

In conclusion, to each eigenvalue $(n+\frac{3}{2})\hbar\omega$ of the energy correspond $\frac{1}{2}(n+1)(n+2)$ states of well-defined angular momentum (lm). For

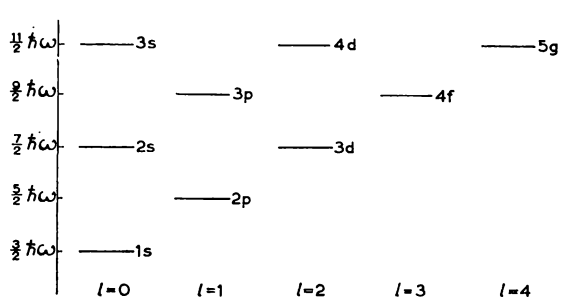


Fig. XII.1. Spectrum of the three-dimensional harmonic oscillator.

each of the possible values of l, there exist $(2l+1)$ eigenstates corresponding, respectively, to the $(2l+1)$ values of m, ranging from $-l$ to $+l$. The values which l may take are

$$n, n-2, \ldots, 0 \quad \text{if} \quad (-)^n = 1 \qquad [\tfrac{1}{2}(n+2) \text{ distinct values}]$$
$$n, n-2, \ldots, 1 \quad \text{if} \quad (-)^n = -1 \qquad [\tfrac{1}{2}(n+1) \text{ distinct values}].$$

The spectroscopic diagram of Fig. XII.1 represents the ground state and first few excited states of the three-dimensional isotropic harmonic oscillator. It is instructive to compare this diagram with the one of the hydrogen atom (Fig. XI.1).

EXERCISES AND PROBLEMS

1. Let a and $a^\dagger$ be two Hermitean conjugate operators such that $[a, a^\dagger] = 1$. We put $N = a^\dagger a$. Show that

(*i*) $[N, a^p] = -pa^p$; $[N, a^{\dagger p}] = +pa^{\dagger p}$ (p integer > 0);

(*ii*) the only algebraic functions of a and $a^\dagger$ which commute with N are the functions of N.

2. Show that the operators a and $a^\dagger$ of Problem XII.1 have no inverse.

3. Form the matrices representing the operators q and p in the $\{N\}$ representation (notation of Ch. XII, § 5). Verify that they are Hermitean and satisfy the commutation relation (XII.2). Set up the eigenvalue problem of q in this representation; show that the spectrum of q is, nondegenerate continuous and extends from $-\infty$ to $+\infty$. Form explicitly the eigenvector corresponding to the eigenvalue 0.

4. One wishes to compare the properties of a quantized oscillator in the state $|n\rangle$ to those of a microcanonical ensemble of classical oscillators of the same energy (Ch. XII, § 10). Show that the statistical distribution of the variable q in that quantum state exhibits oscillations which are tighter the larger n; show also that in the limit where $n \to \infty$ its average value over several oscillations tends toward the corresponding distribution of the ensemble of classical oscillators (use the WKB method).

5. Let χ_0, ϖ_0, η_0 be the respective initial values of the average values

$$\chi = \langle q^2 \rangle - \langle q \rangle^2, \quad \varpi = \langle p^2 \rangle - \langle p \rangle^2, \quad \eta = \langle pq + qp \rangle - 2\langle p \rangle \langle q \rangle$$

relating to a wave packet of an harmonic oscillator. Establish the law of evolution of these mean values as function of time. Show that they are functions of the form $A + B\cos 2\omega t + C \sin 2\omega t$, and that χ and ϖ remain constant if, and only if

$$\eta_0 = 0, \qquad \varpi_0 = m^2\omega^2\chi_0.$$

6. The state of a harmonic oscillator is represented at time zero by the minimum wave packet

$$f(q) = (2\pi\sigma)^{-\frac{1}{4}} \exp\left[\frac{\mathrm{i}}{\hbar}\langle p\rangle q - \frac{(q-\langle q\rangle)^2}{4\sigma}\right].$$

Show that this packet remains minimum in the course of time if and only if $\sigma = \hbar/2m\omega$ (cf. Problem XII.5). We assume henceforth that this condition is fulfilled. Show that then $f(q)$ is the wave function representing the vector

$$|f\rangle = \exp\left(\frac{\mathrm{i}}{\hbar}\langle p\rangle q\right) \exp\left(-\frac{\mathrm{i}}{\hbar}\langle q\rangle p\right) |0\rangle.$$

Deduce from this [making use of the identity (XII.29)] that the function $f(q, t)$ is equal, to within a phase factor, to the expression obtained by substituting in $f(q)$ the mean values $\langle q\rangle$, $\langle p\rangle$ at time t for their values at time 0. Determine the coefficients c_n of the expansion of $|f\rangle$ in a series of eigenvectors of the Hamiltonian, and show that

$$|c_n|^2 = \mathrm{e}^{-\alpha}\frac{\alpha^n}{n!}$$

$$[\alpha = E_{\mathrm{cl.}}/\hbar\omega;\ E_{\mathrm{cl.}} = \frac{1}{2m}(\langle p\rangle^2 + m^2\omega^2\langle q\rangle^2)].$$

7. Verify that Bloch's theorem (Ch. XII, § 12) also applies to the classical harmonic oscillator and that the statistical distribution of $\alpha q + \beta p$ for a quantized harmonic oscillator in thermodynamic equilibrium approaches the classical distribution in the limit where $kT \gg \hbar\omega$.

8. Show that the Hamiltonian of a particle of mass m and charge e in a constant magnetic field $\mathscr{H}$ directed along Oz can be written in the form

$$H = \frac{p_z^2}{2m} + H_\varrho,$$

with

$$H_\varrho = \frac{1}{2m}(p_x^2 + p_y^2) - \frac{e}{2mc}\mathscr{H} l_z + \frac{e^2}{8mc^2}\mathscr{H}^2(x^2+y^2).$$

Show that the operators p_z, l_z, H_ϱ form a complete set of commuting constants of the motion, and that their common eigenfunctions are written in cylindrical coordinates (z,ϱ,θ) in the form $\exp(\mathrm{i}kz)\ \exp(\mathrm{i}\lambda\theta)\ v_{\lambda n}(\varrho)$ [k is any real number; $\lambda = 0, \pm 1, \pm 2, \ldots, \pm\infty$; $n = 0, 1, 2, \ldots, \infty$]. The corresponding eigenvalues are, respectively,

$$\hbar k,\ \hbar\lambda,\ (2n+1)\frac{e\hbar}{2mc}\mathscr{H}.$$

Compare these results with those of Problem II.4.

APPENDIX A

DISTRIBUTIONS, δ-"FUNCTIONS" AND FOURIER TRANSFORMATION

PLAN OF THE APPENDIX

I. Elements of Distribution Theory

1. Concept of Functional and Rigorous Treatment of the Continuous Spectrum.
2. Definition of Distributions.
3. Linear Combination of Distributions.
4. Product of Two Distributions.
5. Series and Integrals of Distributions.
6. Derivative of Distributions.

II. Properties of the δ-"Function"

7. Definition of $\delta(x)$.
8. Representation as the Limit of a Kernel of an Integral Operator.
9. Principal Properties.
10. Derivatives of $\delta(x)$.

III. The Fourier Transformation

11. Fourier Transform of a Function. Definition.
12. Integrable Functions.
13. χ Functions.
14. Fourier Transformation of Distributions. Definition.
15. Tempered Distributions.
16. Square-Integrable Functions.
17. Transformation of Folding into Multiplication.

Table of Fourier Transforms

I. ELEMENTS OF DISTRIBUTION THEORY [1])

1. Concept of Functional and Rigorous Treatment of the Continuous Spectrum

The δ-"function" of Dirac which enables us to treat the continuous spectrum in a manner analogous to the discrete spectrum is not a well-defined mathematical object. If one wishes to introduce into the theory in a rigorous way observables possessing a continuous spectrum, one must set up the eigenvalue problem in a different manner.

In fact, the eigenfunctions of the observables of Wave Mechanics enter only through their scalar product with the wave functions, that is to say by their scalar product with square-integrable functions. Let F be one of these eigenfunctions, and ψ an arbitrarily chosen wave function; the scalar product $\langle \psi, F \rangle$ (notation of Ch. V) may be regarded as an antilinear functional of ψ, or better, as a linear functional of ψ^*. Let us denote the latter by $\hat{F}$; by definition

$$\hat{F}[\psi] = \langle \psi^*, F \rangle.$$

It is the functionals thus associated with each eigenfunction which enter into the theory, and not the eigenfunctions themselves.

Now these functionals belong to a certain class of functionals, called *distributions*, for which one can define essentially the same operations of algebra and analysis as for the functions. Therefore, it is possible to reformulate Wave Mechanics in rigorous fashion by defining the operators of the theory as *operators acting upon the distributions* and not upon the functions; *the eigensolutions of a Hermitean operator are then particular distributions*: they are *the linear and continuous functionals of bounded, square-integrable functions which satisfy the eigenvalue equation of that operator.*

More generally, let X and Ξ be two observables whose spectra are assumed, for simplicity, to be entirely continuous and non-degenerate; let $\langle \xi | x \rangle$ be the unitary transformation matrix permitting the passage from the $\{X\}$ representation to the $\{\Xi\}$ representation. In a rigorous formulation of Quantum Theory, $\langle \xi | x \rangle$ simultaneously represents:

(i) the ensemble of eigensolutions of X in the $\{\Xi\}$ representation,

[1]) Cf. L. Schwartz, *Théorie des distributions* (Paris, Hermann, 1950–1951); see also by the same author, *Les Méthodes Mathématiques de la Physique*, Sorbonne Lectures (Paris, 1955). I. Halperin, *Introduction to the Theory of Distributions*, based on the lectures given by L. Schwartz (Toronto, University of Toronto Press, 1952).

i.e. a certain set of functionals of functions of ξ labelled by the index x;

(*ii*) the ensemble of eigensolutions of Ξ in the $\{X\}$ representation, i.e. a certain set of functionals of functions of x labelled by the index ξ.

In this section, we give the definition of the distributions and state their main properties without proof.

2. Definition of Distributions

Let us denote by $\varphi(x_1, \ldots, x_n)$, or more simply by $\varphi(x)$, a function of n continuous variables $x_1, \ldots, x_n$ whose non-zero values are all contained in a finite domain of these variables, and which has derivatives to all orders with respect to these variables (indefinitely differentiable functions with bounded support).

By definition, *a distribution $T[\varphi]$ is a linear and continuous functional of the functions φ.*

Linearity means that for all linear combinations $\lambda_1\varphi_1 + \lambda_2\varphi_2$, one has

$$T[\lambda_1\varphi_1 + \lambda_2\varphi_2] = \lambda_1 T[\varphi_1] + \lambda_2 T[\varphi_2].$$

Continuity means that for any sequence $\varphi_1, \varphi_2, \ldots, \varphi_j, \ldots$ of functions φ such that $\lim_{j\to\infty} \varphi_j = \varphi$, one has

$$\lim_{j\to\infty} T[\varphi_j] = T[\varphi].$$

To any *locally integrable* function f – that is to say any function whose integral [1]) over any finite interval exists – corresponds a distribution $\hat{f}$ defined by the scalar product

$$\hat{f}[\varphi] = \int f(x)\, \varphi(x)\, \mathrm{d}x = \langle \varphi^*, f \rangle. \tag{A.1}$$

Two locally integrable functions define the same distribution if they are equal almost everywhere (i.e. everywhere except on a set of measure zero). In particular, the wave functions of Wave Mechanics (square-integrable functions) define distributions.

---

[1]) The integrals with which one is dealing throughout this theory are integrals in the sense of Lebesgue. The Lebesgue integral reduces to the integral in the usual sense (Riemann integral) whenever the latter has a meaning; however, the Lebesgue integral exists in cases where the Riemann integral is not defined.

With the function $1/x$ no distribution is associated since this function is not integrable at the point $x=0$. But one can define the distribution

$$\mathrm{PP}\,\frac{1}{x}[\varphi] \equiv \mathrm{PP}\int \frac{\varphi(x)}{x}\,dx \tag{A.2}$$

where PP denotes the Cauchy principal part of the integral

$$\mathrm{PP}\int_{-\infty}^{+\infty} = \lim_{\varepsilon\to 0}\Big\{\int_{-\infty}^{-\varepsilon} + \int_{\varepsilon}^{+\infty}\Big\}.$$

The "Dirac function" $\delta(x)$ defines the distribution

$$\delta[\varphi] = \varphi(0). \tag{A.3}$$

Likewise, the "function" $\delta(x-x_0)$ defines the distribution

$$\delta_{x_0}[\varphi] = \varphi(x_0). \tag{A.4}$$

REMARK. A distribution can eventually be defined over a larger function space than the φ-space. Indeed, if $U[\psi]$ is a linear and continuous functional of the functions ψ of a larger function space than the φ-space, the functional $U[\varphi]$ is well defined and is linear and continuous over φ-space: U is a distribution.

Examples:

δ_{x_0} is defined over the space of the functions $\alpha(x)$ continuous at $x=x_0$:

$$\delta_{x_0}[\alpha] = \alpha(x_0).$$

$\hat{\Psi}$, a distribution corresponding to a square-integrable function, is defined in the space of the square-integrable functions $\psi(x)$:

$$\hat{\Psi}[\psi] = \int \Psi\psi\,dx = \langle\psi^*, \Psi\rangle.$$

The linear and continuous functionals of the wave functions of Wave Mechanics are particular distributions.

3. Linear Combination of Distributions

$T=\lambda_1 T_1+\lambda_2 T_2$ is a distribution defined by

$$T[\varphi] = \lambda_1 T_1[\varphi]+\lambda_2 T_2[\varphi]$$

(λ_1, λ_2 are given complex constants).

4. Product of Two Distributions

If $\hat{f}$ is the distribution associated with a locally integrable function f, and T an arbitrary distribution, the distribution

$$P = \hat{f}T$$

is well defined if T is a linear, continuous functional of the functions $f\varphi$ and one has, by definition

$$P[\varphi] = T[f\varphi]. \tag{A.5}$$

The product of two distributious *does not always exist*. If f has derivatives of all orders, $\hat{f}T$ exists for any T. If f is continuous at the point x_0,

$$(\hat{f}\delta_{x_0})[\varphi] = f(x_0)\,\varphi(x_0). \tag{A.6}$$

If f and g are square-integrable functions, the product $\hat{f}\hat{g}$ is well defined. On the other hand, $[\delta(x)]^2$ has no meaning whatsoever, and neither does $(1/\sqrt{|x|})^2$.

As a special case of eq. (A.6) one has the relation

$$x\delta(x) = 0. \tag{A.7}$$

Conversely, if $xT = 0$, T is a multiple of δ: $T = c\delta$ (c = const.).

Therefore, if $f(x)$ and $g(x)$ are connected by the relation

$$xf(x) = g(x),$$

one necessarily has

$$f(x) = \mathrm{PP}\,\frac{g(x)}{x} + c\delta(x), \tag{A.8}$$

where c is a constant to be determined.

5. Series and Integrals of Distributions

If a set of distributions $T_1, T_2, \ldots, T_j, \ldots$ is such that when $j \to \infty$, $T_j[\varphi]$ has a limit for any φ, this limit is a distribution (i.e. a linear continuous functional of the functions φ):

$$T = \lim_{j\to\infty} T_j.$$

Equivalent statement: if the infinite series $\sum_i T_i[\varphi]$ is summable for any φ, its sum defines a distribution; one says that the *series* of distributions $\sum_i T_i$ is *summable*.

If $T(\lambda)$ is a distribution depending on a parameter λ which can vary continuously in a domain Λ, and if the integral

$$I[\varphi] = \int_{\Lambda} T(\lambda)[\varphi]\, \mathrm{d}\lambda$$

converges for any φ, it defines a distribution

$$I = \int_{\Lambda} T(\lambda)\, \mathrm{d}\lambda.$$

An analogous definition holds for multiple integrals.

In particular, if $f(x, \lambda)$ is an integrable function of x (locally) and of λ, the distribution $\hat{f}(\lambda)$ is integrable in λ, and its integral is the distribution $\hat{g}$ associated with the function

$$g(x) = \int f(x, \lambda)\, \mathrm{d}\lambda.$$

If the function $a(k)$ remains smaller than a positive power of $|k|$ when $|k| \to \infty$:

$$|a(k)| \leqslant A|k|^{\alpha} \quad (A \text{ and } \alpha, \text{ positive constants})$$

the integral $\int_{-\infty}^{+\infty} \mathrm{e}^{\mathrm{i}kx} a(k)\, \mathrm{d}k$ is a distribution.

In particular

$$\int_{-\infty}^{+\infty} \mathrm{e}^{\mathrm{i}kx}\, \mathrm{d}k = 2\pi\delta.$$

6. Derivative of Distributions

By definition, the derivative $\partial T/\partial x_i$ of the distribution T is

$$\frac{\partial T}{\partial x_i}[\varphi] = -T\left[\frac{\partial \varphi}{\partial x_i}\right]. \tag{A.9}$$

In particular, if a locally integrable function is differentiable, the derivative of the corresponding distribution is the distribution corresponding to its derivative. Indeed, upon integrating by parts,

$$\hat{f}'[\varphi] = \int f'(x)\, \varphi(x)\, \mathrm{d}x = -\int f(x)\, \varphi'(x)\, \mathrm{d}x = -\hat{f}[\varphi'].$$

All the properties of the derivatives of the functions apply to the distributions. For instance, the derivative of the product $P = \hat{f}T$ is

$$P' = \hat{f}'T + \hat{f}T'. \tag{A.10}$$

Moreover, certain results pertaining to a more or less restricted class of functions apply to *all* distributions without restriction. They are as follows:

1) *The distributions are differentiable to all orders.*

In particular, the locally summable functions

$$\log |x|, \quad 1/r \quad (r = \sqrt{x^2+y^2+z^2})$$

are differentiable to all orders, since they are distributions:

$$\frac{\mathrm{d}}{\mathrm{d}x} \log |x| = \mathrm{PP}\, \frac{1}{x}. \tag{A.11}$$

$$\Delta\left(\frac{1}{r}\right) = -4\pi\delta \qquad [\delta \equiv \delta(x)\, \delta(y)\, \delta(z)]. \tag{A.12}$$

2) *Differentiation is a linear, continuous operation in the space of the distributions*:

If $\lim_{j\to\infty} T_j = T, \qquad \lim_{j\to\infty} T_j' = T'.$

Hence, if a series is summable, it is differentiable term by term under the summation sign $\sum$. Likewise, if $T(\lambda)$ is summable with respect to the parameter λ:

$$I = \int_\Lambda T(\lambda)\, \mathrm{d}\lambda,$$

$\partial T(\lambda)/\partial x_i$ is certainly summable in the same domain of λ and its integral is equal to $\partial I/\partial x_i$.

II. PROPERTIES OF THE δ-"FUNCTION"

7. Definition of $\delta(x)$

It is customary in physics to use the notation $\delta(x-x_0)$ rather than the more correct notation $\delta_{x_0}[\varphi]$. This notation proves to be quite convenient in practice. $\delta(x-x_0)$ appears as a function whose manipulation is governed by somewhat peculiar rules. Distribution Theory, which provides the mathematical justification for these rules, need not be mentioned explicitly.

By definition, $f(x)$ being a well-defined function at the point $x=x_0$,

$$\int f(x)\, \delta(x-x_0)\, \mathrm{d}x \equiv \delta_{x_0}[f(x)] = f(x_0). \tag{A.13}$$

One therefore has formally

$$\delta(x-x_0) = \begin{cases} 0 & \text{if } x \neq x_0 \\ +\infty & \text{if } x = x_0 \end{cases} \quad \text{and} \quad \int_{-\infty}^{+\infty} \delta(x-x_0)\, \mathrm{d}x = 1 \tag{A.14}$$

$\delta(x-x_0)$ is a generalization of the Kronecker symbol

$$\delta_{mn} = \begin{cases} 0 & \text{if} \quad m \neq n \\ 1 & \text{if} \quad m = n. \end{cases}$$

8. Representation as the Limit of a Kernel of an Integral Operator

$\delta(x-x_0)$ may be considered as the limit of a function which exhibits a very sharp peak about x_0, and whose integral over all space remains constant and equal to 1. For instance:

$$\delta(x-x_0) = \frac{1}{\pi} \lim_{L\to\infty} \sin \frac{L(x-x_0)}{x-x_0} \tag{A.15a}$$

$$= \frac{1}{\pi} \lim_{\varkappa\to\infty} \frac{1-\cos \varkappa(x-x_0)}{\varkappa(x-x_0)^2} \tag{A.15b}$$

$$= \frac{1}{\pi} \lim_{\varepsilon\to+0} \frac{\varepsilon}{(x-x_0)^2+\varepsilon^2} \tag{A.15c}$$

$$= \lim_{\eta\to 0} \frac{E(x-x_0+\eta)-E(x-x_0)}{\eta}. \tag{A.15d}$$

In the last expression, $E(x)$ is the Heaviside function:

$$E(x) = \begin{cases} 1 & \text{if} \quad x > 0 \\ 0 & \text{if} \quad x < 0. \end{cases}$$

(The distribution δ is the derivative of the Heaviside distribution.)

With the property (A.15c) goes the limiting property:

$$\lim_{\varepsilon\to+0} \frac{1}{x-x_0 \pm \mathrm{i}\varepsilon} = \mathrm{PP}\,\frac{1}{x-x_0} \mp \mathrm{i}\pi\delta(x-x_0). \tag{A.15e}$$

9. Principal Properties

The main properties of the function $\delta(x)$ are the following:

$$\delta(x) = \delta(-x) \tag{A.16}$$

$$\delta(ax) = \frac{1}{|a|}\,\delta(x) \qquad (a \neq 0) \tag{A.17}$$

$$\delta[g(x)] = \sum_n \frac{1}{|g'(x_n)|}\,\delta(x-x_n) \qquad \begin{matrix}[g(x_n) = 0,\\ g'(x_n) \neq 0]\end{matrix} \tag{A.18}$$

$$x\delta(x) = 0 \tag{A.19}$$

$$f(x)\,\delta(x-a) = f(a)\,\delta(x-a) \tag{A.20}$$

$$\int \delta(x-y)\,\delta(y-a)\,\mathrm{d}y = \delta(x-a) \tag{A.21}$$

$$\delta(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \mathrm{e}^{\mathrm{i}kx}\,\mathrm{d}k. \tag{A.22}$$

All these equalities state that one side can be replaced by the other when it is multiplied by a regular function and integrated over x. One can prove them all rigorously by means of Distribution Theory (cf. Sec. 1). One can also prove them formally (but not rigorously) by showing that the integrals of the product of each term and $f(x)$ are equal for any sufficiently regular function $f(x)$. Thus relations (A.16), (A.17) and (A.18) are proved by performing a suitable change of variable in these integrals. In expression (A.18), the summation must be taken over all zeros of $g(x)$; the expression has a meaning only if $g(x)$ and $g'(x)$ never vanish simultaneously; for example $\delta(x^2)$ is meaningless.

10. Derivatives of $\delta(x)$

The "function" $\delta(x)$ is differentiable to all orders. Its mth derivative $\delta^{(m)}(x)$ is defined by the property

$$\int_{-\infty}^{+\infty} \delta^{(m)}(x)\,f(x)\,\mathrm{d}x = (-)^m\,f^{(m)}(0), \tag{A.23}$$

valid for any function $f(x)$ which is m times differentiable at the point $x=0$. $\delta^{(m)}(x-x_0)$ can be considered as the limit of one or the other of the mth derivatives of the functions occurring on the right-hand side of eqs. (A.15a), (A.15b) and (A.15c). The following properties, which one can deduce formally (but incorrectly) by the usual procedures of the integral calculus, can be proved rigorously by means of Distribution Theory:

$$\delta^{(m)}(x) = (-)^m\,\delta^{(m)}(-x) \tag{A.24}$$

$$\int \delta^{(m)}(x-y)\,\delta^{(n)}(y-a)\,\mathrm{d}y = \delta^{(m+n)}(x-a) \tag{A.25}$$

$$x^{m+1}\,\delta^{(m)}(x) = 0. \tag{A.26}$$

In particular, the first derivative $\delta'(x)$ has the properties

$$\int_{-\infty}^{+\infty} \delta'(x)\,f(x)\,\mathrm{d}x = -f'(0) \tag{A.27}$$

$$\delta'(x) = -\delta'(-x) \tag{A.28}$$

$$\int \delta'(x-y)\,\delta(y-a)\,\mathrm{d}y = \delta'(x-a) \tag{A.29}$$

$$x\delta'(x) = -\,\delta(x) \tag{A.30}$$

$$x^2\,\delta'(x) = 0 \tag{A.31}$$

$$\delta'(x) = \frac{\mathrm{i}}{2\pi}\int_{-\infty}^{+\infty} k\,\mathrm{e}^{\mathrm{i}kx}\,\mathrm{d}k. \tag{A.32}$$

III. THE FOURIER TRANSFORMATION[1])

11. Fourier Transform of a Function. Definition

If $f(x)$ is a (real or complex) function of the variable x, its Fourier transform, if it exists, is the function

$$F(u) \equiv \mathscr{F}[f] = \left(\frac{\alpha}{2\pi}\right)^{\frac{1}{2}}\int_{-\infty}^{+\infty} \mathrm{e}^{-\mathrm{i}\alpha ux}\,f(x)\,\mathrm{d}x \tag{A.33}$$

where α is a constant fixed once and for all (in Wave Mechanics, one takes $\alpha = 1/\hbar$). Provided certain convergence conditions are fulfilled, $f(x)$ is deduced from $F(u)$ by the inverse Fourier transformation

$$f(x) = \mathscr{F}^{\dagger}[F] = \left(\frac{\alpha}{2\pi}\right)^{\frac{1}{2}}\int_{-\infty}^{+\infty} \mathrm{e}^{\mathrm{i}\alpha ux}\,F(u)\,\mathrm{d}u. \tag{A.33†}$$

More generally, if $f(x_1, x_2, \ldots, x_n)$ is a function of n variables $x_1, \ldots, x_n$, its Fourier transform is

$$F(u_1, \ldots, u_n) \equiv \mathscr{F}[f]$$
$$= \left(\frac{\alpha}{2\pi}\right)^{\frac{1}{2}n}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} \mathrm{e}^{-\mathrm{i}\alpha(u_1x_1+\ldots+u_nx_n)}\,f(x_1, \ldots, x_n)\,\mathrm{d}x_1 \ldots \mathrm{d}x_n \tag{A.34}$$

and the inverse transformation is defined by

$$f(x_1, \ldots, x_n) \equiv \mathscr{F}^{\dagger}[F]$$
$$= \left(\frac{\alpha}{2\pi}\right)^{\frac{1}{2}n}\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} \mathrm{e}^{\mathrm{i}\alpha(u_1x_1+\ldots+u_nx_n)}\,F(u_1, \ldots, u_n)\,\mathrm{d}u_1 \ldots \mathrm{d}u_n. \tag{A.34†}$$

---

[1]) Cf. L. Schwartz, *loc. cit.*, footnote p. 463; also E. C. Titchmarsh, *Introduction to the Theory of Fourier Integrals* (2nd Ed., Oxford University Press, 1948).

Provided the Fourier transform of f exists, one has

$$\mathscr{F}[f(cx_1, \ldots, cx_n)] = \frac{1}{|c|^n} F\left(\frac{u_1}{c}, \ldots, \frac{u_n}{c}\right) \quad (c = \text{arbitrary const.}). \tag{A.35}$$

Likewise, and with the same qualifications, one has

$$\mathscr{F}^{\dagger}[F(cu_1, \ldots, cu_n)] = \frac{1}{|c|^n} f\left(\frac{x}{c}, \ldots, \frac{x_n}{c}\right) \quad (c = \text{arbitrary const.}). \tag{A.35†}$$

In what follows we give without proof the main properties of the Fourier transforms of functions (or of distributions) of a variable. All these properties of the one-dimensional Fourier transformation may be easily extended to any number of dimensions.

12. Integrable Functions

Any integrable function $f(x)$ ($\int_{-\infty}^{+\infty} |f(x)|\,dx < \infty$) has a Fourier transform:

$$F(u) = \mathscr{F}[f].$$

$F(u)$ is: (*i*) continuous,

(*ii*) bounded: $|F(u)| \leqslant \int_{-\infty}^{+\infty} |f(x)|\,dx$ for any u,

(*iii*) zero at infinity: $F(u) \xrightarrow[|u|\to\infty]{} 0$.

If $f(x)$ is m times *continuously differentiable*, and if its m derivatives are integrable,

$$\mathscr{F}[f^{(m)}] = (\mathrm{i}\alpha u)^m F(u). \tag{A.36}$$

If $x^m f(x)$ is integrable, $F(u)$ is m times continuously differentiable and

$$F^{(m)}(u) = \mathscr{F}[(-\mathrm{i}\alpha x)^m f(x)]. \tag{A.37}$$

(The properties of the transformation $\mathscr{F}^{\dagger}$ are deduced from the foregoing by changing i into $-$i in all formulae.)

13. χ Functions

We designate by $\chi(x)$ an indefinitely differentiable function which tends, along with all its derivatives, asymptotically to zero more rapidly than any power of $|x|$:

$$|x|^l \chi^{(m)}(x) \xrightarrow[|x|\to\infty]{} 0 \qquad \text{for any } m \text{ and } l.$$

More generally, $\chi(x_1, \ldots, x_n)$ designates a function of n variables, indefinitely differentiable and such that

$$R^l \frac{\partial^m \chi}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \ldots \partial x_n^{\alpha_n}} \xrightarrow[R \to \infty]{} 0 \qquad [R \equiv (x_1^2 + x_2^2 + \ldots + x_n^2)^{\frac{1}{2}}]$$

for any l, m, and any choice of the indices $\alpha_1, \alpha_2, \ldots, \alpha_n$ $(\alpha_1 + \ldots + \alpha_n = m)$.

The functions φ of Sec. I are particular χ functions; on the other hand, the functions χ are not all φ functions [example: $\exp(-R^2)$].

Since the functions χ are integrable, the properties of the preceding section apply. In addition, we have the following result:

The Fourier transform $\mathscr{F}\chi$ and the inverse Fourier transform $\mathscr{F}^{\dagger}\chi$ of a function χ are likewise χ functions (of the variables $u_1, u_2, \ldots, u_n$). Moreover, the Fourier transformation has the *reciprocity property*:

$$\mathscr{F}\mathscr{F}^{\dagger}\chi = \mathscr{F}^{\dagger}\mathscr{F}\chi = \chi$$

($\mathscr{F}\mathscr{F}^{\dagger}\chi$ means: Fourier transform of $\mathscr{F}^{\dagger}\chi$).

14. Fourier Transformation of Distributions. Definition

If T is a distribution, its Fourier transform $\mathscr{F}T$ is the functional defined by

$$\mathscr{F}T[\varphi] = T[\mathscr{F}\varphi],$$

its inverse Fourier transform is the functional

$$\mathscr{F}^{\dagger}T[\varphi] = T[\mathscr{F}^{\dagger}\varphi].$$

Since $\mathscr{F}\varphi$, $\mathscr{F}^{\dagger}\varphi$ are not necessarily functions of the type φ, it may happen that the functionals $T[\mathscr{F}\varphi]$, $T[\mathscr{F}^{\dagger}\varphi]$ do not exist; in that case T has neither a Fourier transform nor an inverse Fourier transform.

If $\hat{f}$ is the distribution corresponding to a function $f(x)$, and if $F(u)$ is the Fourier transform of $f(x)$, assuming that it exists,

$$\begin{aligned}
\mathscr{F}\hat{f}[\varphi] = \hat{f}[\mathscr{F}\varphi] &= \int_{-\infty}^{+\infty} f(x) \left[\left(\frac{\alpha}{2\pi}\right)^{\frac{1}{2}} \int_{-\infty}^{+\infty} \mathrm{e}^{-\mathrm{i}\alpha u x} \varphi(u)\, \mathrm{d}u \right] \mathrm{d}x \\
&= \int_{-\infty}^{+\infty} \varphi(u) \left[\left(\frac{\alpha}{2\pi}\right)^{\frac{1}{2}} \int_{-\infty}^{+\infty} \mathrm{e}^{-\mathrm{i}\alpha u x} f(x)\, \mathrm{d}x \right] \mathrm{d}u \\
&= \hat{F}[\varphi].
\end{aligned}$$

Thus, the Fourier transform of the distribution $\hat{f}$ is the distribution $\hat{F}$ associated with the Fourier transform of the function f.

15. Tempered Distributions

The tempered distributions are by definition the linear and continuous functionals of the functions χ. They are particular distributions.

All the properties of distributions of Sec. I extend to tempered distributions. It suffices to substitute the functions χ for the functions φ in all statements. In particular, the tempered distributions are differentiable to all orders, and *their derivatives are tempered distributions.*

The square-integrable functions, the functions bounded over all space and, more generally, the locally integrable functions $f(x)$ which *increase sufficiently slowly* at infinity (i.e. for which one can find two positive numbers A and α such that $|f(x)| \leqslant A|x|^{\alpha}$ when $|x| \to \infty$) all define tempered distributions. δ, δ_{x_0}, and all their derivatives are tempered distributions.

The solutions of the eigenvalue problems in Wave Mechanics are linear and continuous functionals of the wave functions, that is to say, of the square-integrable functions $\psi(q_1, \ldots, q_R)$. They are, *a fortiori*, linear and continuous functionals of the functions χ: *they are tempered distributions.*

The interest in tempered distributions stems from their remarkable Fourier transformation properties:

If U_x is a tempered distribution [defined for the functions $\chi(x)$]:

1) *Its Fourier transform V_u and its inverse Fourier transform $V_u{}^{\dagger}$ always exist and they are tempered distributions* [defined for the functions $\chi(u)$]. They are respectively defined by

$$V_u[\chi] \equiv \mathscr{F} U_x[\chi] = U_x[\mathscr{F}\chi] \tag{A.38}$$

$$V_u{}^{\dagger}[\chi] \equiv \mathscr{F}^{\dagger} U_x[\chi] = U_x[\mathscr{F}^{\dagger}\chi]. \tag{A.38†}$$

2) *The transformations $\mathscr{F}$ and $\mathscr{F}^{\dagger}$ are reciprocal:*

$$\mathscr{F}^{\dagger}\mathscr{F} U \equiv \mathscr{F}^{\dagger} V = U \tag{A.39}$$

$$\mathscr{F}\mathscr{F}^{\dagger} U \equiv \mathscr{F} V^{\dagger} = U. \tag{A.40}$$

3) *Differentiation is transformed into multiplication by u, and conversely,* according to the law

$$\mathscr{F}(U_x{}^{(m)}) = (\mathrm{i}\alpha u)^m V_u \tag{A.41}$$

$$\mathscr{F}[(-\mathrm{i}\alpha x)^m U_x] = V_u{}^{(m)}. \tag{A.42}$$

16. Square-Integrable Functions

The square-integrable functions define tempered distributions.

If one adopts the convention not to consider as distinct, two functions which are equal almost everywhere (i.e. everywhere except on a set of measure zero), the properties of the Fourier transformation of tempered distributions apply to square-integrable functions. To these properties are added specific properties of square-integrable functions. The principal theorems are the following:

Theorem I. If $f(x)$ is *square-integrable*, the integral

$$\left(\frac{\alpha}{2\pi}\right)^{\frac{1}{2}} \int_{-\xi}^{+\xi} e^{-i\alpha u x} f(x)\, dx$$

converges in the quadratic mean [1]) *toward a square-integrable function*

$$F(u) \equiv \mathscr{F} f(x) = \left(\frac{\alpha}{2\pi}\right)^{\frac{1}{2}} \operatorname*{lqm}_{\xi\to\infty} \int_{-\xi}^{+\xi} e^{-i\alpha u x} f(x)\, dx. \tag{A.43}$$

Theorem II. The correspondence between $f(x)$ and $F(u)$ is *reciprocal* in the sense that

$$f(x) = \mathscr{F}^{\dagger} F(u) = \left(\frac{\alpha}{2\pi}\right)^{\frac{1}{2}} \operatorname*{lqm}_{\lambda\to\infty} \int_{-\lambda}^{+\lambda} e^{i\alpha u x} F(u)\, du. \tag{A.43†}$$

In fact, for any value of x in the vicinity of which $f(x)$ is of bounded variation

$$\mathscr{F}^{\dagger} F(u) = \lim_{\varepsilon\to 0} \frac{f(x+\varepsilon)+f(x-\varepsilon)}{2}. \tag{A.44}$$

Theorem III. Let the square-integrable functions $f(x)$ and $F(u)$ be Fourier transforms of each other. If the derivative $f'(x)$ is square-

---

[1]) Convergence in the quadratic mean (less restrictive than ordinary convergence) of a function $\Phi(u, \xi)$ toward $\Phi(u)$ as $\xi \to X$, denoted here by the symbol

$$\operatorname*{lqm}_{\xi\to X} \Phi(u, \xi) = \Phi(u),$$

means that

$$\lim_{\xi\to X} \int_{-\infty}^{+\infty} |\Phi(u, \xi) - \Phi(u)|^2\, du = 0,$$

in other words, that $\Phi(u, \xi)$ tends toward $\Phi(u)$ almost everywhere.

integrable, its Fourier transform, $i\alpha u F(u)$, is also square-integrable; conversely, if $i\alpha u F(u)$ is square-integrable, $f(x)$ is differentiable and $f'(x)$ is the inverse Fourier transform of $i\alpha u F(u)$. An analogous property exists for the pair $xf(x)$ and $(i/\alpha)F'(x)$.

N.B. Even if $f(x)$, considered as a function, is not differentiable everywhere, the (tempered) distribution $\hat{f}'_x$ always exists; its Fourier transform is the (tempered) distribution $i\alpha u \hat{F}_u$; the latter, considered as a function, may not be square-integrable.

The Fourier transformation preserves the scalar product of square-integrable functions (Parseval):

Theorem IV. If $F(u)$ and $G(u)$ are the respective Fourier transforms of the square-integrable functions $f(x)$ and $g(x)$, one has

$$\langle g, f \rangle = \langle G, F \rangle, \tag{A.45}$$

or stated differently:

$$\int_{-\infty}^{+\infty} g^*(x)\, f(x)\, \mathrm{d}x = \int_{-\infty}^{+\infty} G^*(u)\, F(u)\, \mathrm{d}u.$$

A particular case ($f = g$) of this theorem is the *conservation of the norm*

$$\int_{-\infty}^{+\infty} |f(x)|^2\, \mathrm{d}x = \int_{-\infty}^{+\infty} |F(u)|^2\, \mathrm{d}u. \tag{A.46}$$

17. Transformation of Folding into Multiplication

By definition, the folding of two functions, if it exists, is the expression

$$f \circ g \equiv \int_{-\infty}^{+\infty} f(x-t)\, g(t)\, \mathrm{d}t. \tag{A.47}$$

One can likewise define (cf. footnote p. 463) – and we shall not state it explicitly here – the folding of two distributions. In particular

$$\delta \circ T = T \quad (T \text{ an arbitrary distribution})$$

$$\delta \circ f = \int_{-\infty}^{+\infty} \delta(x-t)\, f(t)\, \mathrm{d}t = f(x) \quad (f \text{ an arbitrary function}).$$

The folding is a commutative operation: $f \circ g = g \circ f$.

The interest in the folding concept comes from the following theorem:

Theorem V. Provided that it exists, the folding of two functions $f(x)$ and $g(x)$ has as its Fourier transform $\sqrt{2\pi/\alpha}\, F(u)\, G(u)$, an expression in which $F(u)$ and $G(u)$ are the Fourier transforms of $f(x)$ and $g(x)$, respectively; this transformation is reciprocal.

In particular:

Theorem V'. If $f(x)$ is a square-integrable function, $g(x)$ an integrable function, and $F(x)$ and $G(x)$ their respective Fourier transforms, the folding $f \circ g$ and the product $\sqrt{2\pi/\alpha}\, FG$ are square-integrable functions, and the second is the Fourier transform of the first.

TABLE OF FOURIER TRANSFORMS

$f(x) = \left(\frac{\alpha}{2\pi}\right)^{\frac{1}{2}} \int_{-\infty}^{+\infty} e^{i\alpha ux} F(u)\, du$	$F(u) = \left(\frac{\alpha}{2\pi}\right)^{\frac{1}{2}} \int_{-\infty}^{+\infty} e^{-i\alpha ux} f(x)\, dx$
$f\left(\frac{x}{c}\right)$	$\lvert c\rvert\, F(cu)$
$f(-x)$	$F(-u)$
$f^*(x)$	$F^*(-u)$
$F(x)$	$f(-u)$
$xf(x)$	$\frac{i}{\alpha} F'(u)$
$f'(x)$	$i\alpha u\, F(u)$
$f(x-x_0)$	$e^{-i\alpha ux_0} F(u)$
$e^{i\alpha u_0 x} f(x)$	$F(u-u_0)$
$\delta(x)$	$\left(\frac{\alpha}{2\pi}\right)^{\frac{1}{2}}$
$\delta(x-x_0)$	$\left(\frac{\alpha}{2\pi}\right)^{\frac{1}{2}} e^{-i\alpha ux_0}$
$E(x) = \begin{cases} 1 & \text{if } x>0 \\ 0 & \text{if } x<0 \end{cases}$	$\frac{1}{\sqrt{2\pi\alpha}}\left[\pi\delta(u) - i\,PP\,\frac{1}{u}\right]$
$\left(\frac{\varkappa}{\sqrt{\pi}}\right)^{\frac{1}{2}} e^{-\frac{1}{2}\varkappa^2 x^2}$	$\left(\frac{\alpha}{\varkappa\sqrt{\pi}}\right)^{\frac{1}{2}} e^{-\alpha^2 u^2/2\varkappa^2}$ (Re $\varkappa>0$, Re $\varkappa^2>0$)
$\frac{1}{\sqrt{2a}}[E(x+a)-E(x-a)] = \begin{cases} \frac{1}{\sqrt{2a}} & \text{if } \lvert x\rvert<a \\ 0 & \text{if } \lvert x\rvert>a \end{cases}$	$\left(\frac{\alpha a}{\pi}\right)^{\frac{1}{2}} \frac{\sin \alpha a u}{\alpha a u}$ (a real >0)
$\sqrt{\gamma}\, e^{-\gamma\lvert x\rvert}$	$\left(\frac{2\alpha\gamma^3}{\pi}\right)^{\frac{1}{2}} \frac{1}{\gamma^2+\alpha^2u^2}$ (Re $\gamma>0$)
$i\sqrt{2\gamma}\, e^{-\gamma x} E(x) = \begin{cases} i\sqrt{2\gamma}\, e^{-\gamma x} & \text{if } x>0 \\ 0 & \text{if } x<0 \end{cases}$	$\left(\frac{\alpha\gamma}{\pi}\right)^{\frac{1}{2}} \frac{1}{\alpha u - i\gamma}$ (Re $\gamma>0$)

N.B. The functions of the last four lines are normalized to unity:

$$\int_{-\infty}^{+\infty} |f(x)|^2\, dx = \int_{-\infty}^{+\infty} |F(u)|^2\, du = 1.$$

APPENDIX B

SPECIAL FUNCTIONS AND ASSOCIATED FORMULAE

PLAN OF THE APPENDIX

I. Laplace's Equation, Laguerre Polynomials, Coulomb Functions

1. Laplace's Equation and Confluent Hypergeometric Series.
2. Laguerre Polynomials.
3. Eigenfunctions of Hydrogen-like Atoms.
4. Pure Coulomb Wave.
5. Spherical Coulomb Functions.

II. Spherical Bessel Functions

6. Spherical Bessel Functions.

III. Harmonic Oscillator and Hermite Polynomials

7. Hermite Polynomials.
8. Eigenfunctions of the Harmonic Oscillator.

IV. Associated Legendre Functions and Legendre Polynomials; Spherical Harmonics

9. Legendre Polynomials and Associated Legendre Functions.
10. Spherical Harmonics.
11. Various Expansions and Formulae.

I. LAPLACE'S EQUATION, LAGUERRE POLYNOMIALS. COULOMB FUNCTIONS

1. Laplace's Equation and Confluent Hypergeometric Series

Laplace's Equation

$$\left[z \frac{d^2}{dz^2} + (\beta - z) \frac{d}{dz} - \alpha \right] f(z) = 0 \tag{B.1}$$

(α, β are any complex constants).

Confluent Hypergeometric Series $F(\alpha|\beta|z)$.

By definition:

$$\begin{aligned} F(\alpha|\beta|z) &= 1 + \frac{\alpha}{\beta} \cdot \frac{z}{1!} + \frac{\alpha(\alpha+1)}{\beta(\beta+1)} \cdot \frac{z^2}{2!} + \dots \\ &= \sum_{n=0}^{\infty} \frac{\Gamma(\alpha+n)\,\Gamma(\beta)}{\Gamma(\alpha)\,\Gamma(\beta+n)} \cdot \frac{z^n}{n!}. \end{aligned} \tag{B.2}$$

This series:

(*i*) is well-defined for any α and β provided that $\beta \neq -p$ (p integer $\geqslant 0$);

(*ii*) converges in the entire complex plane z;

(*iii*) is a polynomial of degree p (p integer $\geqslant 0$) if $\alpha = -p$, possesses an essential singularity at infinity if $\alpha \neq -p$;

(*iv*) satisfies the *Kummer relation*:

$$F(\alpha|\beta|z) = e^z F(\beta - \alpha|\beta| - z). \tag{B.3}$$

If they exist [1]), the functions

$$F(\alpha|\beta|z) \quad \text{and} \quad z^{1-\beta} F(\alpha - \beta + 1|2 - \beta|z)$$

are solutions of eq. (B.1).

Solutions in Integral Form (Laplace's Method):

If Γ is a contour of the complex plane t such that $t^{\alpha}(1-t)^{\beta-\alpha} e^{zt}$ takes the same value at both limits, the integral

$$\int_{\Gamma} e^{zt}\, t^{\alpha-1}\, (1-t)^{\beta-\alpha-1}\, dt \tag{B.4}$$

---

[1]) If β is not an integer, these two functions exist and are distinct. If $\beta = 1$, they are identical. If $\beta = 0, -1, -2, \dots$, only the function

$$z^{1-\beta} F(\alpha - \beta + 1\,|\,2 - \beta|\, z)$$

exists. If $\beta = 2, 3, \dots$, only the function $F(\alpha|\beta|z)$ exists.

is a particular solution of eq. (B.1).

We assume from now on:

$$\alpha \text{ non-integer}; \qquad \beta = b \qquad (b \text{ integer} > 0).$$

To the closed contour Γ_0 surrounding the points $t=0$ and $t=1$ (Fig. B.1a) [1]) corresponds a solution of the type (B.4). [By convention,

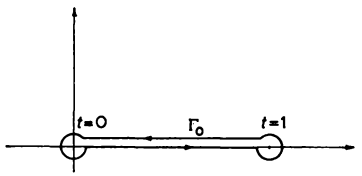


Fig. B.1a.

$\arg t - \arg(1-t) = 0$ on that portion of Γ_0 located on the real axis between 0 and 1 and described in the direction of increasing t.] Since this solution is an entire function of z, it is proportional to $F(\alpha|b|z)$. The coefficient of proportionality is obtained by expanding e^{zt} under the integral sign and making use of the formula:

$$B(x, y) \equiv \frac{\Gamma(x)\,\Gamma(y)}{\Gamma(x+y)} = (1-e^{2\pi i y})^{-1} \int_{\Gamma_s} t^{x-1}\,(1-t)^{y-1}\,dt \qquad \text{(B.5)}$$

$$(x+y = \text{integer}, \quad y \neq \text{integer}).$$

One has:

$$F(\alpha|b|z) = (1-e^{-2\pi i\alpha})^{-1}\,\frac{\Gamma(b)}{\Gamma(\alpha)\,\Gamma(b-\alpha)} \int_{\Gamma_s} e^{zt}\,t^{\alpha-1}\,(1-t)^{b-\alpha-1}\,dt. \qquad \text{(B.6)}$$

To the loops Γ_1, Γ_2 surrounding $t=0$ and $t=1$, respectively (Fig. B.1b) [2]), correspond two solutions of type (B.4) irregular at

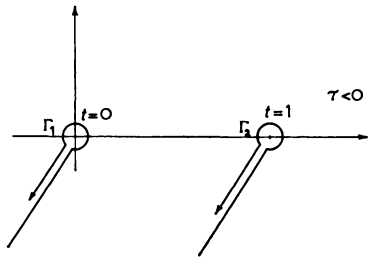


Fig. B.1b.

[1]) Γ_0 is described in the positive sense.

[2]) Γ_1 and Γ_2 are described in the positive sense.

the origin, namely:

$$W_r(\alpha|b|z) = (1-e^{-2\pi i\alpha})^{-1} \frac{\Gamma(b)}{\Gamma(\alpha)\,\Gamma(b-\alpha)} \int_{\Gamma_r} e^{zt}\, t^{\alpha-1}\,(1-t)^{b-\alpha-1}\,dt. \quad \text{(B.7)}$$

$$(r = 1, 2).$$

The condition of convergence of the integral is:

$$\frac{\pi}{2} < \arg z + \tau + 2n\pi < \frac{3\pi}{2}$$

(τ is the argument of the point at infinity of the loops Γ_1 and Γ_2). By convention [1]):

$$-\pi < \arg z < +\pi, \qquad -\pi < \tau < +\pi \qquad \text{sign of } \tau = \text{sign of } (\arg z).$$

At the end of loop Γ_1 and at the beginning of loop Γ_2:

$$\arg t = \tau.$$

At the end of loop Γ_1 and at the beginning of loop Γ_2:

$$\arg (1-t) = \begin{cases} \tau-\pi & \text{if} \quad 0 < \tau < \pi \\ \tau+\pi & \text{if} \quad -\pi < \tau < 0. \end{cases}$$

With these conventions

$$W_1(b-\alpha|b|-z) = e^{-z}\, W_2(\alpha|b|z). \quad \text{(B.8a)}$$

$$W_2(b-\alpha|b|-z) = e^{-z}\, W_1(\alpha|b|z) \quad \text{(B.8b)}$$

$$[-\pi < \arg z < +\pi; \qquad -\pi < \arg(-z) < +\pi]$$

$$F(\alpha|b|z) = W_1(\alpha|b|z) + W_2(\alpha|b|z). \quad \text{(B.9)}$$

Asymptotic expansions of the solutions W_1 *and* W_2 (determined by the saddle-point method):

$$W_1(\alpha|b|z) \underset{|z|\to\infty}{\sim} \frac{\Gamma(b)}{\Gamma(b-\alpha)} (-z)^{-\alpha} \sum_{n=0}^{\infty} \frac{\Gamma(n+\alpha)}{\Gamma(\alpha)} \frac{\Gamma(n+\alpha-b+1)}{\Gamma(\alpha-b+1)} \frac{(-z)^{-n}}{n!} \quad \text{(B.10)}$$

$$[-\pi < \arg(-z) < +\pi].$$

$$W_2(\alpha|b|z) \underset{|z|\to\infty}{\sim} \frac{\Gamma(b)}{\Gamma(\alpha)} e^z\, z^{\alpha-b} \sum_{n=0}^{\infty} \frac{\Gamma(n+1-\alpha)}{\Gamma(1-\alpha)} \cdot \frac{\Gamma(n+b-\alpha)}{\Gamma(b-\alpha)} \frac{z^{-n}}{n!} \quad \text{(B.11)}$$

$$(-\pi < \arg z < +\pi).$$

---

[1]) These conventions are meaningful provided $\operatorname{Im} z \neq 0$. To define W_1 and W_2 on the real axis, one must perform an analytic continuation; the functions obtained are different according to whether $\operatorname{Im} z \to 0+$ or $\operatorname{Im} z \to 0-$.

2. Laguerre Polynomials

Definition [1]:

$$L_p^0 = e^z \frac{d^p}{dz^p}(e^{-z} z^p)$$

$$L_p^k = (-)^k \frac{d^k}{dz^k} L_{p+k}^0 \tag{B.12}$$

$$(k, p = 0, 1, 2, \ldots, \infty)$$

L_p^k is a polynomial of degree p having p zeros between 0 and $+\infty$:

$$L_p^k(z) = \frac{[(p+k)!]^2}{p!\, k!} F(-p|k+1|z)$$
$$= \sum_{s=0}^{p} (-)^s \frac{[(p+k)!]^2}{(p-s)!\,(k+s)!\, s!} z^s. \tag{B.13}$$

In particular $L_0^k = k!$

Laplace's Equation

$$\left[z \frac{d^2}{dz^2} + (k+1-z) \frac{d}{dz} + p \right] L_p^k = 0. \tag{B.14}$$

Generating Function

$$\frac{e^{-zt/(1-t)}}{(1-t)^{k+1}} = \sum_{p=0}^{\infty} \frac{t^p}{(p+k)!} L_p^k(z) \qquad (|t| < 1). \tag{B.15}$$

"Orthonormality" Relations

$$\int_0^\infty e^{-z} z^k L_p^k L_q^k \, dz = \frac{[(p+k)!]^3}{p!} \delta_{pq}. \tag{B.16}$$

3. Eigenfunctions of Hydrogen-like Atoms (Schrödinger Theory)

$$a = \frac{a_0}{Z} = \frac{\hbar^2}{Zm'e^2} \qquad (a_0 \approx \text{radius of the Bohr orbit})$$

Ze = nuclear charge; m' = reduced mass of the electron.

Energy Eigenvalues

$$E_n = -\left(\frac{Ze^2}{\hbar c}\right)^2 \frac{m'c^2}{2n^2} = -\frac{1}{n^2} \frac{Ze^2}{2a}.$$

---

[1] Certain authors designate by L_p^k the polynomial equal to $(-)^k L_{p-k}^k$ in our notation.

Eigenfunctions Normalized to Unity in Spherical Coordinates

$$\psi_{nlm}(r, \theta, \varphi) = a^{-3/2} N_{nl} F_{nl}\left(\frac{2r}{na}\right) Y_l^m(\theta, \varphi). \tag{B.17}$$

$$N_{nl} = \frac{2}{n^2}\sqrt{\frac{(n-l-1)!}{[(n+l)!]^3}}. \tag{B.17a}$$

$$F_{nl}(x) = x^l \, e^{-\frac{1}{2}x} L^{2l+1}_{n-l-1}(x) \tag{B.17b}$$

$(n = 1, 2, \ldots, \infty; \quad l = 0, 1, \ldots, n-1; \quad m = -l, -l+1, \ldots, l).$

Mean Values of Powers of r

Recursion relation between mean values pertaining to the same eigenstate (nlm)

$$\frac{s+1}{n^2}\langle r^s\rangle - (2s+1)\, a \langle r^{s-1}\rangle + \frac{s}{4}[(2l+1)^2 - s^2]\, a^2 \langle r^{s-2}\rangle = 0$$

$$(s > -2l-1)$$

$$\left\langle \frac{1}{r} \right\rangle = \frac{1}{n^2 a}; \qquad \left\langle \frac{1}{r^2} \right\rangle = \frac{2}{(2l+1)\, n^3 a^2}$$

$$\langle r\rangle = \tfrac{1}{2}[3n^2 - l(l+1)]\, a; \qquad \langle r^2\rangle = \tfrac{1}{2}[5n^2 + 1 - 3l(l+1)]\, n^2 a^2.$$

Table of the First Few Radial Functions

$$\varrho = \frac{r}{a}, \quad g_{nl}(\varrho) = N_{nl} F_{nl}\left(\frac{2}{n}\varrho\right)$$

$n = 1 \quad g_{1s} = 2e^{-\varrho}$

$n = 2 \quad g_{2s} = \frac{\sqrt{2}}{2}\left(1 - \frac{1}{2}\varrho\right) e^{-\frac{1}{2}\varrho}, \quad g_{2p} = \frac{\sqrt{6}}{12}\varrho\, e^{-\frac{1}{2}\varrho}$

$n = 3 \quad g_{3s} = \frac{2\sqrt{3}}{9}\left(1 - \frac{2}{3}\varrho + \frac{2}{27}\varrho^2\right) e^{-\frac{1}{2}\varrho},$

$$g_{3p} = \frac{8\sqrt{6}}{27}\left(\varrho - \frac{1}{6}\varrho^2\right) e^{-\frac{1}{2}\varrho}, \quad g_{3d} = \frac{2\sqrt{30}}{955}\varrho^3 e^{-\frac{1}{2}\varrho}.$$

4. Pure Coulomb Wave

$\psi_c(r)$ is the scattering wave of a particle by a pure Coulomb potential $ZZ'e^2/r$.

$$k = \frac{mv}{\hbar} = \text{wave number}$$

v = incident velocity

$$\gamma = \frac{ZZ'e^2}{\hbar v} \tag{B.18}$$

$$\psi_c = e^{-\frac{1}{2}\pi\gamma}\,\Gamma(1+i\gamma)\,e^{ikz}\,F[-i\gamma|1|ik(r-z)] \tag{B.19}$$

$$= \psi_1 + \psi_s \tag{B.20}$$

$$\psi_1 = e^{-\frac{1}{2}\pi\gamma}\,\Gamma(1+i\gamma)\,e^{ikz}\,W_1\,[-i\gamma|1|ik(r-z)] \tag{B.21}$$

$$\psi_s = e^{-\frac{1}{2}\pi\gamma}\,\Gamma(1+i\gamma)\,e^{ikz}\,W_2\,[-i\gamma|1|ik(r-z)]. \tag{B.22}$$

Asymptotic Form

$$\psi_1 \underset{k(r-z)\to\infty}{\sim} \exp\,[ikz+i\gamma\ln k(r-z)]\left\{1+\frac{\gamma^2}{ik(r-z)}+\ldots\right\} \tag{B.23}$$

$$\psi_s \underset{k(r-z)\to\infty}{\sim} f_c(\theta)\,\frac{e^{i(kr-\gamma\ln 2kr)}}{r}\left\{1+\frac{(1+i\gamma)^2}{ik(r-z)}+\ldots\right\} \tag{B.24}$$

$$f_c(\theta) = -\frac{\gamma}{2k\sin^2\frac{1}{2}\theta}\exp\,[-i\gamma\ln(\sin^2\tfrac{1}{2}\theta)+2i\sigma_0] \tag{B.25}$$

$$e^{2i\sigma_0} = \frac{\Gamma(1+i\gamma)}{\Gamma(1-i\gamma)}. \tag{B.26}$$

Behavior at the Origin

$$\psi_c(0) = e^{-\frac{1}{2}\pi\gamma}\,\Gamma(1+i\gamma)$$

$$|\psi_c(0)|^2 = \frac{2\pi\gamma}{e^{2\pi\gamma}-1}. \tag{B.27}$$

5. Spherical Coulomb Functions

Differential Equation

In spherical coordinates, the collision problem of § 4 leads for each value l of the orbital angular momentum, to the *radial equation*:

$$\left[\frac{d^2}{dr^2}+\left(k^2-\frac{l(l+1)}{r^2}-\frac{2\gamma k}{r}\right)\right]y_l = 0. \tag{B.28}$$

The spherical Coulomb functions are special solutions of this equation. They are functions of the argument

$$\varrho = kr.$$

They depend upon the energy through k and γ. One defines the

solution, regular (as r^{l+1}) at the origin, $F_l(\gamma; kr)$, and the irregular solutions, G_l, $u_l^{(+)}$ and $u_l^{(-)}$ [singularity in $(1/r)^l$].

By the change of variable and of function

$$z = -2i\varrho, \qquad y_l = e^{i\varrho}\,\varrho^{l+1}\,v_l,$$

equation (B.28) yields Laplace's differential equation:

$$\left[z\frac{d^2}{dz^2} + (2l+2-z)\frac{d}{dz} - (l+1+i\gamma)\right] v_l = 0,$$

of which one knows the solution that is regular at the origin, $F(l+1+i\gamma|2l+2|z)$, and the two irregular solutions

$$W_{1,2}(l+1+i\gamma|2l+2|z).$$

Definition and Connecting Relations

$$F_l(\gamma;\varrho) = c_l\,e^{i\varrho}\,\varrho^{l+1}\,F(l+1+i\gamma|2l+2|-2i\varrho) \tag{B.29a}$$

$$= c_l\,e^{-i\varrho}\,\varrho^{l+1}\,F(l+1-i\gamma|2l+2|+2i\varrho) \tag{B.29b}$$

$$u_l^{(\pm)}(\gamma;\varrho) = \pm\,2i\,e^{\mp i\sigma_l}\,c_l\,e^{\pm i\varrho}\,\varrho^{l+1}\,W_1(l+1\pm i\gamma|2l+2|\mp 2i\varrho) \tag{B.30a}$$

$$= \pm\,2i\,e^{\mp i\sigma_l}\,c_l\,e^{\mp i\varrho}\,\varrho^{l+1}\,W_2(l+1\mp i\gamma|2l+2|\pm 2i\varrho) \tag{B.30b}$$

$$G_l(\gamma;\varrho) = \tfrac{1}{2}(u_l^{(+)}\,e^{i\sigma_l} + u_l^{(-)}\,e^{-i\sigma_l}). \tag{B.31}$$

c_l and σ_l (Coulomb phase shift) are the following functions of γ:

$$c_l = 2^l\,e^{-\frac{1}{2}\pi\gamma}\,\frac{|\Gamma(l+1+i\gamma)|}{(2l+1)!} \qquad \sigma_l = \arg\,\Gamma(l+1+i\gamma), \tag{B.32}$$

or else:

for $l=0$,

$$c_0 = \left(\frac{2\pi\gamma}{e^{2\pi\gamma}-1}\right)^{\frac{1}{2}} \qquad \sigma_0 = \arg\,\Gamma(1+i\gamma) \tag{B.32a}$$

for $l\neq 0$,

$$c_l = c_0\,\frac{1}{(2l+1)!!}\times\prod_{s=1}^{l}\left(1+\frac{\gamma^2}{s^2}\right)^{\frac{1}{2}} \qquad \sigma_l = \sigma_0 + \sum_{s=1}^{l}\tan^{-1}\frac{\gamma}{s} \tag{B.32b}$$

F_l and G_l are real; $\quad u_l^{(-)} = u_l^{(+)*}$

$$F_l = \frac{1}{2i}\,(u_l^{(+)}\,e^{i\sigma_l} - u_l^{(-)}\,e^{-i\sigma_l}) \tag{B.33}$$

$$u_l^{(\pm)} = e^{\mp i\sigma_l}\,(G_l \pm iF_l). \tag{B.34}$$

Asymptotic Forms: $r \to \infty \qquad [\varrho \gg l(l+1)+\gamma^2]$

$$F_l \underset{\varrho\to\infty}{\sim} \sin(\varrho-\gamma \ln 2\varrho-\tfrac{1}{2}l\pi+\sigma_l) \tag{B.35}$$

$$G_l \underset{\varrho\to\infty}{\sim} \cos(\varrho-\gamma \ln 2\varrho-\tfrac{1}{2}l\pi+\sigma_l) \tag{B.36}$$

$$u_l^{(+)} \underset{\varrho\to\infty}{\sim} \exp[\mathrm{i}(\varrho-\gamma \ln 2\varrho-\tfrac{1}{2}l\pi)] \text{ (outgoing wave)} \tag{B.37}$$

$$u_l^{(-)} \underset{\varrho\to\infty}{\sim} \exp[-\mathrm{i}(\varrho-\gamma \ln 2\varrho-\tfrac{1}{2}l\pi)] \text{ (incoming wave).} \tag{B.38}$$

Behavior at the origin: $r \to 0$

$$F_l \underset{\varrho\to 0}{\sim} c_l\, \varrho^{l+1}\left[1+\frac{\gamma}{l+1}\,\varrho+\ldots\right] \tag{B.39}$$

$$G_l \underset{\varrho\to 0}{\sim} \frac{1}{(2l+1)\,c_l}\,\varrho^{-l}\left[1+\begin{cases} \mathrm{O}(\gamma\varrho\ln\varrho) & \text{if } l=0 \\ \mathrm{O}\left(\dfrac{\gamma}{l}\,\varrho\right) & \text{if } l\neq 0 \end{cases}\right]. \tag{B.40}$$

General Behavior of F_l

As ϱ increases from 0 to ∞, the function F_l increases at first as ϱ^{l+1}, then more and more rapidly (exponential behavior) until

$$\varrho = \gamma + \sqrt{\gamma^2+l(l+1)},$$

where it has a point of inflection; the function then oscillates indefinitely between two extreme values which tend asymptotically toward $+1$ and -1, respectively. The period of these oscillations tends asymptotically toward 2π.

Recursion Formulae

$$(2l+1)\left[\gamma+\frac{l(l+1)}{\varrho}\right]F_l = l\sqrt{\gamma^2+(l+1)^2}\,F_{l+1}+(l+1)\sqrt{\gamma^2+l^2}\,F_{l-1} \quad (l\neq 0) \tag{B.41}$$

$$\left(1+\frac{\gamma^2}{l^2}\right)^{\frac{1}{2}}F_{l-1} = \left(\frac{\mathrm{d}}{\mathrm{d}\varrho}+\frac{l}{\varrho}+\frac{\gamma}{l}\right)F_l. \quad (l\neq 0) \tag{B.42}$$

$$\left(1+\frac{\gamma^2}{l^2}\right)^{\frac{1}{2}}F_l = \left(-\frac{\mathrm{d}}{\mathrm{d}\varrho}+\frac{l}{\varrho}+\frac{\gamma}{l}\right)F_{l-1}. \quad (l\neq 0) \tag{B.43}$$

These relations remain valid if one replaces F_l by $U_l \equiv aF_l+bG_l$ (a, b are arbitrarily fixed coefficients independent of l).

Wronskian Relation

$$G_l \frac{\mathrm{d}F_l}{\mathrm{d}\varrho} - F_l \frac{\mathrm{d}G_l}{\mathrm{d}\varrho} = 1 \tag{B.44}$$

from which ($l \neq 0$):

$$G_l F_{l-1} - F_l G_{l-1} = \frac{l}{\sqrt{l^2+\gamma^2}}. \tag{B.45}$$

If $\gamma = 0$, one obtains the spherical Bessel functions, to within the factor ϱ:

$$\begin{aligned} F_l(0;\varrho) &= \varrho\, j_l(\varrho), & G_l(0;\varrho) &= \varrho\, n_l(\varrho) \\ u_l^{(+)}(0;\varrho) &= \varrho\, h_l^{(+)}(\varrho), & u_l^{(-)}(0;\varrho) &= \varrho\, h_l^{(-)}(\varrho) \end{aligned} \tag{B.46}$$

(for definitions of j_l, n_l, $h_l^{(\pm)}$, cf. following section).

II. SPHERICAL BESSEL FUNCTIONS

6. Spherical Bessel Functions

Differential Equation

In polar coordinates, the Schrödinger equation for the free particle leads, for each value l of the orbital angular momentum, to the *radial equation* [1])

$$\left[\frac{1}{\varrho}\frac{\mathrm{d}^2}{\mathrm{d}\varrho^2}\varrho + 1 - \frac{l(l+1)}{\varrho^2}\right] f_l \equiv \left[\frac{\mathrm{d}^2}{\mathrm{d}\varrho^2} + \frac{2}{\varrho}\frac{\mathrm{d}}{\mathrm{d}\varrho} + 1 - \frac{l(l+1)}{\varrho^2}\right] f_l = 0. \tag{B.47}$$

In the complex plane, f_l exhibits an essential singularity at infinity and, in general, a pole of order $l+1$ at $\varrho = 0$.

The spherical Bessel functions are special solutions of this equation. One defines the solution j_l, which is regular (as r^l) at the origin (proper spherical Bessel function) and the irregular solutions n_l (Neumann function), $h_l^{(+)}$ (Hankel function of the first kind) and $h_l^{(-)}$ (Hankel function of the second kind).

---

[1]) One again finds the radial equation (B.28) in the special case $\gamma = 0$ by putting:

$$\varrho = kr, \qquad y_l = kr\, f_l(kr).$$

Definition [1])

$$j_l(\varrho) = \left(\frac{\pi}{2\varrho}\right)^{\frac{1}{2}} J_{l+\frac{1}{2}}(\varrho), \qquad n_l(\varrho) = (-)^l \left(\frac{\pi}{2\varrho}\right)^{\frac{1}{2}} J_{-l-\frac{1}{2}}(\varrho)$$

$$h_l^{(\pm)}(\varrho) = n_l(\varrho) \pm \mathrm{i} j_l(\varrho)$$

(J_ν designates the ordinary Bessel function of order ν)

$$j_l \text{ and } n_l \text{ are real; } h_l^{(-)} = h_l^{(+)*}. \tag{B.48}$$

Explicit Form

$$\begin{gathered} j_l = R_l \frac{\sin\varrho}{\varrho} + S_l \frac{\cos\varrho}{\varrho}, \qquad n_l = R_l \frac{\cos\varrho}{\varrho} - S_l \frac{\sin\varrho}{\varrho} \\ h_l^{(\pm)} = (R_l \pm \mathrm{i} S_l) \frac{\mathrm{e}^{\pm \mathrm{i}\varrho}}{\varrho}. \end{gathered} \tag{B.49}$$

R_l is a polynomial in $1/\varrho$ with real coefficients of degree l and parity $(-)^l$;
S_l is a polynomial in $1/\varrho$ with real coefficients of degree $l-1$ and parity $(-)^{l-1}$.

$$R_l + \mathrm{i} S_l = \sum_{s=0}^{l} \frac{\mathrm{i}^{s-l}}{2^s s!} \frac{(l+s)!}{(l-s)!} \varrho^{-s} \tag{B.50}$$

$$j_0 = \frac{\sin\varrho}{\varrho}, \qquad n_0 = \frac{\cos\varrho}{\varrho}, \qquad h_0^{(\pm)} = \frac{\mathrm{e}^{\pm \mathrm{i}\varrho}}{\varrho}$$

$$j_1 = \frac{\sin\varrho}{\varrho^2} - \frac{\cos\varrho}{\varrho}, \qquad n_1 = \frac{\cos\varrho}{\varrho^2} + \frac{\sin\varrho}{\varrho}, \qquad h_1^{(\pm)} = \left(\frac{1}{\varrho^2} \mp \frac{\mathrm{i}}{\varrho}\right) \mathrm{e}^{\pm \mathrm{i}\varrho}.$$

Asymptotic Forms: $\varrho \to \infty$ $\qquad [\varrho \gg l(l+1)]$

$$\begin{gathered} j_l \underset{\varrho\to\infty}{\sim} \frac{1}{\varrho} \sin(\varrho - \tfrac{1}{2} l\pi), \qquad n_l \underset{\varrho\to\infty}{\sim} \frac{1}{\varrho} \cos(\varrho - \tfrac{1}{2} l\pi) \\ h_l^{(\pm)} \underset{\varrho\to\infty}{\sim} \frac{1}{\varrho} \exp[\pm \mathrm{i}(\varrho - \tfrac{1}{2} l\pi)] \left[1 \pm \mathrm{i} \frac{l(l+1)}{2\varrho} - \cdots\right]. \end{gathered} \tag{B.51}$$

---

[1]) Most authors denote by n_l the same function *with a change in sign*, and introduce under the name of spherical Hankel function of the first and second kind, respectively, the functions:

$$h_l^{(1)} = -\mathrm{i} h_l^{(+)}, \qquad h_l^{(2)} = \mathrm{i} h_l^{(-)}.$$

Behavior at the Origin: $\varrho \to 0$

$$\begin{aligned} j_l &\underset{\varrho\to 0}{\sim} \frac{\varrho^l}{(2l+1)!!}\left[1 - \frac{\varrho^2}{2(2l+3)} + \ldots\right] \\ n_l &\underset{\varrho\to 0}{\sim} \frac{(2l+1)!!}{(2l+1)}\left(\frac{1}{\varrho}\right)^{l+1}\left[1 + \frac{\varrho^2}{2(2l-1)} + \ldots\right]. \end{aligned} \tag{B.52}$$

General Behavior of j_l

As ϱ increases from 0 to $+\infty$, the function ϱj_l increases first as ϱ^{l+1}, then more and more rapidly (exponential behavior) up to the point $\varrho = \sqrt{l(l+1)}$, where it has a point of inflection. The function then oscillates indefinitely between two extreme values which tend asymptotically toward $+1$ and -1, respectively. The asymptotic form (B.51) is a good approximation when $\varrho \gg \frac{1}{2}l(l+1)$, but the amplitude of the oscillations practically attains its asymptotic value (to within 10 %) as soon as $\varrho \gtrsim 2l$.

Recursion Formulae

Below, $f_l \equiv aj_l + bn_l$, a and b being arbitrarily fixed coefficients independent of l. One has $(l \neq 0)$:

$$(2l+1)f_l = \varrho[f_{l+1} + f_{l-1}] \tag{B.53}$$

$$f_{l-1} = \left[\frac{\mathrm{d}}{\mathrm{d}\varrho} + \frac{l+1}{\varrho}\right] f_l = \frac{1}{\varrho^{l+1}} \frac{\mathrm{d}}{\mathrm{d}\varrho}(\varrho^{l+1} f_l) \tag{B.54}$$

$$f_l = \left[-\frac{\mathrm{d}}{\mathrm{d}\varrho} + \frac{l-1}{\varrho}\right] f_{l-1} = -\varrho^{l-1} \frac{\mathrm{d}}{\mathrm{d}\varrho}\left(\frac{f_{l-1}}{\varrho^{l-1}}\right) \tag{B.55}$$

from which:

$$f_l = \left[\varrho^l \left(-\frac{1}{\varrho}\frac{\mathrm{d}}{\mathrm{d}\varrho}\right)^l\right] f_0. \tag{B.56}$$

Wronskian Relation

$$\varrho^2\left[n_l\left(\frac{\mathrm{d}}{\mathrm{d}\varrho} j_l\right) - j_l\left(\frac{\mathrm{d}}{\mathrm{d}\varrho} n_l\right)\right] = 1 \tag{B.57}$$

from which $(l \neq 0)$

$$\varrho^2[n_l j_{l-1} - j_l n_{l-1}] = 1. \tag{B.58}$$

III. HARMONIC OSCILLATOR AND HERMITE POLYNOMIALS

7. Hermite Polynomials

Definition

$$H_n(z) = (-)^n \, e^{z^2} \left(\frac{d^n}{dz^n} e^{-z^2} \right) \qquad (n = 0, 1, 2, \ldots, \infty) \tag{B.59}$$

H_n is polynomial of degree n, parity $(-)^n$, having n zeros.

$$H_n(z) = \begin{cases} (-)^p \dfrac{(2p)!}{p!} F(-p|\tfrac{1}{2}|z^2) & \text{if } n = 2p \\ (-)^p \, 2 \dfrac{(2p+1)!}{p!} z \, F(-p|\tfrac{3}{2}|z^2) & \text{if } n = 2p+1. \end{cases} \tag{B.60}$$

Differential Equation

$$\left[\frac{d^2}{dz^2} - 2z \frac{d}{dz} + 2n \right] H_n(z) = 0. \tag{B.61}$$

Generating Function

$$\exp(-s^2 + 2sz) = \sum_{n=0}^{\infty} \frac{s^n}{n!} H_n(z). \tag{B.62}$$

Recursion Relations

$$\frac{d}{dz} H_n = 2n \, H_{n-1} \tag{B.63}$$

$$\left(2z - \frac{d}{dz} \right) H_n = H_{n-1} \tag{B.64}$$

$$2z \, H_n = H_{n+1} + 2n \, H_{n-1}.$$

Table of the First Six Hermite Polynomials

$$H_0 = 1, \qquad H_1 = 2z$$
$$H_2 = 4z^2 - 2, \qquad H_3 = 8z^3 - 12z,$$
$$H_4 = 16z^4 - 48z^2 + 12, \qquad H_5 = 32z^5 - 160z^3 + 120z.$$

8. Eigenfunctions of the Harmonic Oscillator

$u_n(Q)$ is the eigenfunction normalized to unity, with the eigenvalue

$$E_n = (n + \tfrac{1}{2}) \, \hbar\omega \qquad (n = 0, 1, \ldots, \infty)$$

(phase chosen in such a way that relation (B. 68) is satisfied, and that $u_0(0)$ is real and positive).

Eigenvalue Equation $[Q=(m\omega/\hbar)^{\frac{1}{2}}q]$.

$$\tfrac{1}{2}\left(Q^2-\frac{\mathrm{d}^2}{\mathrm{d}Q^2}\right)u_n(Q)=(n+\tfrac{1}{2})u_n. \tag{B.65}$$

Generating Function

$$\pi^{-\frac{1}{4}}\,\mathrm{e}^{-\frac{1}{2}Q^2}\,\mathrm{e}^{-\frac{1}{2}t^2+\sqrt{2}Qt}=\sum_{n=0}^{\infty}\frac{t^n}{\sqrt{n!}}u_n(Q). \tag{B.66}$$

Orthonormality and Closure Relations

$$\int_{-\infty}^{+\infty}u_nu_p\mathrm{d}Q=\delta_{np}$$

$$\sum_{n=0}^{\infty}u_n^*(Q)u_n(Q')=\delta(Q-Q').$$

Recursion Relations

$$\frac{1}{\sqrt{2}}\left(Q+\frac{\mathrm{d}}{\mathrm{d}Q}\right)u_n=\sqrt{n}\,u_{n-1} \tag{B.67}$$

$$\frac{1}{\sqrt{2}}\left(Q-\frac{\mathrm{d}}{\mathrm{d}Q}\right)u_n=\sqrt{n+1}\,u_{n+1} \tag{B.68}$$

$$Qu_n=\sqrt{\frac{n+1}{2}}\,u_{n+1}+\sqrt{\frac{n}{2}}\,u_{n-1}. \tag{B.69}$$

Parity $(-)^n$: $u_n(-Q)=(-)^n\,u_n(Q)$.

Expression as a Function of Hermite Polynomials

$$u_n=(\sqrt{\pi}2^n n!)^{-\frac{1}{2}}\,\mathrm{e}^{-\frac{1}{2}Q^2}\,H_n(Q). \tag{B.70}$$

IV. ASSOCIATED LEGENDRE FUNCTIONS AND LEGENDRE POLYNOMIALS; SPHERICAL HARMONICS

9. Legendre Polynomials and Associated Legendre Functions

Definitions

Legendre polynomial P_l $(l=0, 1, 2, \ldots, \infty)$:

$$P_l(u)=\frac{1}{2^l l!}\frac{\mathrm{d}^l}{\mathrm{d}u^l}(u^2-1)^l \tag{B.71}$$

P_l is a polynomial of degree l, parity $(-)^l$, having l zeros in the interval $(-1, +1)$.

Associated Legendre function

P_l^m ($l = 0, 1, 2, \ldots, +\infty$; $m = 0, 1, 2, \ldots, l$):

$$P_l^m(u) = (1-u^2)^{\frac{1}{2}m} \frac{\mathrm{d}^m}{\mathrm{d}u^m} P_l(u)$$
$$= \frac{(1-u^2)^{\frac{1}{2}m}}{2^l l!} \frac{d^{l+m}}{\mathrm{d}u^{l+m}} (u^2-1)^l \qquad (-1 \leqslant u \leqslant 1) \qquad \text{(B.72)}$$

$P_l^m(u)$ is the product of $(1-u^2)^{\frac{1}{2}m}$ and a polynomial of degree $(l-m)$ and parity $(-)^{l-m}$, having $(l-m)$ zeros in the interval $(-1, +1)$. In particular:

$$\begin{aligned} m &= l \qquad P_l^l = (2l-1)!!\,(1-u^2)^{\frac{1}{2}l} \\ m &= 0 \qquad P_l^0 = P_l(u) \end{aligned} \qquad \text{(B.73)}$$

$P_l(u)$ is a particular associated Legendre function.

Differential Equations

$$\left[(1-u^2)\frac{\mathrm{d}^2}{\mathrm{d}u^2} - 2u\frac{\mathrm{d}}{\mathrm{d}u} + l(l+1) - \frac{m^2}{1-u^2}\right] P_l^m = 0. \qquad \text{(B.74)}$$

Generating Functions

$$\frac{1}{\sqrt{1-2tu+t^2}} = \sum_{l=0}^{\infty} t^l P_l(u) \qquad \text{(B.75)}$$

$(|t| < 1)$

$$(2m-1)!!\,(1-u^2)^{\frac{1}{2}m} \frac{t^m}{[1-2tu+t^2]^{m+\frac{1}{2}}} = \sum_{l=m}^{\infty} t^l P_l^m(u). \qquad \text{(B.76)}$$

Orthonormality Relations

$$\int_{-1}^{+1} P_k^m P_l^m \,\mathrm{d}u = \frac{2}{2l+1} \frac{(l+m)!}{(l-m)!} \delta_{kl}. \qquad \text{(B.77)}$$

Recursion Relations

$$(2l+1)u\, P_l^m = (l+1-m)\, P_{l+1}^m + (l+m)\, P_{l-1}^m \qquad \text{(B.78)}$$

$$(1-u^2)\frac{\mathrm{d}}{\mathrm{d}u} P_l^m = -lu\, P_l^m + (l+m)\, P_{l-1}^m \qquad \text{(B.79)}$$

$$= (l+1)u\, P_l^m - (l+1-m)\, P_{l+1}^m \qquad \text{(B.80)}$$

(these relations are also valid when $l=0$, using the convention $P_{-1}=0$).

Particular Values

$$\begin{aligned} &P_l(1) = 1, \qquad P_l(-1) = (-)^l \\ &\text{if } m \neq 0, \qquad P_l^m(1) = P_l^m(-1) = 0 \end{aligned} \tag{B. 81}$$

$$P_l^m(0) = \begin{cases} (-)^p \dfrac{(2p+2m)!}{2^l\, p!\, (p+m)!} & \text{if } l-m=2p \\ 0 & \text{if } l-m=2p+1. \end{cases}$$

Table of the First Five Legendre Polynomials

$$P_0 = 1, \qquad P_1 = u \qquad P_2 = \tfrac{1}{2}(3u^2-1),$$
$$P_3 = \tfrac{1}{2}(5u^3-3u) \qquad P_4 = \tfrac{1}{8}(35u^4-30u^2+3).$$

10. Spherical Harmonics

Operators L_x, L_y, L_z in Polar Coordinates

L_x, L_y, L_z are Hermitean differential operators, defined (in a system of units where $\hbar = 1$) by

$$\boldsymbol{L} \equiv \frac{1}{\mathrm{i}} (\boldsymbol{r} \times \nabla).$$

One takes Oz as polar axis; (r, θ, φ) are the polar coordinates of $\boldsymbol{r}$; $\Omega \equiv (\theta, \varphi)$ designates the set of two angular coordinates ($\varphi = 0$: plane zOx; $\varphi = \pi/2$: plane zOy). The element of solid angle is:

$$\mathrm{d}\Omega = \sin\theta \, \mathrm{d}\theta \, \mathrm{d}\varphi.$$

In polar coordinates:

$$L_z = \frac{1}{\mathrm{i}} \frac{\partial}{\partial \varphi} \tag{B.82}$$

$$L_\pm \equiv L_x \pm \mathrm{i} L_y = \mathrm{e}^{\pm \mathrm{i}\varphi} \left[\pm \frac{\partial}{\partial \theta} + \mathrm{i} \cot\theta \frac{\partial}{\partial \varphi} \right] \tag{B.83}$$

$$\boldsymbol{L}^2 \equiv L_x^2 + L_y^2 + L_z^2 = -\left[\frac{1}{\sin\theta} \frac{\partial}{\partial \theta} \left(\sin\theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial \varphi^2} \right]. \tag{B.84}$$

Definition of the Spherical Harmonics $Y_l^m(\theta, \varphi)$:

Eigenfunctions common to the operators $\boldsymbol{L}^2$ and L_z:

$$\begin{aligned} \boldsymbol{L}^2\, Y_l^m &= l(l+1)\, Y_l^m \\ L_z\, Y_l^m &= m\, Y_l^m \end{aligned} \tag{B.85}$$

$$(l = 0, 1, 2, \ldots, \infty; \quad m = -l, -l+1, \ldots, l). \tag{B.86}$$

One completes their definition by adopting the conventions:

a) the Y_l^m are normalized to unity on the unit sphere;

b) their phases are such that the recursion relations (B.89) are satisfied, and that $Y_l^0(0, 0)$ is real and positive.

Orthonormality and Closure Relations

$$\int Y_l^{m*}\, Y_{l'}^{m'}\, \mathrm{d}\Omega \equiv \int_0^{2\pi} \mathrm{d}\varphi \int_0^{\pi} \sin\theta\, \mathrm{d}\theta\, Y_l^{m*}(\theta, \varphi)\, Y_{l'}^{m'}(\theta, \varphi) = \delta_{mm'}\, \delta_{ll'} \tag{B.87}$$

$$\sum_{l=0}^{\infty} \sum_{m=-l}^{+l} Y_l^{m*}(\theta, \varphi)\, Y_l^m(\theta', \varphi') = \frac{\delta(\theta-\theta')\, \delta(\varphi-\varphi')}{\sin\theta} \equiv \delta(\Omega-\Omega'). \tag{B.88}$$

The Y_l^m form a complete orthonormal set of square-integrable functions on the unit sphere.

Recursion Relations

$$\begin{aligned} L_{\pm}\, Y_l^m &= [l(l+1) - m(m \pm 1)]^{\frac{1}{2}}\, Y_l^{m\pm 1} \\ &= [(l \mp m)(l+1 \pm m)]^{\frac{1}{2}}\, Y_l^{m\pm 1} \end{aligned} \tag{B.89}$$

$$\cos\theta\, Y_l^m = \left[\frac{(l+1+m)(l+1-m)}{(2l+1)(2l+3)}\right]^{\frac{1}{2}} Y_{l+1}^m + \left[\frac{(l+m)(l-m)}{(2l+1)(2l-1)}\right]^{\frac{1}{2}} Y_{l-1}^m. \tag{B.90}$$

Parity $(-)^l$

Under a space reflection $(\theta, \varphi) \to (\pi-\theta, \varphi+\pi)$:

$$Y_l^m(\pi-\theta, \varphi+\pi) = (-)^l\, Y_l^m(\theta, \varphi). \tag{B.91}$$

Complex Conjugation

$$Y_l^{m*}(\theta, \varphi) = (-)^m\, Y_l^{-m}(\theta, \varphi). \tag{B.92}$$

Connection with the Associated Legendre Functions $(m \geqslant 0)$

$$Y_l^m(\theta, \varphi) = (-)^m \left[\frac{(2l+1)}{4\pi} \frac{(l-m)!}{(l+m)!}\right]^{\frac{1}{2}} P_l^m(\cos\theta)\, \mathrm{e}^{\mathrm{i}m\varphi}. \tag{B.93}$$

Y_l^m is the product of $\mathrm{e}^{\mathrm{i}m\varphi}$ $\sin^{|m|}\theta$ and a polynomial of degree $(l-|m|)$ in $\cos\theta$ and parity $(-)^{l-m}$. In particular:

$$m = 0 \qquad Y_l^0 = \sqrt{\frac{2l+1}{4\pi}}\, P_l(\cos\theta) \tag{B.94}$$

$$m = l \qquad Y_l^l = (-)^l \left[\frac{(2l+1)}{4\pi} \frac{(2l)!}{2^{2l}\,(l!)^2}\right]^{\frac{1}{2}} \sin^l\theta\, \mathrm{e}^{\mathrm{i}l\varphi}. \tag{B.95}$$

Harmonic Polynomials and *Spherical Harmonics*

The $(2l+1)$ homogeneous polynomials of degree l in x, y, z

$$\begin{gathered}\mathscr{Y}_l^m(\boldsymbol{r}) \equiv r^l\, Y_l^m(\theta, \varphi) \\ (m = -l,\ -l+1,\ \ldots,\ +l)\end{gathered} \tag{B.96}$$

form a set of $(2l+1)$ linearly independent harmonic polynomials of degree l [1]:

$$\triangle\, \mathscr{Y}_l^m(\boldsymbol{r}) = 0. \tag{B.97}$$

Table of the First Few Spherical Harmonics

$$Y_0^0 = \frac{1}{\sqrt{4\pi}}, \qquad Y_1^0 = \sqrt{\frac{3}{4\pi}}\cos\theta, \qquad Y_2^0 = \sqrt{\frac{5}{16\pi}}\,(3\cos^2\theta - 1),$$

$$Y_3^0 = \sqrt{\frac{7}{16\pi}}\,(5\cos^3\theta - 3\cos\theta),$$

$$Y_1^1 = -\sqrt{\frac{3}{8\pi}}\sin\theta\, e^{i\varphi}, \qquad Y_2^1 = -\sqrt{\frac{15}{8\pi}}\sin\theta\cos\theta\, e^{i\varphi},$$

$$Y_3^1 = -\sqrt{\frac{21}{64\pi}}\sin\theta\,(5\cos^2\theta - 1)\, e^{i\varphi},$$

$$Y_2^2 = \sqrt{\frac{15}{32\pi}}\sin^2\theta\, e^{2i\varphi}, \qquad Y_3^2 = \sqrt{\frac{105}{32\pi}}\sin^2\theta\cos\theta\, e^{i\varphi},$$

$$Y_3^3 = -\sqrt{\frac{35}{64\pi}}\sin^3\theta\, e^{3i\varphi}.$$

11. Various Expansions and Formulae

Addition Theorem

$$\frac{2l+1}{4\pi}\, P_l(\cos\alpha) = \sum_{m=-l}^{+l} Y_l^{m*}(\theta_1, \varphi_1)\, Y_l^m(\theta_2, \varphi_2) \tag{B.98}$$

[α = angle between directions (θ_1, φ_1) and (θ_2, φ_2)].

---

[1]) (B.97) arises from the operator identity which holds for any function finite at the point $r = 0$:

$$\triangle \equiv \frac{1}{r}\frac{\mathrm{d}^2}{\mathrm{d}r^2}\, r - \frac{\boldsymbol{L}^2}{r^2}.$$

By definition, $h(x, y, z)$ is a harmonic polynomial if it is homogeneous in x, y, z and satisfies the equation: $\triangle h = 0$. There exist $(2l + 1)$ linearly independent harmonic polynomials of degree l.

Green's Functions of the Operators Δ *and* $\Delta + k^2$ [1])

$$\frac{1}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|} = \sum_{l=0}^{\infty} \frac{r_<^l}{r_>^{l+1}} P_l(\cos\alpha) \tag{B.99}$$

$$\frac{e^{ik|\boldsymbol{r}_1 - \boldsymbol{r}_2|}}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|} = k \sum_{l=0}^{\infty} (2l+1)\, j_l(kr_<)\, h_l^{(+)}(kr_>)\, P_l(\cos\alpha) \tag{B.100}$$

$$\frac{\cos(k|\boldsymbol{r}_1 - \boldsymbol{r}_2|)}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|} = k \sum_{l=0}^{\infty} (2l+1)\, j_l(kr_<)\, n_l(kr_>)\, P_l(\cos\alpha) \tag{B.101}$$

(α = angle between the directions of $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$; $r_<$ = the smaller of the lengths r_1 and r_2; $r_>$ = the larger of the lengths r_1 and r_2).

Formulae (B.100) and (B.101) are valid for any k, even if k is complex.

Expansion of the Plane Wave and the Pure Coulomb Scattering Wave

Polar axis $\equiv z$ axis = direction of the incident wave vector $\boldsymbol{k}$

$$e^{ikz} = \sum_{l=0}^{\infty} (2l+1)\, i^l\, j_l(kr)\, P_l(\cos\theta) \tag{B.102}$$

$$\psi_c = \frac{1}{kr} \sum_{l=0}^{\infty} (2l+1)\, i^l\, e^{i\sigma_l}\, F_l(\gamma; kr)\, P_l(\cos\theta) \tag{B.103}$$

$$f_c(\theta) = \frac{1}{k} \sum_{l=0}^{\infty} (2l+1)\, e^{i\sigma_l} \sin\sigma_l\, P_l(\cos\theta). \tag{B.104}$$

The definitions of γ, ψ_c, $f_c(\theta)$, F_l, σ_l are those of §§ 4 and 5 [equations (B.18), (B.19), (B.25), (B.29) and (B.32)].

With a different choice of polar axis, the expansions (B. 102), (B.103), and (B.104) remain valid since θ represents the angle between the directions of $\boldsymbol{k}$ and $\boldsymbol{r}$.

By applying the addition theorem, one derives the expansions of the same expressions in series of spherical harmonics having the arguments (θ_k, φ_k) and (θ_r, φ_r). For instance:

$$e^{i\boldsymbol{k}\cdot\boldsymbol{r}} = 4\pi \sum_{l=0}^{\infty} \sum_{m=-l}^{+l} i^l\, j_l(kr)\, Y_l^{m*}(\theta_k, \varphi_k)\, Y_l^m(\theta_r, \varphi_r). \tag{B.105}$$

---

[1]) $\Delta\left(\frac{1}{r}\right) = -4\pi\delta(\boldsymbol{r})$, $(\Delta + k^2)\frac{e^{ikr}}{r} = -4\pi\delta(\boldsymbol{r})$, $(\Delta + k^2)\frac{\cos kr}{r} = -4\pi\delta(\boldsymbol{r})$.

Mon Dieu! ma chère, que ton père
a la forme enfoncée dans la matière!
(*Les Précieuses Ridicules*, VI).

PART THREE

SYMMETRIES AND INVARIANCE

NOTE TO THE READER

Pages 498–504 have been omitted from this edition. In the original two-book edition, these pages contained the index to Volume I. The complete index to both Volumes I and II is reprinted on pages 1125–1136.

CHAPTER XIII

ANGULAR MOMENTUM IN QUANTUM MECHANICS

1. Introduction

Symmetry properties of the equations of motion play a large role in Quantum Mechanics as in Classical Mechanics. A systematic and general examination of symmetries and of their consequences will be made in Chapter XV. The present chapter is devoted to symmetries with respect to rotations – one of the most important types. In Quantum Mechanics, as in Classical Mechanics, the rotation of a system brings in its angular momentum, and the rotational invariance of the equations of motion is reflected in the angular momentum being a constant of the motion. Differences with Classical Mechanics arise because the angular momentum is not an ordinary vector but a vector operator, whose three components do not commute.

In section I, we define the angular momentum by the commutation rules for its components J_x, J_y, J_z [relations (XIII.3)] and solve the eigenvalue problem for $\mathbf{J}^2$ and J_z using only these rules and the fact that these three components are observables. The method, due to Dirac, is analogous in many ways to the treatment of the harmonic oscillator in Chapter XII.

Section II is devoted to a special case, the orbital angular momentum of a particle and the construction of the corresponding eigenfunctions (the spherical harmonics).

In section III we establish the connection between rotations and angular momentum. The rotation of a physical system is effected by a certain operator which depends on the components of the total angular momentum, and whose form is given by equation (XIII.60). We show that invariance of the equations of motion under rotation is equivalent to the condition that the Hamiltonian commute with the three angular momentum components, whence the law of conservation of angular momentum.

Experiment shows that most particles possess an intrinsic angular momentum, the spin. The notion of spin is examined in section IV.

Section V is devoted to the important problem of the addition of angular momenta.

The different operators of Quantum Mechanics may be characterized

by their law of transformation under rotations. In particular, there are scalar operators (invariant under rotation), vector operators, and more generally, irreducible tensor operators, whose transformation properties are especially simple. These operators are characterized as well by their very simple commutation rules with the components of angular momentum, from which follow certain noteworthy properties of their representations (Wigner–Eckart theorem). These are treated, with their principal applications, in the sixth and last section of this chapter.

Appendix C is a complement to the present chapter in which are assembled the most important formulae and the principal properties of the different coefficients associated with rotations and with the addition of angular momenta.

I. EIGENVALUES AND EIGENFUNCTIONS OF ANGULAR MOMENTUM

2. Definition of Angular Momentum

We have already encountered the angular momentum operator in the treatment of single particle quantum systems. By definition, the angular momentum $\boldsymbol{l}$ of the particle is given by

$$\boldsymbol{l} \equiv \boldsymbol{r} \times \boldsymbol{p}, \tag{XIII.1}$$

$\boldsymbol{r}$ and $\boldsymbol{p}$ being respectively the position and momentum vectors of the given particle. In Wave Mechanics, $\boldsymbol{l}$ is represented by the vector operator [1] $(-\mathrm{i})\,\boldsymbol{r} \times \boldsymbol{\nabla}$ whose three components are differential operators satisfying the commutation rules

$$[l_x, l_y] = \mathrm{i}\, l_z, \qquad [l_y, l_z] = \mathrm{i}\, l_x, \qquad [l_z, l_x] = \mathrm{i}\, l_y. \tag{XIII.2}$$

Each of them commutes with the square of the angular momentum

$$\boldsymbol{l}^2 \equiv l_x^2 + l_y^2 + l_z^2,$$

i.e.

$$[\boldsymbol{l}, \boldsymbol{l}^2] = 0.$$

These properties have already been derived in § V.18. Recall that the last of them is a simple consequence of relations (XIII.2).

---

[1] In the present chapter the units are chosen so as to have $\hbar = 1$.

$\boldsymbol{r}$, $\boldsymbol{p}$, and $\boldsymbol{l}$ are *vector operators*. A vector operator $\boldsymbol{B}$ is defined by its components B_x, B_y, B_z along three orthogonal axes, where B_x, B_y, B_z are operators in the ordinary sense of the term. Given these three particular components, we can define the component B_u of $\boldsymbol{B}$ in an arbitrary direction $\boldsymbol{u}$ defined by the unit vector $\boldsymbol{u}(u_x, u_y, u_z)$:

$$B_u \equiv (\boldsymbol{u} \cdot \boldsymbol{B}) = u_x B_x + u_y B_y + u_z B_z.$$

We may thus define the components of $\boldsymbol{B}$ along any other system of orthogonal axes. The different manipulations of vector algebra (sum, scalar product, vector product, etc.) apply without change to vector operators if we observe a few precautions regarding the order of the operators involved (cf. § IX.2).

Consider now a quantum system of N particles. As above, we can define the angular momentum of the nth particle $\boldsymbol{l}^{(n)} = \boldsymbol{r}^{(n)} \times \boldsymbol{p}^{(n)}$. The total angular momentum of the system is the vector sum of the angular momenta of the N particles,

$$\boldsymbol{L} = \sum_{n=1}^{N} \boldsymbol{l}^{(n)}.$$

Since each individual angular momentum satisfies the commutation relations (XIII.2), and since the components of any one of them commute with those of any other, we have

$$[L_x, L_y] = \sum_n \sum_{n'} [l_x^{(n)}, l_y^{(n')}] = \sum_n [l_x^{(n)}, l_y^{(n)}] = \mathrm{i} \sum_n l_z^{(n)} = \mathrm{i} L_z.$$

We similarly obtain the two relations resulting from circular permutation of the indices. Thus the components of total angular momentum have the same commutation rules as those of the individual angular momenta themselves.

We are thus led to the following definition of angular momentum: *A vector operator $\boldsymbol{J}$ is an angular momentum if its components are observables satisfying the commutation relations*:

$$[J_x, J_y] = \mathrm{i} J_z, \qquad [J_y, J_z] = \mathrm{i} J_x, \qquad [J_z, J_x] = \mathrm{i} J_y. \tag{XIII.3}$$

From these three relations we can deduce analogous relations for the components of $\boldsymbol{J}$ along any other system of axis. Let J_u, J_v, J_w be the components of $\boldsymbol{J}$ along three orthogonal axes whose unit vectors $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w}$ are so oriented that $\boldsymbol{w} = \boldsymbol{u} \times \boldsymbol{v}$. We can easily prove the relation

$[J_u, J_v] = \mathrm{i}\, J_w$ and the two relations resulting from circular permutation of the indices [1]). More generally, if $\boldsymbol{a}$ and $\boldsymbol{b}$ are any two vectors (or two vector operators that commute with each other and also with $\boldsymbol{J}$), we have

$$[\boldsymbol{a}\cdot\boldsymbol{J}, \boldsymbol{b}\cdot\boldsymbol{J}] = \mathrm{i}\,((\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{J}). \tag{XIII.4}$$

3. Characteristic Algebraic Relations

The square of the angular momentum,

$$\boldsymbol{J}^2 \equiv J_x^2 + J_y^2 + J_z^2,$$

commutes with J_x, J_y and J_z. This property is a consequence of commutation relations (XIII.3), and can be deduced in the same way as the analogous property for the angular momentum $\boldsymbol{l}$ introduced in § 2. We write symbolically,

$$[\boldsymbol{J}, \boldsymbol{J}^2] = 0 \tag{XIII.5}$$

In consequence, $\boldsymbol{J}^2$ commutes with any function of the components of $\boldsymbol{J}$.

Let us introduce the two Hermitean conjugate operators

$$J_+ = J_x + \mathrm{i}\,J_y, \qquad J_- = J_x - \mathrm{i}\,J_y. \tag{XIII.6}$$

The three operators J_+, J_-, and J_z completely define the vector operator $\boldsymbol{J}$ and turn out to be more convenient for algebraic manipulations than J_x, J_y, and J_z. Their commutation relations can be deduced from relations (XIII.3):

$$[J_z, J_+] = J_+ \tag{XIII.7a}$$

$$[J_z, J_-] = -J_- \tag{XIII.7b}$$

$$[J_+, J_-] = 2J_z. \tag{XIII.7c}$$

---

[1]) These relations can also be written in the condensed form ($i, j, k = u, v$ or w)

$$[J_i, J_j] = \mathrm{i}\sum_k \varepsilon_{ijk}\, J_k \quad \text{or} \quad \sum_{ij} \varepsilon_{ijk}\, J_i\, J_j = \mathrm{i}\, J_k,$$

where ε_{ijk} is the completely antisymmetrical tensor with three indices:

$$\varepsilon_{ijk} = \begin{cases} 0 \text{ if two indices are equal.} \\ +1 \text{ if } i, j, k \text{ are obtained by even permutation of } u, v, w. \\ -1 \text{ if } i, j, k \text{ are obtained by odd permutation of } u, v, w. \end{cases}$$

The second form is equivalent to the equation between vector operators,

$$\boldsymbol{J}\times\boldsymbol{J} = \mathrm{i}\,\boldsymbol{J}.$$

Also, from equation (XIII.5), we have

$$[\boldsymbol{J}^2, J_+] = [\boldsymbol{J}^2, J_-] = [\boldsymbol{J}^2, J_z] = 0. \tag{XIII.8}$$

$\boldsymbol{J}^2$ is given by the expression

$$\boldsymbol{J}^2 = \tfrac{1}{2}(J_+ J_- + J_- J_+) + J_z^2$$

from which we deduce, with the aid of (XIII.7*c*), the two identities

$$J_- J_+ = \boldsymbol{J}^2 - J_z(J_z + 1) \tag{XIII.9a}$$

$$J_+ J_- = \boldsymbol{J}^2 - J_z(J_z - 1). \tag{XIII.9b}$$

4. Spectrum of $\boldsymbol{J}^2$ and J_z

Since $\boldsymbol{J}^2$ commutes with each of the components of $\boldsymbol{J}$, one can form a complete set of common eigenfunctions of $\boldsymbol{J}^2$ and one of its components, J_z for example. The fact that J_x, J_y, and J_z are *Hermitean* and obey *commutation relations* (XIII.3) imposes severe limitations on the eigenvalue spectrum.

$\boldsymbol{J}^2$ is a positive-definite Hermitean operator since it is the sum of positive-definite Hermitean operators [1]). Its eigenvalues are necessarily positive (or zero). They shall henceforth be written in the form $j(j+1)$, and labelled with the real quantum number $j(\geqslant 0)$.

Let $|jm\rangle$ be an eigenvector of $\boldsymbol{J}^2$ and J_z corresponding to the eigenvalues $j(j+1)$ and m respectively. We shall say that $|jm\rangle$ represents the state of angular momentum (jm). If $\boldsymbol{J}^2$ and J_z do not form a complete set of commuting observables, there may exist several linearly independent states (jm). In this case, $|jm\rangle$ is a particular ket vector chosen once and for all in the subspace of angular momentum (jm). The following argument is valid for any vector thus chosen. The only conditions imposed on $|jm\rangle$ are:

$$\boldsymbol{J}^2 |jm\rangle = j(j+1)|jm\rangle,$$

$$J_z |jm\rangle = m|jm\rangle.$$

Consider the vectors $J_+|jm\rangle$ and $J_-|jm\rangle$.

---

[1]) This is a consequence of the Hermitean character of J_x, J_y and J_z. For any $|u\rangle$, $\langle u|J_x^2|u\rangle$ is the norm of the vector $J_x|u\rangle$ and it is therefore $\geqslant 0$. Since J_y^2 and J_z^2 have the same property, $\langle u|\boldsymbol{J}^2|u\rangle \geqslant 0$ for any $|u\rangle$.

From identities (XIII.9*a*) and (XIII.9*b*),

$$J_- J_+ |jm\rangle = [j(j+1) - m(m+1)]\,|jm\rangle \equiv (j-m)(j+m+1)\,|jm\rangle \tag{XIII.10a}$$

$$J_+ J_- |jm\rangle = [j(j+1) - m(m-1)]\,|jm\rangle \equiv (j+m)(j-m+1)\,|jm\rangle. \tag{XIII.10b}$$

Thus the norms of the vectors $J_+|jm\rangle$ and $J_-|jm\rangle$ are respectively

$$\langle jm|J_- J_+|jm\rangle = (j-m)(j+m+1)\langle jm|jm\rangle$$
$$\langle jm|J_+ J_-|jm\rangle = (j+m)(j-m+1)\langle jm|jm\rangle.$$

By a fundamental axiom of Hilbert space these norms cannot be negative, and therefore

$$(j-m)(j+m+1) \geqslant 0, \qquad (j+m)(j-m+1) \geqslant 0,$$

from which

$$-j \leqslant m \leqslant j.$$

Moreover, since the vanishing of its norm is a necessary and sufficient condition for the vanishing of a vector,

$$J_+|jm\rangle = 0 \text{ if, and only if } (j-m)(j+m+1) = 0;$$

similarly,

$$J_-|jm\rangle = 0 \text{ if, and only if } (j+m)(j-m+1) = 0.$$

Since m is necessarily in the interval $(-j, +j)$, these vanishing conditions reduce to $m=j$ and $m=-j$ respectively.

If $m \neq j$, the (non-null) vector $J_+|jm\rangle$ is a vector of angular momentum $(j, m+1)$; i.e. following rule (XIII.8):

$$\boldsymbol{J}^2 J_+|jm\rangle = J_+ \boldsymbol{J}^2|jm\rangle = j(j+1)J_+|jm\rangle$$

and, since from (XIII.7*a*)

$$J_z J_+ = J_+(J_z+1) \tag{XIII.11a}$$

we have

$$J_z J_+|jm\rangle = J_+(J_z+1)|jm\rangle = (m+1)J_+|jm\rangle.$$

Using the fact that J_- commutes with $\boldsymbol{J}^2$ and that, from (XIII.7*b*),

$$J_z J_- = J_-(J_z-1) \tag{XIII.11b}$$

we obtain an analogous result for $J_-|jm\rangle$.

Summarizing, we have the important theorem:
If $|jm\rangle$ *is a vector of angular momentum* (jm) *and of norm* N:

(i) *necessarily*

$$-j \leqslant m \leqslant j; \tag{XIII.12}$$

(ii) *if* $m=j, J_+|jm\rangle = 0$;
if $m \neq j, J_+|jm\rangle$ *is necessarily a vector of angular momentum* $(j, m+1)$ *and of norm* $[j(j+1)-m(m+1)]N$;

(iii) *if* $m=-j, J_-|jm\rangle = 0$;
if $m \neq -j, J_-|jm\rangle$ *is necessarily a vector of angular momentum* $(j, m-1)$ *and of norm* $[j(j+1)-m(m+1)]N$.

We now consider the vectors obtained by repeated application of the operator J_+ to $|jm\rangle$:

$$J_+|jm\rangle, J_+^2|jm\rangle, \ldots, J_+^p|jm\rangle, \ldots \tag{XIII.13}$$

We know that $-j \leqslant m \leqslant j$. If $m=j, J_+|jm\rangle = 0$. If $m<j$, $J_+|jm\rangle$ is a non-null vector of angular momentum $(j, m+1)$. It therefore has the properties (i)–(iii) characteristic of any common eigenvector of $\boldsymbol{J}^2$ and J_z: necessarily $m+1 \leqslant j$. If $m+1=j$, $J_+^2|jm\rangle = 0$. If $m+1<j$, $J_+^2|jm\rangle$ is a non-null vector of angular momentum $(j, m+2)$ and therefore it too has the properties (i)–(iii). One may thus, step-by-step, continue the analysis of the properties of the vectors (XIII.13). Clearly this sequence must terminate somewhere; otherwise we could form eigenvectors of J_z having eigenvalues larger than any given number, in contradiction with (XIII.12) according to which the eigenvalues of J_z cannot be larger than j. Therefore, there exists an integer $p(\geqslant 0)$ such that $J_+^p|jm\rangle$ is a non-null vector of angular momentum $(j, m+p)$ and upon which the action of J_+ gives 0; we therefore have $m+p=j$. Thus we have shown that $j-m$ is integral $(\geqslant 0)$ and that the p vectors

$$J_+|jm\rangle, J_+^2|jm\rangle, \ldots, J_+^p|jm\rangle \tag{XIII.14}$$

represent states of well-defined angular momentum, corresponding all to the same eigenvalue $j(j+1)$ of $\boldsymbol{J}^2$ and to the eigenvalues

$$m+1, m+2, \ldots, m+p=j,$$

of J_z respectively.

A similar examination of the vectors obtained by repeated application of J_- to $|jm\rangle$ shows that $j+m \equiv q$ is also an integer $\geqslant 0$, and that the q vectors

$$J_-|jm\rangle, J_-^2|jm\rangle, \ldots, J_-^q|jm\rangle \tag{XIII.15}$$

represent states of well-defined angular momentum all corresponding to the same eigenvalue $j(j+1)$ of $\boldsymbol{J}^2$, and to the eigenvalues

$$m-1, m-2, \ldots, m-q=-j,$$

of J_z respectively. Since p and q are both non-negative integers, their sum $p+q=2j$ is also a non-negative integer.

Bringing together these results, we have the following fundamental theorem:

(A) *The only possible eigenvalues of* $\boldsymbol{J}^2$ *are of the form* $j(j+1)$, *where* j *is a non-negative, integral or half-integral* [1]) *number*:

$$j=0, \tfrac{1}{2}, 1, \tfrac{3}{2}, 2, \ldots, \infty.$$

(B) *The only possible eigenvalues of* J_z *are the integral and half-integral numbers*:

$$m=0, \pm\tfrac{1}{2}, \pm 1, \pm\tfrac{3}{2}, \pm 2, \ldots, \pm\infty.$$

(C) *If* $j(j+1)$ *and* m *are the respective eigenvalues of* $\boldsymbol{J}^2$ *and* J_z *corresponding to a common eigenstate of these two operators* – that is, to a state of angular momentum (jm) – *the only possible values of* m *are the* $(2j+1)$ *quantities*

$$-j, -j+1, \ldots, +j.$$

5. Eigenvectors of $\boldsymbol{J}^2$ and J_z. Construction of the Invariant Subspaces $\mathscr{E}^{(j)}$

Starting from a vector $|j\,m\rangle$ of well-defined angular momentum one can construct in all $2j+1$ vectors of well-defined angular momentum by repeated application of the operators J_+ and J_-. In general these vectors are not normalized to unity but by proceeding in the following way one may easily construct eigenvectors which are.

---

[1]) Throughout this book "half-integral" will be understood to mean "half-odd-integral" i.e. the values $\frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \ldots$.

Suppose that the norm of $|j\,m\rangle$ is equal to one. $J_+|j\,m\rangle$ is zero if $m=j$; if $m<j$ it is a vector of angular momentum $(j, m+1)$. Denote by $|j\,m+1\rangle$ the vector of norm 1 defined by

$$J_+|j\,m\rangle = c_m|j\,m+1\rangle.$$

From the expression given above for the norm of $J_+|j\,m\rangle$ it follows that

$$|c_m|^2 = [j\,(j+1) - m\,(m+1)].$$

We fix the phase of $|j\,m+1\rangle$ so that c_m is both real and positive. Thus:

$$J_+|j\,m\rangle = \sqrt{j\,(j+1) - m\,(m+1)}\,|j\,m+1\rangle.$$

Multiplying both sides by J_- and using (XIII.10a), we have

$$J_-|j\,m+1\rangle = \sqrt{j\,(j+1) - m\,(m+1)}\,|j\,m\rangle.$$

The vector $|j\,m+1\rangle$ may be treated in the same way. It suffices to replace everywhere m by $m+1$. If $m+1=j$, $J_+|j\,m+1\rangle = 0$. If $m+1 \neq j$ we form the vector $|j\,m+2\rangle$, of angular momentum $(j, m+2)$ and of norm 1 with its phase fixed by a convention identical to the one above. And so forth until we obtain $|j\,j\rangle$.

In the same way, by repeated application of J_-, we successively form the vectors $|j\,m-1\rangle, \ldots, |j-j\rangle$ of norm 1 and of angular momentum $(j, m-1), \ldots, (j, -j)$ respectively; their phase is fixed by an identical convention.

Thus, starting from $|j\,m\rangle$, we form a series of $(2j+1)$ orthonormal vectors

$$|j\,j\rangle\,|j\,j-1\rangle \ldots |j\,m\rangle \ldots |j-j\rangle \qquad \text{(XIII.16)}$$

satisfying the eigenvalue equations

$$\mathbf{J}^2|j\,\mu\rangle = j\,(j+1)\,|j\,\mu\rangle \qquad \text{(XIII.17)}$$

$$J_z|j\,\mu\rangle = \mu\,|j\,\mu\rangle \qquad \text{(XIII.18)}$$

and whose relative phases have been chosen in such a way that they are obtained one from another by the relations

$$J_+|j\,\mu\rangle = \sqrt{j\,(j+1) - \mu\,(\mu+1)}\,|j\,\mu+1\rangle \qquad \text{(XIII.19)}$$

$$J_-|j\,\mu\rangle = \sqrt{j\,(j+1) - \mu\,(\mu-1)}\,|j\,\mu-1\rangle. \qquad \text{(XIII.20)}$$

In particular

$$J_+ |j\, j\rangle = J_- |j - j\rangle = 0. \tag{XIII.21}$$

These $(2j+1)$ vectors span a certain subspace $\mathscr{E}^{(j)}$. Since the operators J_+, J_-, J_z transform these vectors one into another, they transform any vector of $\mathscr{E}^{(j)}$ into a vector of $\mathscr{E}^{(j)}$; in other words, they leave $\mathscr{E}^{(j)}$ invariant. Any function $F(\boldsymbol{J})$ of the components of $\boldsymbol{J}$, being a function only of the operators J_+, J_-, J_z also leaves $\mathscr{E}^{(j)}$ invariant. In section III we shall see that an overall rotation of the quantum system corresponds to the application of an operator of the type $F(\boldsymbol{J})$ to the state vector; therefore, any rotation of the whole system leaves $\mathscr{E}^{(j)}$ invariant.

6. Standard Representation $\{\boldsymbol{J}^2\, J_z\}$

If $\boldsymbol{J}^2$ and J_z do not form a complete set of commuting observables, there will be many systems of basis vectors common to these two operators. Even when they do form a complete set, the phase of each basis vector may be arbitrarily fixed.

Among the representations with $\boldsymbol{J}^2$ and J_z diagonal there are certain in which manipulations of the angular momentum are particularly simple, and which are therefore preferable. They will be called *standard representations* $\{\boldsymbol{J}^2\, J_z\}$. They are those representations in which the basis vectors corresponding to a specified value of the quantum number j can be grouped in one or several series of $(2j+1)$ vectors connected by relations (XIII.19–20). To each series there corresponds a subspace $\mathscr{E}^{(j)}$ and the entire Hilbert space is the direct sum of such subspaces.

To set up a standard representation we may proceed in the following way. Among the eigenvectors of $\boldsymbol{J}^2$ belonging to the eigenvalue $j(j+1)$ we consider those for which $J_z = j$. They form a certain subspace $\mathscr{F}^{(j)}$ of Hilbert space which according to the case can have one, several or an infinite number of dimensions. It is always possible to choose a complete set of orthonormal vectors $|\tau\, j\, j\rangle$ in $\mathscr{F}^{(j)}$. The index τ distinguishes these vectors of angular momentum $(j\, j)$ one from another; it can, according to the case, take one, several or an infinite number of values (discrete or continuous; we shall suppose them discrete). We have by hypothesis

$$\langle \tau\, j\, j | \tau'\, j\, j\rangle = \delta_{\tau\tau'}.$$

With each of these vectors $|\tau\, j\, j\rangle$ we can associate the $2j$ vectors obtained by repeated application of J_- according to the method of the preceding paragraph, thus forming a $(2j+1)$-dimensional subspace $\mathscr{E}^{(j)}$; to distinguish these subspaces one from another we shall denote them $\mathscr{E}(\tau\, j)$. Thus the $(2j+1)$ basis vectors of the subspace $\mathscr{E}(\tau\, j)$ are

$$|\tau\, j\, j\rangle, |\tau\, j\, j-1\rangle, \ldots, |\tau\, j-j\rangle.$$

These vectors are orthonormal and satisfy the fundamental relations

$$\boldsymbol{J}^2 |\tau\, j\, \mu\rangle = j\,(j+1)\,|\tau\, j\, \mu\rangle \tag{XIII.22}$$

$$J_z |\tau\, j\, \mu\rangle = \mu\, |\tau\, j\, \mu\rangle \tag{XIII.23}$$

$$J_+ |\tau\, j\, \mu\rangle = \sqrt{j\,(j+1) - \mu\,(\mu+1)}\, |\tau\, j\, \mu+1\rangle \tag{XIII.24}$$

$$J_- |\tau\, j\, \mu\rangle = \sqrt{j\,(j+1) - \mu\,(\mu-1)}\, |\tau\, j\, \mu-1\rangle. \tag{XIII.25}$$

From (XIII.24) and (XIII.25) we easily deduce the following important relations (Problem XIII.1):

$$|\tau\, j\, \pm\mu\rangle = \sqrt{\frac{(j+\mu)!}{(2j)!\,(j-\mu)!}}\, J_{\mp}^{j-\mu} |\tau\, j\, \pm j\rangle \tag{XIII.26}$$

$$|\tau\, j\, \pm j\rangle = \sqrt{\frac{(j+\mu)!}{(2j)!\,(j-\mu)!}}\, J_{\pm 1}^{j-\mu} |\tau\, j\, \pm \mu\rangle. \tag{XIII.27}$$

It is easily shown that the subspaces $\mathscr{E}(\tau\, j)$ (given j, variable τ) are mutually orthogonal and that together they form the subspace $\mathscr{E}_j$ for the eigenvalue $j\,(j+1)$ of $\boldsymbol{J}^2$. The demonstration is as follows.

The basis vectors $|\tau\, j\, \mu\rangle$ and $|\tau'\, j\, \mu'\rangle$ of subspaces $\mathscr{E}(\tau\, j)$ and $\mathscr{E}(\tau'\, j)$ $(\tau \neq \tau')$ are certainly orthogonal if $\mu \neq \mu'$, as they correspond to different eigenvalues of J_z; the same is true if $\mu = \mu'$ since by repeated application of (XIII.24)

$$\langle \tau'\, j\, \mu\, |\tau\, j\, \mu\rangle = \langle \tau'\, j\, \mu+1\, |\tau\, j\, \mu+1\rangle = \ldots = \langle \tau'\, j\, j\, |\tau\, j\, j\rangle = \delta_{\tau\tau'}.$$

To show that any eigenvector of $\boldsymbol{J}^2$ corresponding to the eigenvalue $j\,(j+1)$ is a linear combination of the $|\tau\, j\, \mu\rangle$ (τ and μ variable, j fixed) it suffices to show that any vector $|\bar{\omega}\, j\, \mu\rangle$ of angular momentum $(j\, \mu)$ is a linear combination of the basis vectors $|\tau\, j\, \mu\rangle$ of the same angular momentum. If $\mu = j$ this is true by our initial hypothesis. If $\mu \neq j$ we can reduce the problem to this case by applying the operator $J_+^{j-\mu}$ to the vector $|\bar{\omega}\, j\, \mu\rangle$ and using relations (XIII.26) and (XIII.27).

Thus, starting from a complete set of orthonormal vectors of angular momentum $(j\,j)$, we have constructed the basis vectors of a standard representation $\{\mathbf{J}^2 J_z\}$ in the subspace $\mathscr{E}_j$ corresponding to the eigenvalue $j(j+1)$ of $\mathbf{J}^2$. Repeating the same operation for all possible eigenvalues of $\mathbf{J}^2$ we obtain a standard basis for the entire Hilbert space.

Note the particularly simple form of the matrices representing the components of $\mathbf{J}$ in such a representation (cf. Problem XIII.2). From (XIII.23), (XIII.24) and (XIII.25),

$$\begin{aligned}
&\langle \tau\, j\, \mu \,|J_z\,|\tau'\, j'\, \mu'\rangle = \mu\, \delta_{\tau\tau'}\, \delta_{jj'}\, \delta_{\mu\mu'} \\
&\langle \tau\, j\, \mu \,|J_\pm\,|\tau'\, j'\, \mu'\rangle = \sqrt{j\,(j+1)-\mu\mu'}\; \delta_{\tau\tau'}\, \delta_{jj'}\, \delta_{\mu\mu'\pm 1}.
\end{aligned} \tag{XIII.28}$$

7. Conclusion

The preceding study of the properties of angular momentum is exclusively based on the commutation relations of the latter and on the fact that its components are Hermitean operators in a Hilbert space. From this alone we have been able to show that the quantum number j can take only integral or half-integral values and that to each eigenvalue $j(j+1)$ of $\mathbf{J}^2$ there corresponds one or several series of $(2j+1)$ linearly independent vectors; the vectors of a series are obtained one from another by application of the operators J_- and J_+ and correspond to the $(2j+1)$ possible values of the quantum number m:

$$-j, -j+1, \ldots, +j.$$

However, these hypotheses are not sufficient for the complete solution of the eigenvalue problem. It remains to determine:

(i) Which of the integral and half-integral numbers actually make up the spectrum of j;

(ii) How many series of $(2j+1)$ linearly independent vectors correspond to each of these values of j.

The answer to these questions depends on the problem considered. From the commutation relations alone one cannot *a priori* exclude the case in which j takes a single value (integral or half-integral) and in which there is just one set of $(2j+1)$ linearly independent vectors corresponding to this eigenvalue; the state vector space is

then $(2j+1)$-dimensional. We have an example of this in section IV, in connection with spin.

Another important special case is that of the angular momentum $\boldsymbol{l}$ of a particle, as defined in equation (1). We shall see in section II, where this case is examined in detail, that the spectrum of j is then formed of all the integers from 0 to ∞, all half-integral values being excluded.[1])

II. ORBITAL ANGULAR MOMENTUM AND THE SPHERICAL HARMONICS

8. The Spectrum of l^2 and l_z

Let us return to the one-particle quantum system first considered in § 2. Choosing the z axis as the polar direction, we can express the operator $\boldsymbol{l}^2$ and the components of $\boldsymbol{l}$ as functions of the polar angles (θ, φ) and of their derivatives [eqs. (B.82–84)]. In what follows the radial variable may be ignored. We wish to find the function $F_l^m(\theta, \varphi)$ that satisfies the two eigenvalue equations

$$\boldsymbol{l}^2 F_l^m(\theta, \varphi) = l(l+1) F_l^m(\theta, \varphi) \qquad \text{(XIII.29)}$$

$$l_z F_l^m(\theta, \varphi) = m F_l^m(\theta, \varphi). \qquad \text{(XIII.30)}$$

Since any wave function is a single-valued function of $\boldsymbol{r}$, $F_l^m(\theta, \varphi)$ must remain unchanged when φ is replaced by $\varphi + 2\pi$ [2]). Equation (XIII.30) has already been studied in § V.12. Since $l_z = -\mathrm{i}\partial/\partial\varphi$, $F_l^m(\theta, \varphi)$ is necessarily of the form $f_l^m(\theta)\mathrm{e}^{\mathrm{i}m\varphi}$ with m integral. Since m is integral, l must also be integral: *there is no half-integral orbital angular momentum.*

In order to determine among the integers ($\geqslant 0$) the eigenvalues of $\boldsymbol{l}$ and their degeneracy, we form the eigenfunctions $F_{l'}^{l}(\theta, \varphi)$

---

[1]) Problem XIII.15 treats another case in which $\boldsymbol{J}^2$ and J_z form a complete set of commuting observables (a single series of $(2j+1)$ vectors for each value of j) and in which j can take all integral or half-integral values. Due to the simultaneous presence of integral and half-integral values in this spectrum, the treatment of $\boldsymbol{J}$ as an angular momentum imposes certain qualifications as to the physical significance of the different observables of the system (cf. § 15).

[2]) Also, $F_l^m(0, \varphi)$ and $F_l^m(\pi, \varphi)$ do not depend on φ, conditions automatically satisfied by the eigenfunctions found below.

corresponding to angular momentum (l, l). Such a function is defined by the equations

$$l_z \; F_l{}^l(\theta, \varphi) = l \; F_l{}^l(\theta, \varphi) \tag{XIII.31}$$

$$l_+ \; F_l{}^l(\theta, \varphi) = 0. \tag{XIII.32}$$

These equations correspond to the system (XIII.29–30) for the case $m = l$ since from (XIII.9a)

$$\boldsymbol{l}^2 = l_z(l_z + 1) + l_- l_+$$

and therefore (XIII.32) and (XIII.31) lead to (XIII.29) and vice versa. The system of first-order partial differential equations (XIII. 31–32) is easily solved once given the differential operators l_z and l_+ [eqs. (B.82–83)]. From (XIII.31),

$$F_l{}^l(\theta, \varphi) = f_l(\theta)\, e^{il\varphi}.$$

Substituting this expression in (XIII.32), we obtain the differential equation

$$\left(\frac{d}{d\theta} - l \cot\theta\right) f_l(\theta) = 0,$$

whose solution is $\sin^l\theta$ to within a constant. For each integral value of $l\,(\geqslant 0)$ there exists one, and only one, eigenfunction of angular momentum (l, l) (defined up to an arbitrary constant), namely

$$\sin^l\theta\, e^{il\varphi}.$$

Thus, the spectrum of $\boldsymbol{l}^2$ is the sequence of numbers $l(l+1)$, where *l takes all integral values from* 0 *to* $+\infty$. To each eigenvalue $l(l+1)$ there correspond $(2l+1)$ eigenvalues m of l_z – the $(2l+1)$ integers in the interval $(-l, +l)$. To each such pair (lm) there corresponds *one, and only one eigenstate* (if we limit ourselves to functions of the angles θ and φ): the spectrum of $\boldsymbol{l}^2$ and l_z is entirely non-degenerate.

9. Definition and Construction of the Spherical Harmonics

The common eigenfunction of $\boldsymbol{l}^2$ and l_z corresponding to the eigenvalues $(l\, m)$ is undefined to the extent of an arbitrary constant. We shall fix this constant by normalizing to unity and adopting a suitable phase convention. We thus obtain the spherical harmonic of order $(l\, m)$, $Y_l{}^m(\theta, \varphi)$.

The $Y_l^m(\theta, \varphi)$ form an *orthonormal set* of functions of θ and φ (the volume element in the scalar-product integral being $d\Omega = \sin\theta \, d\theta \, d\varphi$). It is also a *complete* set. The proof of completeness will not be given here.

The phases are fixed in the following way [1]). We first require that the Y_l^m form a standard basis. For this it is sufficient that they satisfy equations (XIII.24–25) written in the $\{\theta \varphi\}$ representation [eq. (B.89)]. The relative phases of the $(2l+1)$ spherical harmonics corresponding to the same value of l are thus determined and it remains but to fix the phase of one of them, $Y_l^0(\theta, \varphi)$ say. This we do by requiring that $Y_l^0(0, 0)$ be real and positive.

If we denote by $|l\, m\rangle$ the vectors represented by the spherical harmonic $Y_l^m(\theta, \varphi)$ in the $\{\theta \varphi\}$ representation, the different vectors thus defined satisfy equations (XIII.24–25); and therefore, [eq. (XIII.26)],

$$|l\, m\rangle = \sqrt{\frac{(l+m)!}{(2l)!\,(l-m)!}}\, l_-^{\,l-m}\, |l\, l\rangle$$

$$= \sqrt{\frac{(l-m)!}{(2l)!\,(l+m)!}}\, l_+^{\,l+m}\, |l-l\rangle.$$

In other words

$$Y_l^m(\theta, \varphi) = \sqrt{\frac{(l+m)!}{(2l)!\,(l-m)!}}\, l_-^{\,l-m}\, Y_l^l(\theta, \varphi) \qquad \text{(XIII.33)}$$

$$= \sqrt{\frac{(l-m)!}{(2l)!\,(l+m)!}}\, l_+^{\,l+m}\, Y_l^{-l}(\theta, \varphi). \qquad \text{(XIII.34)}$$

The differential operators l_- and l_+ are explicitly given in the appendix [eq. (B.83)]. From the expression given there, we deduce that

$$l_\pm\, e^{i\mu\varphi} f(\theta) = \mp\, e^{i(\mu\pm1)\varphi} \left(\sin^{1\pm\mu}\theta \frac{d}{d(\cos\theta)} \sin^{\mp\mu}\theta \right) f(\theta).$$

For any $f(\theta)$ the expression in brackets on the right-hand side must be considered as an operator acting on the function $f(\theta)$ to the right

---

[1]) This particularly convenient convention is the one adopted by most authors. However, for discussions involving invariance under time reversal, it is preferable to use the functions $\mathcal{Y}_l^m(\theta,\varphi) \equiv i^l\, Y_l^m(\theta, \varphi)$ when dealing with wave functions in configuration space (cf. § XV.22).

of it. Therefore, the repeated application of l_+ or l_- to the function $e^{i\mu\varphi} f(\theta)$ gives (μ and s integral)

$$l_\pm{}^s e^{i\mu\varphi} f(\theta) = (\mp)^s e^{i(\mu\pm s)\varphi} \left(\sin^{s\pm\mu}\theta \frac{d^s}{d(\cos\theta)^s} \sin^{\mp\mu}\theta \right) f(\theta). \quad \text{(XIII.35)}$$

From the discussion of § 8, we know that

$$Y_l{}^l(\theta, \varphi) = c_l \sin^l\theta\, e^{il\varphi},$$

where c_l is a constant whose modulus is determined by the normalization condition on $Y_l{}^l$, which gives

$$|c_l| = (4\pi)^{-\frac{1}{2}} \frac{\sqrt{(2l+1)!}}{2^l\, l!}, \quad \text{(XIII.36)}$$

and whose phase remains to be determined in agreement with the convention adopted above.

Substituting this expression for $Y_l{}^l$ in equation (XIII.33) we obtain, with the aid of identity (XIII.35)

$$Y_l{}^m(\theta, \varphi) = c_l \sqrt{\frac{(l+m)!}{(2l)!\,(l-m)!}}\, e^{im\varphi} \sin^{-m}\theta \frac{d^{l-m}}{d(\cos\theta)^{l-m}} \sin^{2l}\theta. \quad \text{(XIII.37)}$$

When $m = -l$, this formula gives

$$Y_l{}^{-l}(\theta, \varphi) = (-)^l c_l\, e^{-il\varphi} \sin^l\theta. \quad \text{(XIII.38)}$$

Substituting this expression for $Y_l{}^{-l}$ in equation (XIII.34) and again using (XIII.35), we obtain a new expression for $Y_l{}^m$ equivalent to the preceding one:

$$Y_l{}^m(\theta, \varphi) = (-)^m c_l \sqrt{\frac{(l-m)!}{(2l)!\,(l+m)!}}\, e^{im\varphi} \sin^m\theta \frac{d^{l+m}}{d(\cos\theta)^{l+m}} \sin^{2l}\theta. \quad \text{(XIII.39)}$$

When $m = 0$, these two expressions are identical

$$Y_l{}^0(\theta, \varphi) = c_l \sqrt{\frac{1}{(2l)!}} \frac{d^l}{d(\cos\theta)^l} (1-\cos^2\theta)^l.$$

To within a multiplicative constant, this is just the Legendre poly-

nomial $P_l(\cos\theta)$ [eq. (B.71)]:

$$Y_l^0(\theta,\varphi)=(-)^l\frac{c_l\,2^l\,l!}{\sqrt{(2l)!}}P_l(\cos\theta)$$

$$=(-)^l\frac{c_l}{|c_l|}\sqrt{\frac{2l+1}{4\pi}}P_l(\cos\theta).$$

The condition that $Y_l^0(0,0)$ be real and positive imposes the choice of phase

$$\frac{c_l}{|c_l|}=(-)^l.$$

Expressions (XIII.37) and (XIII.38) for $Y_l^m(\theta,\varphi)$ are thus completely specified.

Most of the properties of spherical harmonics given in Appendix B can be read directly from the condensed forms (XIII.37) and (XIII.38). Note in particular that

$$Y_l^{-m}(\theta,\varphi)=(-)^m\,Y_l^{m*}(\theta,\varphi),$$

that Y_l^m is the product of $e^{im\varphi}\sin^{|m|}\theta$ by a polynomial of degree $l-|m|$ and of parity $(-)^{l-|m|}$ in $\cos\theta$, and that the *parity* of Y_l^m (Problem XIII. 4) is $(-)^l$, i.e.

$$Y_l^m(\pi-\theta,\varphi+\pi)=(-)^l\,Y_l^m(\theta,\varphi).$$

III. ANGULAR MOMENTUM AND ROTATIONS

10. Definition of Rotation. Euler Angles

In this paragraph we review certain properties of rotations in ordinary space.

By definition, a rotation about a given point O is an overall displacement of the points of space in which the point O remains fixed. In such a displacement, each point P takes up a new position P′ and there is a one-to-one correspondence between P and P′. One could also define a rotation about O as a one-to-one correspondence between the points of space in which the point O corresponds to

itself and which conserves both distances (and therefore angles) and the sense of coordinate axes [1]).

A unit vector $\boldsymbol{u}$ and an angle φ define a particular rotation $\mathscr{R}_{\boldsymbol{u}}(\varphi)$, of angle φ about the oriented axis defined by $\boldsymbol{u}$ (the positive sense of rotation being defined in the usual way). There are an infinite number of ways of thus specifying the same rotation. $\mathscr{R}_{\boldsymbol{u}}(\varphi)$ is equal to $\mathscr{R}_{\boldsymbol{u}_1}(\varphi)$ if

$$\begin{matrix} \boldsymbol{u}_1 = \boldsymbol{u} \\ \varphi_1 = \varphi + 2n\pi \end{matrix} \quad \text{or} \quad \begin{matrix} \boldsymbol{u}_1 = -\boldsymbol{u} \\ \varphi_1 = -\varphi + 2n\pi \end{matrix} \qquad (n \text{ any integer}).$$

We say that the rotation is *infinitesimal* if $\varphi = \varepsilon$ is an infinitesimal. It is easy to write down the transform $\boldsymbol{V}'$, of a vector $\boldsymbol{V}$ in an infinitesimal rotation $\mathscr{R}_{\boldsymbol{u}}(\varepsilon)$:

$$\boldsymbol{V}' \simeq \boldsymbol{V} + \varepsilon(\boldsymbol{u} \times \boldsymbol{V}) \qquad (\varepsilon \ll 1). \tag{XIII.40}$$

Another method of specifying a rotation consists in giving its Euler angles α, β, γ. Let $Oxyz$ be the system obtained by the rotation of the Cartesian system $OXYZ$, and Ou one of the two *directed* axes perpendicular to OzZ (Fig. XIII.1). The Euler angles are [2])

$$\alpha = (Oy, Ou), \qquad \beta = (Oz, OZ), \qquad \gamma = (Ou, OY).$$

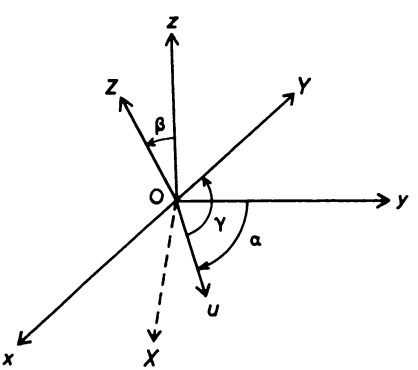


Fig. XIII.1. Definition of the Euler angles.

[1]) In contrast with reflections which conserve distances but change the sense of coordinate axes.

[2]) The definition adopted here differs slightly from the one generally adopted in the theory of the gyroscope.

The rotation is the result of the following sequence of three rotations:

(i) a rotation of angle α about Oz, $\mathscr{R}_Z(\alpha)$ (Oy goes into Ou)

(ii) a rotation of angle β about Ou, $\mathscr{R}_u(\beta)$ (Oz goes into OZ)

(iii) a rotation of angle γ about OZ, $\mathscr{R}_Z(\gamma)$ (Ou goes into OY)

We denote the resulting rotation by $\mathscr{R}(\alpha\,\beta\,\gamma)$ and write

$$\mathscr{R}(\alpha\,\beta\,\gamma)=\mathscr{R}_Z(\gamma)\,\mathscr{R}_u(\beta)\,\mathscr{R}_z(\alpha). \tag{XIII.41}$$

The angles α, β, γ are algebraic quantities with a positive or negative sign according as the corresponding rotations about the axes Oz, Ou, OZ are positive or negative. The system $Oxyz$ being chosen once and for all, the same rotation can be defined by several different sets of Euler angles. A necessary and sufficient condition for

$$\mathscr{R}(\alpha\,\beta\,\gamma)=\mathscr{R}(\alpha_1\,\beta_1\,\gamma_1)$$

is that:

$$\begin{aligned}
&\alpha_1=\alpha+2\pi n_\alpha && \alpha_1=\alpha+\pi+2\pi n_\alpha\\
&\beta_1=\beta+2\pi n_\beta \quad \text{or} && \beta_1=-\beta+2\pi n_\beta\\
&\gamma_1=\gamma+2\pi n_\gamma && \gamma_1=\gamma-\pi+2\pi n_\gamma
\end{aligned} \tag{XIII.42}$$

(n_α, n_β, n_γ any integers).

With each rotation $\mathscr{R}$ can be associated a certain 3×3 matrix defined in the following way. We choose, once and for all, a right-handed Cartesian system $Oxyz$ with unit vectors $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$ in the directions Ox, Oy, Oz respectively. Under rotation these transform respectively into 3 new vectors, $\boldsymbol{A}_1$, $\boldsymbol{A}_2$, $\boldsymbol{A}_3$ forming a new right-handed Cartesian system $OXYZ$. Each of the vectors $\boldsymbol{A}_j$ is a linear combination of the vectors $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$ [1]):

$$\boldsymbol{A}_j\equiv\mathscr{R}[\boldsymbol{a}_j]=\boldsymbol{a}_i\,\mathscr{R}_{ij},\qquad \mathscr{R}_{ij}=(\boldsymbol{a}_i\cdot\boldsymbol{A}_j).$$

The coefficients $\mathscr{R}_{ij}$ of these three linear combinations are the elements of a 3×3 matrix which we denote by the same letter $\mathscr{R}$ as the rotation itself. The rotation is completely defined by this matrix, i.e. if $\boldsymbol{V}\equiv\boldsymbol{a}_jV_j$ is any vector in space defined by its coordinates (V_1, V_2, V_3) along

---

[1]) In this entire section we shall systematically adopt the convention of summing over repeated indices; thus:

$$\boldsymbol{a}_i\,\mathscr{R}_{ij}=\boldsymbol{a}_1\,\mathscr{R}_{1j}+\boldsymbol{a}_2\,\mathscr{R}_{2j}+\boldsymbol{a}_3\,\mathscr{R}_{3j}.$$

the axis $Oxyz$, it is transformed by the rotation into the vector

$$\mathbf{V}' \equiv \mathscr{R}[\mathbf{V}] = \mathbf{A}_j \, V_j = \mathbf{a}_i \, \mathscr{R}_{ij} \, V_j.$$

The components of $\mathbf{V}'$ along the axes $Oxyz$ are

$$V'_i = \mathscr{R}_{ij} \, V_j. \tag{XIII.43}$$

They are obtained from those of $\mathbf{V}$ by application of the matrix $\mathscr{R}$.

Since the $\mathbf{A}_i$ form a Cartesian system, the real matrix $\mathscr{R}$ is orthogonal and unimodular:

$$\mathscr{R} = \mathscr{R}^*, \qquad \tilde{\mathscr{R}} = \mathscr{R}^{-1}, \qquad \det \mathscr{R} = 1.$$

The system of axes $Oxyz$ having been chosen once and for all, the matrix associated with a rotation is uniquely defined. Conversely, to each real, orthogonal, unimodular matrix there corresponds one, and only one rotation.

The elements of the matrix associated with the rotation $\mathscr{R}(\alpha \, \beta \, \gamma)$ are given explicitly as functions of its Euler angles in Appendix C [formula (C.45)]. As an example we give the transformation law for the coordinates of the above-mentioned vector $\mathbf{V}$ in a rotation $\mathscr{R}_z(\alpha)$ of angle α about Oz:

$$\begin{aligned} V'_1 &= V_1 \cos \alpha - V_2 \sin \alpha \\ V'_2 &= V_1 \sin \alpha + V_2 \cos \alpha \\ V'_3 &= V_3. \end{aligned} \tag{XIII.44}$$

The product of two rotations $\mathscr{R}_1$ and $\mathscr{R}_2$, namely the transformation $\mathscr{R} \equiv \mathscr{R}_2 \mathscr{R}_1$ obtained by successively effecting $\mathscr{R}_1$ and then $\mathscr{R}_2$, is also a rotation. Relation (XIII.41) is an example of such a product. It is not easy to write the Euler angles of $\mathscr{R}$ as a function of those of $\mathscr{R}_1$ and $\mathscr{R}_2$ but its associated matrix is easily obtained being the product of the matrices associated with $\mathscr{R}_1$ and $\mathscr{R}_2$:

$$\mathscr{R} = \mathscr{R}_2 \mathscr{R}_1.$$

11. Rotation of a Physical System. Rotation Operator

In discussing a given rotation in relation to a physical problem — what follows is also true for any spatial transformation — one can adopt two points of view which must be clearly distinguished. The

first (sometimes called passive) consists in rotating the reference axis, keeping fixed each point P of space and the physical quantities attached to it. The second (sometimes called active) consists in keeping the axes fixed and rotating the physical system itself. The two points of view are equivalent. That we rotate the coordinate axes, or that we rotate the physical system itself in the opposite direction, amounts to exactly the same thing. Unless otherwise specified, in what follows we shall adopt the second of these viewpoints (rotation of the physical system).

In Quantum Mechanics, more care is required in defining the "rotation of a physical system" than in Classical Mechanics due to the relation between dynamical variables and dynamical states being much less direct. For simplicity, we first consider the case of a single particle. Denote by a a possible dynamical state of the particle and by $\psi(\boldsymbol{r})$ the corresponding wave function. Denote by a' the state obtained by subjecting the particle to a certain rotation $\mathscr{R}$, and by $\psi'(\boldsymbol{r})$ the wave-function corresponding to a':

$$a' \equiv \mathscr{R}[a], \qquad \psi'(\boldsymbol{r}) \equiv \mathscr{R}[\psi(\boldsymbol{r})].$$

When we say that the state a goes over into the state a' in the rotation $\mathscr{R}$, we mean that whatever observations we make upon the system in state a', the results of these observations can be deduced by a rotation $\mathscr{R}$ from the results that the same observation would yield if made upon the system in state a. Consider, for example, a position measurement. The probability distributions for states a and a' are $|\psi(\boldsymbol{r})|^2$ and $|\psi'(\boldsymbol{r})|^2$ respectively. The above statement asserts that the latter is obtained from the former by a rotation $\mathscr{R}$, i.e. *the value of the second function at a given point* $\boldsymbol{r}$ *is equal to the value of the first at the point* $\boldsymbol{r}_1$, *which transforms into* $\boldsymbol{r}$ *in the rotation* $\mathscr{R}$;

$$|\psi'(\boldsymbol{r})|^2 = |\psi(\boldsymbol{r}_1)|^2, \qquad \boldsymbol{r}_1 = \mathscr{R}^{-1}\boldsymbol{r}. \tag{XIII.45}$$

Similarly, if $\varphi(\boldsymbol{p})$ and $\varphi'(\boldsymbol{p})$ are the momentum space wave functions corresponding respectively to ψ and ψ', we must have

$$|\varphi'(\boldsymbol{p})|^2 = |\varphi(\boldsymbol{p}_1)|^2, \qquad \boldsymbol{p}_1 = \mathscr{R}^{-1}\boldsymbol{p}. \tag{XIII.46}$$

In order to satisfy all these conditions it is clearly sufficient that the

value of the function ψ' at $\boldsymbol{r}$ equal the value of the function ψ at $\boldsymbol{r}_1$, i.e.

$$\psi'(\boldsymbol{r}) \equiv \mathscr{R}[\psi(\boldsymbol{r})] = \psi(\mathscr{R}^{-1}\boldsymbol{r}). \tag{XIII.47}$$

It can be shown that this is also a necessary condition [1]; the wave function is therefore unambiguously defined.

Relation (XIII.47) sets up a one-to-one correspondence between ψ and ψ'. It is clear that this correspondence is linear. In other words there exists an operator R such that

$$\psi' = R\,\psi.$$

R is unitary since the norms of ψ and ψ' are equal:

$$\int |\psi'(\boldsymbol{r})|^2 \mathrm{d}\boldsymbol{r} = \int |\psi(\mathscr{R}^{-1}\boldsymbol{r})|^2 \mathrm{d}\boldsymbol{r} = \int |\psi(\boldsymbol{r}_1)|^2 \mathrm{d}\boldsymbol{r}_1$$

(to obtain the last integral we have made the change of variable $\boldsymbol{r}_1 = \mathscr{R}^{-1}\boldsymbol{r}$ and used the fact that the volume element $\mathrm{d}\boldsymbol{r}$ is conserved in the rotation $\mathscr{R}^{-1}$).

All of this may easily be generalized to a system of N particles, the wave function $\psi(\boldsymbol{r}^{(1)}, \boldsymbol{r}^{(2)}, \ldots, \boldsymbol{r}^{(N)})$ transforming in the rotation $\mathscr{R}$ into

$$\mathscr{R}[\psi(\boldsymbol{r}^{(1)}, \ldots, \boldsymbol{r}^{(N)})] = \psi(\mathscr{R}^{-1}\boldsymbol{r}^{(1)}, \ldots, \mathscr{R}^{-1}\boldsymbol{r}^{(N)}) = R\,\psi(\boldsymbol{r}^{(1)}, \ldots, \boldsymbol{r}^{(N)}). \tag{XIII.48}$$

The rotation operator R is, as above, linear and unitary.

In all generality, a unitary operator R is associated with each rotation $\mathscr{R}$ of a given physical system; the application of R to the vector $|a\rangle$ representing the dynamical state of the system before rotation gives the vector $|a'\rangle$ representing its dynamical state after rotation:

$$R\,R^\dagger = R^\dagger\,R = 1 \tag{XIII.49}$$

$$|a'\rangle = R\,|a\rangle. \tag{XIII.50}$$

From (XIII.50) we can easily deduce the law for the transformation of the density operator by referring directly to its definition. Let ϱ be the density operator representing a certain (pure or mixed)

---

[1] The general proof will be given in § XV.6. The function ψ' is actually defined by these conditions only to within a phase factor. This factor is fixed if we require that the operators R defined below form a group isomorphic with the rotations; this is what is done here.

state of the system, ϱ' the one representing the state resulting from the rotation $\mathscr{R}$. We have

$$\varrho' \equiv \mathscr{R}[\varrho] = R\,\varrho\,R^{\dagger}. \tag{XIII.51}$$

12. Rotation of Observables

Besides rotating the system itself, we can also rotate the instruments with which we observe it. Having defined the law for the transformation of state vectors, we must now define the law for the transformation of the observables representing the various measuring operations that can be made on the system.

Let Q be an observable and $Q' = \mathscr{R}[Q]$ its transform in the rotation $\mathscr{R}$. Physically, the observable Q represents a measuring operation and the transformation of Q into Q' represents an overall rotation of the measuring instrument. Therefore, the average value of measurements of Q made on the system in state $|a\rangle$ is equal to the average value of measurements of Q' made on the system in state $|a'\rangle \equiv \mathscr{R}[|a\rangle]$, i.e.,

$$\langle a|Q|a\rangle = \langle a'|Q'|a'\rangle.$$

Since $|a'\rangle = R|a\rangle$, this may be written

$$\langle a|Q|a\rangle = \langle a|R^{\dagger}\,Q'\,R|a\rangle.$$

Since this must hold for every $|a\rangle$, we have (cf. § VII.5)

$$Q = R^{\dagger}\,Q'\,R,$$

i.e.,

$$Q' = R\,Q\,R^{\dagger}. \tag{XIII.52}$$

Thus, in a rotation $\mathscr{R}$, the observables undergo the same unitary transformation as the state vectors.

In particular, if an observable S represents a *scalar* quantity – that is, invariant under rotation [1] – then for any R,

$$S' \equiv R\,S\,R^{\dagger} = S.$$

---

[1] This definition of a scalar will be adopted throughout the present chapter. Later on quantities invariant under rotation will be classed as scalars and pseudo-scalars. The first are unchanged in a reflection, the second are multiplied by -1.

Since R is unitary, this may be written

$$[R, S]=0. \tag{XIII.53}$$

Thus an observable invariant under rotation commutes with all of the rotation operators.

Vector operators provide another particularly interesting case. We follow the notation of § 10 and denote the vector operator with components $K_i=(\mathbf{K}\cdot\boldsymbol{a}_i)$ by $\mathbf{K}$. If we apply the rotation $\mathscr{R}$ to the operator K_1, the component of $\mathbf{K}$ along Ox, the operator obtained, K_1', is the component of $\mathbf{K}$ along OX. In general $\mathscr{R}\,[\mathbf{K}\cdot\boldsymbol{a}]=\mathbf{K}\cdot\boldsymbol{a}'$, where $\boldsymbol{a}'=\mathscr{R}\,[\boldsymbol{a}]$; thus

$$K_i' \equiv \mathscr{R}\,[K_i]=\mathbf{K}\cdot\mathbf{A}_i=K_j\,\mathscr{R}_{ji}.$$

Hence the transformation law for the Cartesian components of $\mathbf{K}$ is

$$K_i' \equiv R\,K_i\,R^\dagger=\tilde{\mathscr{R}}_{ij}\,K_j. \tag{XIII.54}$$

Note that $\tilde{\mathscr{R}}$, the inverse of $\mathscr{R}$, appears here rather than $\mathscr{R}$ itself as in (XIII.43): the components of $\mathbf{K}$ transform in the rotation $\mathscr{R}$ like those of a vector in the rotation $\mathscr{R}^{-1}$.

13. Angular Momentum and Infinitesimal Rotations

We are now in a position to establish the fundamental relation between the angular momentum of a system and its infinitesimal rotation operators.

We first consider the case of the single particle of § 11. According to transformation law (XIII.47), the rotation $\mathscr{R}_z\,(\alpha)$ of angle α about Oz, transforms the function $\psi\,(x, y, z)$ [cf. eq. (XIII.44)] into

$$\mathscr{R}_z\,(\alpha)\,[\psi\,(x, y, z)]=\psi\,(x\cos\alpha+y\sin\alpha, -x\sin\alpha+y\cos\alpha, z).$$

In particular the infinitesimal rotation $\mathscr{R}_z\,(\varepsilon)$ gives, taking only terms of the first order in ε in the Taylor expansion of the right-hand side about the point (x, y, z),

$$\begin{aligned} R_z\,(\varepsilon)\,[\psi\,(x, y, z)] &\simeq \psi\,(x+y\varepsilon, -x\varepsilon+y, z) \\ &\simeq \psi\,(x, y, z)+\varepsilon\left(y\frac{\partial\psi}{\partial x}-x\frac{\partial\psi}{\partial y}\right) \\ &\simeq (1-\mathrm{i}\varepsilon l_z)\,\psi\,(x, y, z). \end{aligned}$$

The last line follows from the definition of the differential operator l_z ($\hbar = 1$). The infinitesimal rotation operator is therefore of the form:

$$R_z(\varepsilon) \simeq 1 - i\varepsilon l_z.$$

The same argument applied to the infinitesimal rotation about $\boldsymbol{u}$ gives

$$R_{\boldsymbol{u}}(\varepsilon) = 1 - i\varepsilon (\boldsymbol{l} \cdot \boldsymbol{u}).$$

For a system of N particles we have a similar result. Starting from (XIII.48) rather than (XIII.47) as above, an analogous treatment gives

$$\mathscr{R}_z(\varepsilon) \simeq 1 - i\varepsilon L_z,$$

and more generally

$$R_{\boldsymbol{u}}(\varepsilon) \simeq 1 - i\varepsilon (\boldsymbol{L} \cdot \boldsymbol{u}),$$

where $\boldsymbol{L}$ is the total angular momentum of the system.

In summary:

If $\boldsymbol{J}$ is the total angular momentum of a system, its component along any axis $\boldsymbol{u}$ is related to the operator of infinitesimal rotation about that axis by the relation

$$\boxed{R_{\boldsymbol{u}}(\varepsilon) \simeq 1 - i\varepsilon (\boldsymbol{J} \cdot \boldsymbol{u})} \qquad \text{(XIII.55)}$$

When the system has no classical analogue this fundamental relation serves as the definition of total angular momentum.

For this definition to be coherent we must be sure that the operator $(\boldsymbol{J} \cdot \boldsymbol{u})$ is the component along $\boldsymbol{u}$ of a certain vector operator $\boldsymbol{J}$. It is sufficient [1]), for this, that to each *infinitesimal* rotation $\mathscr{R}_{\boldsymbol{u}}(\varepsilon)$ there corresponds *one, and only one* infinitesimal-rotation operator $R_{\boldsymbol{u}}(\varepsilon)$. By the transformation law for vectors (XIII.40), the operation $\mathscr{R}_{\boldsymbol{u}}(\varepsilon)$ is equivalent, to first order in ε, to the product of operations $\mathscr{R}_x(\varepsilon u_x)\, \mathscr{R}_y(\varepsilon u_y)\, \mathscr{R}_z(\varepsilon u_z)$, so that

$$\begin{aligned} R_{\boldsymbol{u}}(\varepsilon) &\simeq R_x(\varepsilon u_x)\, R_y(\varepsilon u_y)\, R_z(\varepsilon u_z) \\ &\simeq 1 - i\varepsilon (u_x J_x + u_y J_y + u_z J_z). \end{aligned}$$

It follows from this definition that any scalar operator S commutes

---

[1]) This amounts to supposing that the rotation operators form a group.

with the components of $\boldsymbol{J}$ [eq. (XIII.53)]:

$$[(\boldsymbol{u}\cdot\boldsymbol{J}), S]=0. \tag{XIII.56}$$

Relation (XIII.55) also provides the commutation rules for the components of $\boldsymbol{J}$ with those of a general vector operator $\boldsymbol{K}$. Let $K_a \equiv \boldsymbol{K}\cdot\boldsymbol{a}$ be the component of $\boldsymbol{K}$ along a given unit vector $\boldsymbol{a}$. By definition, its transform in the rotation $\mathscr{R}_u(\varepsilon)$ is

$$K_a' \equiv R_u(\varepsilon)\, K_a\, R_u^{\dagger}(\varepsilon) \simeq K_a - \mathrm{i}\varepsilon\,[J_u, K_a].$$

However, by the law for the transformation of the vector $\boldsymbol{a}$ [eq. (XIII.40)],

$$K_a' = \boldsymbol{K}\cdot\boldsymbol{a}' \simeq \boldsymbol{K}\cdot[\boldsymbol{a}+\varepsilon\,(\boldsymbol{u}\times\boldsymbol{a})].$$

Equating the terms of first order in ε in these two expressions, we find

$$[J_u, K_a] = \mathrm{i}\boldsymbol{K}\cdot(\boldsymbol{u}\times\boldsymbol{a}),$$

i.e.,

$$\boxed{[(\boldsymbol{u}\cdot\boldsymbol{J}), (\boldsymbol{a}\cdot\boldsymbol{K})] = \mathrm{i}((\boldsymbol{u}\times\boldsymbol{a})\cdot\boldsymbol{K})} \tag{XIII.57}$$

Substitution of the operator $\boldsymbol{J}$ for $\boldsymbol{K}$ gives us back the commutation relations characteristic of angular momentum [eq. (XIII.4)].

The following definition of the *total* angular momentum is equivalent to the one given above:

If the fundamental observables of a system are the scalar operators $S_1, S_2, \ldots$ and the components of the vector operators $\boldsymbol{K}_1, \boldsymbol{K}_2, \ldots$, the *total* angular momentum of the system is by definition a *vector operator* $\boldsymbol{J}$ whose components commute with all of the S and satisfy the commutation relations (XIII.57) with the components of the $\boldsymbol{K}$.

If relations (XIII.57) are not satisfied by all of the vectors $\boldsymbol{K}_1$, $\boldsymbol{K}_2, \ldots$, $\boldsymbol{J}$ is not the total angular momentum of the system, even if it satisfies the commutation rules (XIII.4) characteristic of an angular momentum operator. Thus, in the N particle case considered in § 11, any vector operator formed by summing a certain number of individual angular momenta $\boldsymbol{l}^{(i)}$ will verify relations (XIII.4), but only the sum $\boldsymbol{L}$ of *all* the $\boldsymbol{l}^{(i)}$ corresponds to the definition of total angular momentum.

14. Construction of the Operator $R\ (\alpha \beta\ \gamma)$

Any finite rotation can be looked upon as a succession of infinitesimal rotations. The corresponding rotation operator is the product of the corresponding infinitesimal rotation operators. Since these are well-defined functions of the total angular momentum [eq. (XIII.55)], any finite rotation operator can also be expressed as a function of the total angular momentum.

The rotation $\mathscr{R}_u\ (\varphi)$ is a succession of infinitesimal rotations about the u axis. In particular,

$$\mathscr{R}_u\ (\varphi + \mathrm{d}\varphi) = \mathscr{R}_u\ (\mathrm{d}\varphi)\ \mathscr{R}_u\ (\varphi).$$

Putting $J_u \equiv (J \cdot u)$ and applying formula (XIII.55), this gives

$$\begin{aligned} R_u\ (\varphi + \mathrm{d}\varphi) &= R_u\ (\mathrm{d}\varphi)\ R_u\ (\varphi) \\ &= (1 - \mathrm{i}J_u\ \mathrm{d}\varphi)\ R_u\ (\varphi), \end{aligned}$$

or again

$$\frac{\mathrm{d}}{\mathrm{d}\varphi} R_u\ (\varphi) = -\mathrm{i}J_u\ R_u\ (\varphi) \qquad (R_u\ (0) = 1).$$

This differential equation is easily integrated to give

$$R_u\ (\varphi) = \mathrm{e}^{-\mathrm{i}\varphi J_u}. \tag{XIII.58}$$

Consider now the rotation $\mathscr{R}\ (\alpha\ \beta\ \gamma)$ defined by its Euler angles (α, β, γ). It was seen in § 10 that it may be looked upon as a succession of rotations of angles α, β, γ about axes Oz, Ou, OZ respectively (Fig. XIII.1). We therefore have

$$R\ (\alpha\ \beta\ \gamma) = R_Z\ (\gamma)\ R_u\ (\beta)\ R_z\ (\alpha).$$

With the aid of (XIII.58), the three rotations on the right-hand side can be expressed in terms of the angular momentum components J_Z, J_u, and J_z:

$$R\ (\alpha\ \beta\ \gamma) = \mathrm{e}^{-\mathrm{i}\gamma J_Z}\ \mathrm{e}^{-\mathrm{i}\beta J_u}\ \mathrm{e}^{-\mathrm{i}\alpha J_z}. \tag{XIII.59}$$

Note the order of the three exponentials on the right-hand side.

We shall put this expression in a form where only the components of angular momentum along the coordinate axes appear. The rotation $\mathscr{R}_z\ (\alpha)$ takes the operator J_y over into the operator J_u and so, by

the law for the transformation of operators (XIII.52),

$$J_u = R_z(\alpha)\, J_y\, R_z^{\dagger}(\alpha) = e^{-i\alpha J_z} J_y\, e^{+i\alpha J_z}.$$

Thus

$$e^{-i\beta J_u} = e^{-i\alpha J_z}\, e^{-i\beta J_y}\, e^{+i\alpha J_z}.$$

Substituting this expression into the right-hand side of (XIII.59), we get

$$R(\alpha\,\beta\,\gamma) = e^{-i\gamma J_z}\, e^{-i\alpha J_z}\, e^{-i\beta J_y}.$$

Similarly, J_Z is obtained from J_z by successive application of the rotations $\mathscr{R}_z(\alpha)$ and $\mathscr{R}_u(\beta)$ and can thus be eliminated like J_u above to give finally:

$$\boxed{R(\alpha\,\beta\,\gamma) = e^{-i\alpha J_z}\, e^{-i\beta J_y}\, e^{-i\gamma J_z}} \qquad \text{(XIII.60)}$$

15. Rotation through an Angle 2π and Half-integral Angular Momenta

From equation (XIII.58)

$$R_u(2\pi) = e^{-2\pi i J_u}.$$

Although *a rotation through* 2π *about an axis* $\boldsymbol{u}$ brings us back to our starting point, the corresponding rotation operator is not necessarily equal to 1. It is diagonal in the representation in which J_u is diagonal and its diagonal elements are $+1$ or -1 according as the corresponding eigenvalue of J_u is integral or half-integral.

Call D the function of J^2, with eigenvalue $+1$ for j integral and -1 for j half-integral. D is an observable, and has the following characteristic properties:

(i) $\frac{1}{2}(1+D)$ is the projector onto the subspace of integral j;

(ii) $\frac{1}{2}(1-D)$ is the projector onto the subspace of half-integral j;

(iii) $D^2 = 1$;

(iv) D commutes with all of the rotation operators:

$$[D, R] = 0.$$

Clearly

$$R_u(2\pi) = D. \qquad \text{(XIII.61)}$$

In order to have $R_u(2\pi) = 1$, it is necessary that the angular momen-

tum take only integral values. On the other hand, we shall always have:

$$R_u(4\pi) = D^2 = 1.$$

The existence of a one-to-one correspondence between the infinitesimal rotations and the infinitesimal R operators [definition (XIII.55)] in no way implies that there is a similar correspondence for finite rotations. There are infinitely many ways of writing a finite rotation as a product of infinitesimal rotations, the operators $R_u(\varphi)$ and $R(\alpha\,\beta\,\gamma)$ discussed above correspond each to one of these ways. There is no *a priori* reason to suppose that another way would give the same operator.

It can be shown (the proof will not be given here) that to each finite rotation $\mathscr{R}$ there corresponds in all two operators, R' and R'', differing by a "rotation through 2π",

$$R'' = DR'. \tag{XIII.62}$$

In the physical systems hitherto encountered, the total angular momentum could take only integral values, in which case $D = 1$, $R' = R''$, so that to each rotation $\mathscr{R}$ there corresponds one, and only one operator for the rotation of ket vectors. If the system has half-integral angular momentum states, the operators R' and R'' are not identical.

The occurrence of two different operators to describe the same rotation requires some discussion. That the rotation of a ket through 2π does not give the same ket raises no difficulty of principle so long as no observable effect is produced. Clearly the results of an experiment are not modified if beforehand one rotates some of the instruments of observation through an angle 2π; two identical counters occupying the same position will necessarily give the same answers. Therefore, if an observable Q represents a measureable quantity, it must be invariant under a rotation through 2π; more generally, if we effect a certain rotation $\mathscr{R}$ upon Q, the observable obtained must not depend on the particular path followed in so doing:

$$R'\,Q\,R'^{\dagger} = R''\,Q\,R''^{\dagger}.$$

The invariance under "rotation through 2π" is sufficient to guarantee

this more general property. Formally it may be written

$$[D, Q] = 0. \tag{XIII.63}$$

By definition, an observable is a Hermitean operator having a complete set of eigenvectors. Any operator representing a physical quantity must be an observable – a necessary condition for the self-consistence of Quantum Mechanics. However, the converse is by no means necessarily true. We will give the name of *physical observable* to an observable associated with a physically measurable quantity. The foregoing analysis shows that any physical observable must obey relation (XIII.63) [1]). In studying a physical system, one usually implicitly assumes that all the observables of the system are physical observables; although this hypothesis often facilitates discussion, it is not essential. More restrictive conditions may hold without any serious modification of the interpretation of the theory. Relation (XIII.63) is just one of these restrictions; we shall encounter others in the discussion of identical particles [2]).

In view of the foregoing discussion, no principle of Quantum Mechanics opposes the existence of half-integral angular momenta. Indeed, they are observed in nature.

16. Irreducible Invariant Subspaces. Rotation Matrices $R^{(J)}$

Expression (XIII.60) shows that any rotation operator is a function of the components of total angular momentum, as stated at the end of § 5. The vectors of a space $\mathscr{E}^{(J)}$ of the type constructed in that paragraph are therefore carried over in a rotation into vectors of $\mathscr{E}^{(J)}$ i.e., *the space* $\mathscr{E}^{(J)}$ *is invariant under rotation* [3]).

Better, if $|u\rangle$ is an arbitrarily chosen vector of that space, then the set of vectors $R|u\rangle$ obtained from $|u\rangle$ by rotation span the whole of $\mathscr{E}^{(J)}$. We say of a space having this property that it is *irreducible*

---

[1]) The observables of all the physical systems studied in this book (including angular momenta with half-integral eigenvalues) will all satisfy XIII.63. The distinction between observable and physical observable is then purely academic. However, one can imagine systems for which the observables do not all satisfy (XIII.53); Problem XIII.15 provides an example.

[2]) For a general discussion of relations of the type (XIII.63) and their consequences (superselection rules) cf. Wick, Wightman and Wigner, Phys. Rev. **88** (1952) 101.

[3]) We shall henceforth use the capitals J, M to denote the quantum numbers of total angular momentum.

with respect to rotations. If, on the other hand, there existed in $\mathscr{E}^{(J)}$ at least one vector $|v\rangle$ such that the set of vectors $R\,|v\rangle$ only partially spanned $\mathscr{E}^{(J)}$, then $\mathscr{E}^{(J)}$ would be reducible with respect to rotations.

The irreducibility of $\mathscr{E}^{(J)}$ can be demonstrated as follows. Let the space spanned by the vectors $R\,|u\rangle$ be denoted by $\mathscr{E}_1{}^{(J)}$. $J_+\,|u\rangle$ belongs to $\mathscr{E}_1{}^{(J)}$ for we have

$$J_+\,|u\rangle \equiv (J_x + \mathrm{i}J_y)\,|u\rangle = \frac{1}{\varepsilon}\,(1 - \mathrm{i} + \mathrm{i}R_x(\varepsilon) - R_y(\varepsilon))\,|u\rangle.$$

The same is true for $J_-\,|u\rangle$. More generally, any vector obtained by application of J_+ or J_- to a vector of $\mathscr{E}_1{}^{(J)}$ belongs to $\mathscr{E}_1{}^{(J)}$. Consider now the expansion $|u\rangle = \sum_M |JM\rangle\langle JM\,|u\rangle$ and denote by m the smallest value of M for which $\langle J\,M\,|u\rangle \neq 0$. Following the methods of § 5, the vector $J_+{}^{J-m}\,|u\rangle$ is a non-null vector proportional to $|J\,J\rangle$; thus $|J\,J\rangle$ belongs to $\mathscr{E}_1{}^{(J)}$, and since by repeated application of J_- to $|J\,J\rangle$ we form all of the $|J\,M\rangle$, they too belong to $\mathscr{E}_1{}^{(J)}$. Therefore, $\mathscr{E}_1{}^{(J)}$ contains a complete set of basis vectors for $\mathscr{E}^{(J)}$; the two spaces are therefore identical. Q.E.D.

As pointed out in § 6, the ket-vector space of a physical system is formed by the direct sum of a certain number of $(2J+1)$-dimensional subspaces $\mathscr{E}(\tau J)$, where τ represents the set of quantum numbers which distinguish between those corresponding to the same eigenvalue of $\boldsymbol{J}^2$. Each of the $\mathscr{E}(\tau J)$ is an *irreducible invariant subspace with respect to rotations*. In a standard representation $\{\boldsymbol{J}^2 J_z\}$, the components of $\boldsymbol{J}$ are represented in each of these subspaces by very simple τ-independent matrices. Similarly, each rotation operator $R(\alpha\,\beta\,\gamma)$ is represented in each of the $\mathscr{E}(\tau J)$ by a certain $(2J+1)$-dimensional matrix, $R^{(J)}(\alpha\,\beta\,\gamma)$, *depending on* $\boldsymbol{J}$ *but independent of the quantum numbers* τ. By definition:

$$\begin{aligned} R^{(J)}_{MM'}(\alpha\,\beta\,\gamma) &\equiv \langle \tau\,J\,M\,|R(\alpha\,\beta\,\gamma)\,|\tau\,J\,M'\rangle \\ &\equiv \langle J\,M\,|\mathrm{e}^{-\mathrm{i}\alpha J_z}\,\mathrm{e}^{-\mathrm{i}\beta J_y}\,\mathrm{e}^{-\mathrm{i}\gamma J_z}\,|J\,M'\rangle. \end{aligned} \qquad \text{(XIII.64)}$$

These matrices constitute a particularly convenient representation of the operators $R(\alpha\,\beta\,\gamma)$ and are commonly used each time it is necessary to change the orientation of the state vectors or observables. They are called *rotation matrices*. Their principal properties and the explicit expressions for some of them are given in Appendix C (Section IV).

From the very definition of these matrices it follows that the $(2j+1)$

basis vectors of a subspace $\mathscr{E}(\tau J)$ transform in a given rotation $\mathscr{R}(\alpha\,\beta\,\gamma)$ according to the law

$$R(\alpha\,\beta\,\gamma)\,|\tau\,J\,M\rangle = \sum_{M'} |\tau\,J\,M'\rangle\, R^{(J)}_{M',M}(\alpha\,\beta\,\gamma). \qquad \text{(XIII.65)}$$

One can easily demonstrate the converse, namely: if $(2J+1)$ vectors $|u_M\rangle$ $(M=-J,\,-J+1,\,\ldots,\,+J)$ transform one into another by rotation according to the law

$$R(\alpha\,\beta\,\gamma)\,|u_M\rangle = \sum_{M'} |u_{M'}\rangle\, R^{(J)}_{M',M}(\alpha\,\beta\,\gamma) \qquad \text{(XIII.66)}$$

then they satisfy the eigenvalue equations

$$\boldsymbol{J}^2\,|u_M\rangle = J(J+1)\,|u_M\rangle, \qquad J_z\,|u_M\rangle = M\,|u_M\rangle,$$

and are obtained one from another by application of J_+ and J_- in accordance with the relations (XIII.24–25).

17. Rotational Invariance and Conservation of Angular Momentum. Rotational Degeneracy

The invariance of a quantity under rotation can always be expressed as a special property of angular momentum. This is because any rotation can be expressed as a product of infinitesimal rotations and therefore if a quantity is invariant under these, it is invariant under any rotation whatever. Through relations (XIII.55) the angular momentum appears in the condition of invariance under infinitesimal rotations.

Thus, for a wave function or for a ket $|\rangle$ to be rotationally invariant, a necessary and sufficient condition is that the application of an arbitrary component of the total angular momentum give zero:

$$\boldsymbol{J}\,|\rangle = 0.$$

In fact, it is sufficient that

$$\boldsymbol{J}^2\,|\rangle = 0. \qquad \text{(XIII.67)}$$

Such is the case for the wave functions of a particle in the s-state; functions of this type depend only on the variable r. It is also the case for wave functions for several particles depending only on the distances between the particles and on the angles of the position vectors with respect to one another [1]).

---

[1]) This is to be compared with the property $(l+l')\,P_l(\cos\alpha) = 0$ which appears in the proof of the addition theorem (Problem XIII.5).

Similarly, for an observable S to be invariant under rotation [condition (XIII.53)], a necessary and sufficient condition is that it commute with the components of angular momentum.

$$[\boldsymbol{J}, S]=0. \tag{XIII.68}$$

The invariance of the Hamiltonian under rotation deserves particular attention. If we have

$$[R, H]=0 \text{ for any } R \tag{XIII.69}$$

then *the equations of motion are invariant under rotation*: two state vectors that are the transforms one of the other in a given rotation at time t_0, will continue to be so related throughout the course of time. This is obvious, for if $|\psi(t)\rangle$ satisfies the Schrödinger equation, we have, whatever R,

$$\left(i\hbar\frac{\partial}{\partial t}-H\right)R|\psi(t)\rangle=R\left(i\hbar\frac{\partial}{\partial t}-H\right)|\psi(t)\rangle=0$$

and therefore $R|\psi(t)\rangle$ is also a solution of the Schrödinger equation.

Similarly, if $|\rangle$ is an eigenvector of H, all the vectors $R|\rangle$ that can be obtained by rotation are also eigenvectors of H belonging to the same eigenvalue i.e., the subspace of each eigenvalue of H is invariant with respect to rotations.

All of the consequences of invariance under rotation of the equations of motion are contained in the relations

$$[\boldsymbol{J}, H]=0 \tag{XIII.70}$$

which express the invariance of H with respect to infinitesimal rotations.

When these relations are satisfied the operators $\boldsymbol{J}^2$, J_z and H commute, and the solving of the eigenvalue problem for H is considerably simplified – we may seek the eigenfunctions of H among the common eigenfunctions of $\boldsymbol{J}^2$ and J_z. Moreover, the energy spectra corresponding to the same value of J are the same, the eigenfunctions belonging to the $(2J+1)$ possible values of M being obtained one from another by repeated application of J_+ or J_-. In other words, the energy eigenvalues are independent of M. To each eigenvalue E_J corresponding to a given value of J there corresponds one or several series of $(2J+1)$ eigenvectors; the eigenvectors of a given series are obtained one from another by repeated application of J_+ or J_- and span an

irreducible invariant subspace with respect to rotations. This type of degeneracy is called *rotational degeneracy.*

The case of a particle in a central field (Ch. IX) is a good illustration of the present discussion. The Hamiltonian of a particle in a central field must obviously be invariant under rotation; one verifies directly that it commutes with the three components of the angular momentum $\boldsymbol{l}$. The method of Chapter IX consists precisely in seeking the eigenfunctions of H among the common eigenfunctions of $\boldsymbol{l}^2$ and l_z belonging to the eigenvalues $l(l+1)$ and m respectively, i.e., among the functions of form

$$\chi_l(r)\, Y_l^m(\theta, \varphi).$$

Such a problem reduces to solving a second-order differential equation in r. Moreover, since m does not appear in that equation, we can form from each of the radial functions thus determined in all $(2l+1)$ eigenfunctions of H belonging to the same eigenvalue.

As previously stated, we have here a striking analogy between Classical Mechanics and Quantum Mechanics. When the equations of motion of a classical system are invariant under rotations of the coordinate axes, the total angular momentum of the system is conserved, a property which permits us to obtain first integrals of the motion and which simplifies considerably the solving of the equations. In the same way, the invariance under rotation of the equations of motion in Quantum Mechanics leads to the conservation of the total angular momentum; however, here the conservation laws are not so simply expressed due to the fact that the components of angular momentum do not commute.

IV. SPIN

18. The Hypothesis of Electron Spin

The Schrödinger theory as it results from the simple application of the correspondence principle cannot explain the properties of complex atoms, even leaving relativistic corrections aside. Two important modifications are required. Neither have any analogue in classical mechanics that would have permitted their existence to be foreseen. The one consists in retaining only those solutions of the Schrödinger theory that have certain well-defined symmetry properties

in a permutation of the coordinates of the electrons. It is known as the Pauli principle and will be studied in Chapter XIV. It may be ignored for the purpose of the present discussion. The other is the hypothesis of electron spin.

The principal evidence supporting this hypothesis comes from the study of the behavior of complex atoms in a magnetic field (Zeeman effect, Stern–Gerlach experiment).

The Schrödinger equation for an atom of Z (spinless) electrons has already been written down [eq. (II.30)]. If we suppose the nucleus to be infinitely heavy its position will coincide with the center of mass, and the Hamiltonian in the center of mass system is simply

$$H_0 = \sum_{i=1}^{Z} \left(\frac{\boldsymbol{p}_i^2}{2m} - \frac{Ze^2}{r_i} \right) + \sum_{i<j} \frac{e^2}{|\boldsymbol{r}_i - \boldsymbol{r}_j|}. \qquad \text{(XIII.71)}$$

To obtain the Hamiltonian of the same atom placed in the static magnetic field described by the potential $\boldsymbol{A}(\boldsymbol{r})$ we need only to replace each $\boldsymbol{p}_i$ by $\boldsymbol{p}_i - e\boldsymbol{A}(\boldsymbol{r}_i)/c$. In particular, for a constant magnetic field $\mathscr{H}$, $\boldsymbol{A} = \frac{1}{2}(\mathscr{H} \times \boldsymbol{r})$ and

$$\left(\boldsymbol{p} - \frac{e}{c}\boldsymbol{A} \right)^2 = \boldsymbol{p}^2 - \frac{e}{c}(\boldsymbol{A}\cdot\boldsymbol{p} + \boldsymbol{p}\cdot\boldsymbol{A}) + \frac{e^2}{c^2}\boldsymbol{A}^2$$

$$= \boldsymbol{p}^2 - \frac{e}{c}(\mathscr{H}\cdot\boldsymbol{l}) + \frac{e^2}{4c^2}\mathscr{H}^2 r_\perp^2,$$

where $r_\perp^2$ is the square of the projection of $\boldsymbol{r}$ on the plane perpendicular to the field $\mathscr{H}$. The Hamiltonian becomes

$$H = H_0 - \frac{e}{2mc}\mathscr{H}\cdot\boldsymbol{L} + \frac{e^2}{8mc^2}\mathscr{H}^2 \sum_{i=1}^{Z} r_{i\perp}^2;$$

$\boldsymbol{L}$ is the total angular momentum of the Z electrons: $\boldsymbol{L} \equiv \sum_i (\boldsymbol{r}_i \times \boldsymbol{p}_i)$. In the phenomena that we shall study, the contribution of the third term in the above expression is entirely negligible [1]. Thus we have,

---

[1] This term is the main factor in atomic diamagnetism. Knowing that $\langle r^2 \rangle \simeq 10^{-16}$ cm^2, we can evaluate its order of magnitude $(Ze^2/12\,mc^2)\,\mathscr{H}^2\langle r^2\rangle$. The ratio of this to the level distance $e\hbar\mathscr{H}/2\,mc$ found below is about $10^{-9}Z\mathscr{H}$ (gauss), a negligibly small quantity, even for a very strong field and a very heavy atom. Thus its neglect can in no way be held responsible for the conflicts which appear further along.

to a very good approximation

$$H = H_0 - \frac{e}{2mc}(\mathcal{H} \cdot \mathbf{L}). \tag{XIII.72}$$

This is just what we would have if each electron in circulating in its orbit induced a magnetic moment

$$\mu = \frac{e}{2mc} l$$

proportional to its angular momentum, the constant of proportionality (the gyromagnetic ratio) being precisely equal to the value given by the classical theory of this effect, namely $e/2mc$. In this interpretation the total magnetic moment of the atom is equal to the sum of the Z individual magnetic moments, i.e.

$$\mathcal{M} = \frac{e}{2mc} \mathbf{L},$$

and the energy of the atom placed in the field $\mathcal{H}$ differs from the energy in the absence of the field by the magnetic energy term $-(\mathcal{M} \cdot \mathcal{H})$.

A certain number of important properties can be deduced simply by inspection of expression (XIII.72) if we take into account that H_0, being invariant under rotation, commutes with the three components of $\mathbf{L}$.

Let us take the direction of $\mathcal{H}$ as the z axis. H_0, $\mathbf{L}^2$ and L_z have a common set of eigenvectors $|n\, L\, M\rangle$ and the corresponding eigenvalues of H_0, $E_0{}^{nL}$, are independent of M and are $(2L+1)$-fold degenerate [1]).

H is a function of H_0 and L_z [eq. (XIII.72)] and therefore has the same set of eigenfunctions, the eigenvalue of the vector $|nLM\rangle$ being

$$E^{nLM} = E_0{}^{nL} - M\mu_B \mathcal{H}, \tag{XIII.73}$$

---

[1]) If several of these eigenvalues accidentally coincide (as in the hydrogen atom) the degeneracy is higher. Suppose $E_0{}^{nL} = E_0{}^{n'L'}$; the degeneracy is then of order $(2L + 1) + (2L' + 1)$. In this case the argument that follows above needs some modification of detail. However, the conclusions drawn are valid if we everywhere replace L by the larger of L and L'. In particular the result that each Zeeman "multiplet" has an odd number of equidistant levels is not modified.

where

$$\mu_B = \frac{e\hbar}{2mc} \qquad \text{(Bohr magneton).} \qquad \text{(XIII.74)}$$

Since M can take all integral values from $-L$ to $+L$, each level $E_0{}^{nL}$ gets split under the effect of the field $\mathscr{H}$ into $(2L+1)$ distinct, equidistant levels distributed according to (XIII.73). We therefore have the following theoretical predictions:

(i) Each level $E_0{}^{nL}$ of the atomic spectrum is split by the field $\mathscr{H}$ into a "multiplet" of $(2L+1)$ equidistant levels;

(ii) These levels are distributed on either side of $E_0{}^{nL}$ in such a way that their distance from $E_0{}^{nL}$ averages to zero;

(iii) The distance between two neighboring levels is a quantity $\mu_B \mathscr{H}$ that is independent of the atom considered and proportional to $\mathscr{H}$.

These theoretical predictions are only partially confirmed by experiment. There are two important discrepancies:

(a) in atoms of odd Z, the multiplets are all even, which is just what one would have if L were half-integral;

(b) the distance between neighbors in the same multiplet is found to be $g\mu_B \mathscr{H}$, the factor g (the Landé factor) varying from one multiplet to another within rather large limits.

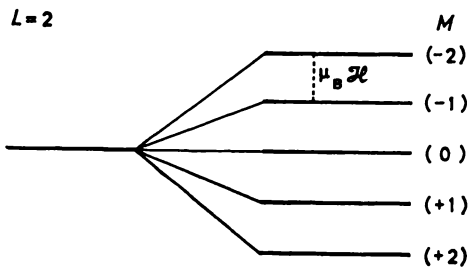


Fig. XIII.2. Zeeman effect for a D state ($L = 2$). On the left, energy level in absence of field; on the right, corresponding levels when $\mathscr{H} \neq 0$.

The existence of half-integral angular momenta is directly established by the Stern–Gerlach experiment (§ I.10). As the atoms making up the beam are almost all in their ground state, the number of spots observed on the screen is equal to the multiplicity of the ground state.

With silver atoms we observe in all two spots, hence the ground state of the silver atom has angular momentum $\frac{1}{2}$. More generally, atoms having odd Z invariably give an even number of spots, result characteristic of a half-integral angular momentum.

Points (a) and (b) show up simultaneously in the study of the anomalous Zeeman effect. The spectral results in general permit the simultaneous determination of the multiplicity of the states between which the optical transitions are effected, and their respective Landé g-factors.

To surmount these difficulties it is necessary to introduce half-integral angular momenta and gyromagnetic ratios differing from $e/2mc$. This is what is very simply done by the hypothesis of electron spin (Uehlenbeck and Goudsmit, 1925):

Each electron has an intrinsic angular momentum or spin $\mathbf{s}$, *of magnitude* $\frac{1}{2}\hbar$ (*spin* $\frac{1}{2}$) *with which is associated the magnetic moment*

$$\boldsymbol{\mu}_s = g_s \frac{e}{2mc} \mathbf{s}. \tag{XIII.75}$$

g_s is an adjustable constant. The theory is in excellent agreement with experiment if we take

$$g_s \simeq 2. \tag{XIII.76}$$

This value for g_s is explained by the relativistic theory of the electron (Ch. XX).

Experiment shows that the nucleons (protons and neutrons) also have a spin $\frac{1}{2}$, which can be directly revealed by measuring the associated magnetic moment [1]).

In the rest of this section we develop the non-relativistic theory of particles with spin $\frac{1}{2}$ (*the Pauli Theory*).

19. Spin $\frac{1}{2}$ and the Pauli Matrices

Let $\mathbf{s}$ be the intrinsic angular momentum (or spin vector) of a particle of spin $\frac{1}{2}$. By hypothesis, $\mathbf{s}^2$ has just the one eigenvalue

[1]) If $\boldsymbol{\mu}_p$, $\mathbf{s}_p$, M_p denote respectively the magnetic moment, the spin and the mass of the proton, one has [cf. eq. (XIII.75)]:

$$\boldsymbol{\mu}_p = g_p \frac{e}{2M_pc} \mathbf{s}_p.$$

There is an analogous formula for the neutron. Experiment gives $g_p = 5.59$ and $g_n = -3.83$.

$s(s+1) = \frac{1}{2} \times \frac{3}{2} = \frac{3}{4}$. Each component, for example s_z, can take one or the other of the values $\pm \frac{1}{2}$. We suppose these to be non-degenerate. Consequently the components of $\mathbf{s}$ are operators acting in a two-dimensional space in which a possible basis is constituted by the two eigenvectors

$$|+\rangle \equiv |\tfrac{1}{2} \ +\tfrac{1}{2}\rangle, \qquad |-\rangle \equiv |\tfrac{1}{2} \ -\tfrac{1}{2}\rangle$$

of $\mathbf{s}^2$ and s_z.

If we adopt this basis we can easily write down the matrices representing the operators s_x, s_y, s_z. They are particular J_x, J_y, J_z matrices with elements given by equations (XIII.28).

In addition to the commutation relations characteristic of angular momenta, the components of $\mathbf{s}$ verify the special relations

$$s_x^2 = s_y^2 = s_z^2 = \tfrac{1}{4}, \qquad s_+^2 = s_-^2 = 0.$$

Since

$$s_+^2 = (s_x + \mathrm{i}\, s_y)^2 = (s_x^2 - s_y^2) + \mathrm{i}\,(s_x\, s_y + s_y\, s_x),$$

we can deduce that

$$s_x\, s_y + s_y\, s_x = 0,$$

i.e. that the operators s_x, s_y, s_z anticommute [1]).

It is convenient to introduce the *Pauli matrices* $\boldsymbol{\sigma} \equiv (\sigma_x, \sigma_y, \sigma_z)$ defined by the equation

$$\mathbf{s} = \tfrac{1}{2}\boldsymbol{\sigma}. \tag{XIII.77}$$

They are given explicitly by

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Their principal properties, which can be deduced from their definition and which can easily be verified in their explicit form, are summarized by the following equations:

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = 1. \tag{XIII.78}$$

$$\sigma_x\, \sigma_y = -\sigma_y\, \sigma_x = \mathrm{i}\sigma_z. \tag{XIII.79a}$$

$$\sigma_y\, \sigma_z = -\sigma_z\, \sigma_y = \mathrm{i}\sigma_x. \tag{XIII.79b}$$

$$\sigma_z\, \sigma_x = -\sigma_x\, \sigma_z = \mathrm{i}\sigma_y. \tag{XIII.79c}$$

$$\sigma_x\, \sigma_y\, \sigma_z = \mathrm{i}. \tag{XIII.80}$$

$$\mathrm{Tr}\ \sigma_x = \mathrm{Tr}\ \sigma_y = \mathrm{Tr}\ \sigma_z = 0. \tag{XIII.81}$$

$$\det \sigma_x = \det \sigma_y = \det \sigma_z = -1. \tag{XIII.82}$$

---

[1]) Two operators A, B anticommute if $AB + BA = 0$

From these we can deduce the important identity (Problem XIII.9)

$$(\boldsymbol{\sigma}\cdot\boldsymbol{A})\,(\boldsymbol{\sigma}\cdot\boldsymbol{B})=(\boldsymbol{A}\cdot\boldsymbol{B})+\mathrm{i}\boldsymbol{\sigma}\cdot(\boldsymbol{A}\times\boldsymbol{B}), \tag{XIII.83}$$

$\boldsymbol{A}$ and $\boldsymbol{B}$ being any two vectors [1]).

Since $\boldsymbol{s}$ is the angular momentum, the rotation operator $R_u^{(s)}(\varphi)$ effecting the transformation of vectors of that space in the rotation $\mathscr{R}_u(\varphi)$ is, from (XIII.58)

$$R_u^{(s)}(\varphi)\equiv \mathrm{e}^{-\frac{1}{2}\mathrm{i}\varphi\sigma_u},$$

where $\sigma_u\equiv(\boldsymbol{u}\cdot\boldsymbol{\sigma})$. Expanding the exponential and separately summing the terms even and odd in σ_u with the aid of the relations

$$\sigma_u{}^{2p}=1, \qquad \sigma_u{}^{2p+1}=\sigma_u$$

[cf. eq. (XIII.83)] we obtain the very simple expression

$$R_u^{(s)}(\varphi)=\cos\tfrac{1}{2}\varphi-\mathrm{i}\sigma_u\sin\tfrac{1}{2}\varphi. \tag{XIII.84}$$

Observe that the operator of rotation through 2π is equal to -1, in agreement with the results of § 15.

The operator representing the rotation $\mathscr{R}(\alpha\,\beta\,\gamma)$ is, from (XIII.60)

$$R^{(s)}(\alpha\,\beta\,\gamma)=\mathrm{e}^{-\frac{1}{2}\mathrm{i}\alpha\sigma_z}\,\mathrm{e}^{-\frac{1}{2}\mathrm{i}\beta\sigma_y}\,\mathrm{e}^{-\frac{1}{2}\mathrm{i}\gamma\sigma_z}. \tag{XIII.85}$$

Its explicit form can be calculated in the same way as that of $\mathscr{R}_u(\varphi)$ and is given in the Appendix [formula (C.74)].

The vectors of the space considered here bear some analogy to those of ordinary space. The latter are geometric entities with three components and which under rotation transform one into another following a well-determined law. Such is also the case with those considered here [transformation law (XIII.85)] except that they have two components instead of three. They are called *spinors*.

20. Observables and Wave Functions of a Spin $\frac{1}{2}$ Particle. Spinor Fields

Consider a particle of spin $\frac{1}{2}$.

Its fundamental variables may be put into two categories, the

[1]) or two vector operators providing their components commute with those of $\boldsymbol{\sigma}$. In this case the order of $\boldsymbol{A}$ and $\boldsymbol{B}$ in the two sides of the identity must be respected. Example:

$$(\boldsymbol{\sigma}\cdot\boldsymbol{r})\,(\boldsymbol{\sigma}\cdot\boldsymbol{p})=(\boldsymbol{r}\cdot\boldsymbol{p})+\mathrm{i}\,\boldsymbol{\sigma}\cdot(\boldsymbol{r}\times\boldsymbol{p}).$$

orbital variables and the intrinsic, or spin, variables. The first are the components of position $\boldsymbol{r}$ and momentum $\boldsymbol{p}$; they verify the fundamental commutation relations ($\hbar = 1$)

$$[r_i, p_j] = i\delta_{ij}.$$

The second are the components of the spin $\boldsymbol{s}$; they verify the commutation relations

$$[s_i, s_j] = i\varepsilon_{ijk}\, s_k,$$

and in addition must satisfy the supplementary condition $\boldsymbol{s}^2 = \frac{3}{4}$.

As the orbital variables commute with the spin variables, the space $\mathscr{E}$ of the state vectors of the particle is the tensor product

$$\mathscr{E} = \mathscr{E}^{(0)} \otimes \mathscr{E}^{(s)}$$

of the orbital space $\mathscr{E}^{(0)}$ and of the spin space $\mathscr{E}^{(s)}$ (cf. § VIII.7). $\mathscr{E}^{(0)}$ is the state space of a particle without spin, $\mathscr{E}^{(s)}$ is the two-dimensional space constructed in the preceding paragraph.

To represent the vectors of $\mathscr{E}$, one usually chooses the representation with $\boldsymbol{r}$ and s_z diagonal; each vector $|\psi\rangle$ is then represented by the wave function

$$\psi(\boldsymbol{r}, \mu) \equiv \langle \boldsymbol{r}\, \mu\, |\psi\rangle, \tag{XIII.86}$$

a function of the continuous variables $\boldsymbol{r} \equiv (x, y, z)$, and of the discrete variable μ representing the eigenvalue of s_z and taking the two values $\pm \frac{1}{2}$.

The total angular momentum of the particle is

$$\boldsymbol{j} \equiv \boldsymbol{l} + \boldsymbol{s}. \tag{XIII.87}$$

The fundamental variables of the system are the components of the three vector-operators $\boldsymbol{r}$, $\boldsymbol{p}$, $\boldsymbol{s}$. Clearly $\boldsymbol{j}$ verifies with each of these the commutation relations (XIII.57) characterizing the total angular momentum, for $\boldsymbol{l} \equiv \boldsymbol{r} \times \boldsymbol{p}$ verifies these relations with $\boldsymbol{r}$ and $\boldsymbol{p}$ and commutes with $\boldsymbol{s}$, and $\boldsymbol{s}$ verifies these relations with itself and commutes with $\boldsymbol{r}$ and $\boldsymbol{p}$.

The rotation operator $R(\alpha\, \beta\, \gamma)$ can now be deduced [eq. (XIII.60)]. Since $\boldsymbol{l}$ and $\boldsymbol{s}$ commute, it takes the form of a product of two commuting operators:

$$R(\alpha\, \beta\, \gamma) = R^{(s)}(\alpha\, \beta\, \gamma)\, R^{(0)}(\alpha\, \beta\, \gamma) \tag{XIII.88}$$

$R^{(s)}(\alpha\,\beta\,\gamma)$, defined by equation (XIII.85), *rotates the spin*, while $R^{(0)}(\alpha\,\beta\,\gamma)$, defined by

$$R^{(0)}(\alpha\,\beta\,\gamma)=\mathrm{e}^{-\mathrm{i}\alpha l_z}\,\mathrm{e}^{-\mathrm{i}\beta l_y}\,\mathrm{e}^{-\mathrm{i}\gamma l_z}$$

rotates the orbital variables.

In a rotation through 2π, $R^{(0)}=1$ and $R^{(s)}=-1$, and therefore all kets change sign in such a rotation. However, all of the fundamental observables are invariant in a rotation through 2π and therefore, following the discussion of § 15, their physical interpretation raises no difficulty.

It is often convenient to put

$$\psi(\boldsymbol{r},\,\pm\tfrac{1}{2})=\psi_{\pm}(\boldsymbol{r})$$

and to write the wave function $\psi(\boldsymbol{r},\mu)$ in the form of a *two component wave function*:

$$\psi\equiv\begin{pmatrix}\psi_{+}(\boldsymbol{r})\\ \psi_{-}(\boldsymbol{r})\end{pmatrix}.$$

For each value of $\boldsymbol{r}$, ψ represents a ket vector of the space $\mathscr{E}^{(s)}$, namely

$$\langle\boldsymbol{r}\,|\psi\rangle\equiv\psi_{+}(\boldsymbol{r})\,|+\rangle+\psi_{-}(\boldsymbol{r})\,|-\rangle. \tag{XIII.89}$$

In other words, the wave function may be regarded as a spinor field [1]).

The extension of these considerations to a system containing Z particles of spin $\frac{1}{2}$ is straightforward. The state space for the whole system is the tensor product of the state spaces for each individual particle. In particular, the total spin space has 2^Z dimensions and is the tensor product of the Z individual spin spaces. A system of Pauli matrices, $\boldsymbol{\sigma}^{(i)}$, is then introduced for each individual spin. An

---

[1]) In a rotation $\mathscr{R}(\alpha\beta\gamma)$, the spinor field ψ transforms into

$$\mathscr{R}[\psi]=R\,\psi=R^{(\frac{1}{2})}\begin{pmatrix}\psi_{+}(\mathscr{R}^{-1}\,\boldsymbol{r})\\ \psi_{-}(\mathscr{R}^{-1}\,\boldsymbol{r})\end{pmatrix}.$$

This is a direct result of (XIII.88); $R^{(\frac{1}{2})}$ is the rotation matrix corresponding to $J=\frac{1}{2}$. We may compare this transformation law to formula XIII.47 for a scalar field. We have an analogous formula for a vector field, with $R^{(1)}$ in place of $R^{(\frac{1}{2})}$ (cf. § 21).

overall rotation of the spins can be expressed with the aid of the total spin:

$$\boldsymbol{S} = \tfrac{1}{2} \sum_{i=1}^{Z} \boldsymbol{\sigma}^{(i)}. \qquad \text{(XIII.90)}$$

An overall rotation through 2π of the spins is represented by the operator $(-)^Z$.

21. Vector Fields and Particles of Spin 1

It is well to stress the parallel between the concept of a spinor field and the more familiar concept of a vector field.

Let $\boldsymbol{A}(\boldsymbol{r})$ be a vector field associated with a physical system. It might, for example, be a magnetic or electric field or, as we shall see, the wave function of a particle of spin 1.

Let us examine how $\boldsymbol{A}(\boldsymbol{r})$ transforms under rotation. Let $\boldsymbol{A}'(\boldsymbol{r})$ be the rotated field resulting from an overall rotation $\mathscr{R}$ of the physical system:

$$\boldsymbol{A}' \equiv \mathscr{R}\,[\boldsymbol{A}].$$

The field $\boldsymbol{A}'$ at the point $\boldsymbol{r}$ is obtained by applying the rotation $\mathscr{R}$ to the vector $A(\boldsymbol{r}_1)$ representing the field $\boldsymbol{A}$ at the point $\boldsymbol{r}_1 \equiv \mathscr{R}^{-1}\boldsymbol{r}$, i.e. [cf. eq. (XIII.43) and (XIII.47)]:

$$A_i'(\boldsymbol{r}) = \mathscr{R}_{ij}\, A_j(\mathscr{R}^{-1}\boldsymbol{r}) \qquad (i = x, y, z).$$

Thus for a rotation α about Oz we find [cf. eq. (XIII.44)]

$$\boldsymbol{A}' \equiv \mathscr{R}_z(\alpha)\,[\boldsymbol{A}] \qquad \boldsymbol{r}_1 \equiv (x\cos\alpha + y\sin\alpha, -x\sin\alpha + y\cos\alpha, z)$$
$$A'_x(\boldsymbol{r}) = A_x(\boldsymbol{r}_1)\cos\alpha - A_y(\boldsymbol{r}_1)\sin\alpha$$
$$A'_y(\boldsymbol{r}) = A_x(\boldsymbol{r}_1)\sin\alpha + A_y(\boldsymbol{r}_1)\cos\alpha$$
$$A'_z(\boldsymbol{r}) = A_z(\boldsymbol{r}_1).$$

In particular, the infinitesimal rotation ε about Oz gives

$$\mathscr{R}_z(\varepsilon)\,[\boldsymbol{A}] = (1 - \mathrm{i}\varepsilon\,(l_z + s_z))\,\boldsymbol{A}, \qquad \text{(XIII.91)}$$

where l_z is the above-defined differential operator and s_z the operator defined by

$$s_z \begin{pmatrix} A_x(\boldsymbol{r}) \\ A_y(\boldsymbol{r}) \\ A_z(\boldsymbol{r}) \end{pmatrix} = \begin{pmatrix} -\mathrm{i}A_y(\boldsymbol{r}) \\ \mathrm{i}A_x(\boldsymbol{r}) \\ 0 \end{pmatrix}.$$

s_z transforms each component of the field at a given point into a

particular linear combination of the three components of the field at the same point. $\boldsymbol{A}$ being defined by its three Cartesian components A_x, A_y and A_z, s_z is represented by the matrix

$$s_z = \begin{pmatrix} 0 & -\mathrm{i} & 0 \\ \mathrm{i} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

One similarly defines the operators s_x and s_y; their representative matrices are

$$s_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -\mathrm{i} \\ 0 & \mathrm{i} & 0 \end{pmatrix} \qquad s_y = \begin{pmatrix} 0 & 0 & \mathrm{i} \\ 0 & 0 & 0 \\ -\mathrm{i} & 0 & 0 \end{pmatrix}.$$

We easily show that s_x, s_y, s_z verify the commutation relations characteristic of the components of an angular momentum. We denote this angular momentum by $\boldsymbol{s}$; the calculation of its square gives:

$$\boldsymbol{s}^2 = 2$$

which corresponds to an angular momentum $s = 1$. By definition, we shall say that $\boldsymbol{s}$ is the intrinsic angular momentum, or spin, of the vector field.

A field such as $\boldsymbol{A}(\boldsymbol{r})$ can describe a particle of spin 1. Put:

$$A_i(\boldsymbol{r}) \equiv A(\boldsymbol{r}, i) \qquad (i = x, y \text{ or } z),$$

$A(\boldsymbol{r}, i)$ is a wave function which depends not only on the position variables but also on an index i which may take 3 values and which constitutes an internal variable describing the orientation of the particle. The scalar product of two wave functions of this type is:

$$\langle B, A \rangle \equiv \sum_i \int B^*(\boldsymbol{r}, i)\, A(\boldsymbol{r}, i)\, \mathrm{d}\boldsymbol{r} = \int (\boldsymbol{B}^* \cdot \boldsymbol{A})\, \mathrm{d}\boldsymbol{r}. \qquad \text{(XIII.92)}$$

An operator such as $\boldsymbol{l}$ acts on the position variables alone, while $\boldsymbol{s}$ acts on the internal variable i. Clearly $\boldsymbol{l}$ and $\boldsymbol{s}$ commute since they act on different variables. The operator of infinitesimal rotation about the z axis is defined by equation (XIII.91); we likewise obtain the operator of infinitesimal rotation about any other axis; applying definition (XIII.55) we find the total angular momentum of the particle:

$$\boldsymbol{j} \equiv \boldsymbol{l} + \boldsymbol{s}$$

[cf. eq. (XIII.87)].

More generally, any linear transformation of a vector field may

be looked upon as the action of a certain linear operator that can be expressed as a function of the three fundamental vector operators:

$$\boldsymbol{r}, \quad \boldsymbol{p} \equiv -i\nabla, \quad \boldsymbol{s}.$$

In particular, we have the important identity

$$\text{curl} \equiv \boldsymbol{s} \cdot \boldsymbol{p} \tag{XIII.93}$$

which may easily be verified from the definition of the curl and the explicit form of the matrices s_x, s_y, s_z given above.

The concepts of scalar product, of rotation, and more generally, of linear transformation, are independent of the representation chosen. The wave function $\boldsymbol{A}(\boldsymbol{r}, i)$ represents the dynamical state of the particle in a representation in which the basis vectors for the internal variable correspond to unit vectors along each of the three axes Ox, Oy, Oz; these basis vectors, $|x\rangle$, $|y\rangle$, $|z\rangle$ are respectively eigenvectors of s_x, s_y, s_z belonging to the eigenvalue 0. (cf. Problem XIII.10.) It is often more convenient to choose the representation where the basis vectors are the eigenvectors of s_z, $|+\rangle$, $|0\rangle$, $|-\rangle$, belonging respectively to the eigenvalues $+1$, 0, -1. These are obtained one from another by the standard law defined in § 6. In this new representation, s_x, s_y and s_z are represented by matrices verifying relations (XIII.28) (with $j = j' = 1$) and the ket $|A\rangle$ associated with the vector field $\boldsymbol{A}$ is represented by the wave function

$$A(\boldsymbol{r}, \mu) \equiv A_\mu(\boldsymbol{r}) \qquad (\mu = +, 0, -),$$

from the definition [cf. eq. (XIII.89)]

$$\langle \boldsymbol{r} | A \rangle \equiv A_+(\boldsymbol{r}) | + \rangle + A_0(\boldsymbol{r}) | 0 \rangle + A_-(\boldsymbol{r}) | - \rangle.$$

We have:

$$\begin{aligned} A_+ &= -\frac{\sqrt{2}}{2}(A_x - iA_y) \\ A_0 &= A_z \\ A_- &= \frac{\sqrt{2}}{2}(A_x + iA_y). \end{aligned} \tag{XIII.94}$$

22. Spin-dependent Interactions in Atoms

The existence of intrinsic magnetic moment leads to spin-dependent terms in the Hamiltonian of *an electron in an electromagnetic field.* In particular, in the presence of a magnetic field $\mathcal{H}(\boldsymbol{r})$ we have

the *direct coupling* term suggested by the correspondence principle:

$$-\boldsymbol{\mu}\cdot\mathscr{H}(\boldsymbol{r}) \equiv -\mu_B\boldsymbol{\sigma}\cdot\mathscr{H},$$

where $\boldsymbol{\mu}$ is the intrinsic magnetic moment defined by equations (XIII.75–76).

This is not the only additional term. Even with a purely electrostatic potential it is clear that terms due to *spin-orbit coupling* will appear, for an electron moving in such a potential "sees" a magnetic field, which field can interact with $\boldsymbol{\mu}$. This argument from Classical Mechanics can be used as a guide for the empirical determination of the spin-orbit coupling. However, since we are dealing here with a relativistic effect (tending to zero in the limit $v \ll c$), it is preferable that it be derived from the relativistic equation of the electron. This can be done by making an expansion in v/c and retaining non-zero terms of the lowest order. This problem will be studied in Chapter XX. In a spherically symmetrical potential $V(r)$, the spin-orbit interaction is obviously invariant under rotation, and therefore commutes with all three components of the total angular momentum $\boldsymbol{j}$. The expression given by the relativistic theory is

$$\frac{\hbar^2}{2m^2c^2}(\boldsymbol{l}\cdot\boldsymbol{s})\frac{1}{r}\frac{\mathrm{d}V}{\mathrm{d}r}. \qquad \text{(XIII.95)}$$

For the same reasons, the Hamiltonian H_0 of the Z electrons of a complex atom contains spin-orbit terms in addition to the Coulomb terms shown in eq. (XIII.71). These additional terms commute with the total angular momentum

$$\boldsymbol{J}=\boldsymbol{L}+\boldsymbol{S}$$

but, unlike the rest of H_0, they do not commute with $\boldsymbol{L}$ and $\boldsymbol{S}$ separately. Moreover, although their contribution to the total energy is relatively very small (except for the heaviest atoms), their presence results in a qualitative modification in the atomic spectrum – removal of the degeneracy – and therefore can never be ignored [1]).

The Hamiltonian H of an atom in a constant magnetic field $\mathscr{H}$ is obtained by applying the treatment of § 18 to the Hamiltonian without external field, H_0, and adding the direct magnetic interaction

---

[1]) For completeness, one should also mention the modifications due to the existence of the magnetic moment of the nucleus (hyperfine structure).

terms $-\sum_i \boldsymbol{\mu}^{(i)} \cdot \mathscr{H}$. If we neglect the "diamagnetic term" in $\mathscr{H}^2$, as we did in equation (XIII.72), we find

$$H = H_0 - \frac{e}{2mc} [\mathscr{H} \cdot (\boldsymbol{L} + 2\boldsymbol{S})]. \tag{XIII.96}$$

23. Spin-dependent Nucleon-Nucleon Interactions

As a further example of spin-dependent interactions, we consider the interaction of two nucleons. Let M_0 be the mass of the nucleons, $\boldsymbol{r} = \boldsymbol{r}_1 - \boldsymbol{r}_2$ their relative position, $\boldsymbol{p} = \frac{1}{2}(\boldsymbol{p}_1 - \boldsymbol{p}_2)$ their relative momentum, $\frac{1}{2}\boldsymbol{\sigma}_1$ and $\frac{1}{2}\boldsymbol{\sigma}_2$ their respective spins. The motion of the center of mass separates completely from the relative motion; the dynamical variables and dynamical states considered below refer exclusively to the relative motion. The orbital angular momentum is

$$\boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p},$$

the total spin

$$\boldsymbol{S} = \tfrac{1}{2}(\boldsymbol{\sigma}_1 + \boldsymbol{\sigma}_2) \tag{XIII.97}$$

and the total angular momentum

$$\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{S}. \tag{XIII.98}$$

The Hamiltonian is of the form

$$H = \frac{\boldsymbol{p}^2}{M_0} + V.$$

The four types of rotationally-invariant interaction most commonly proposed are

$$V_1(r) \tag{XIII.99a}$$

$$V_2(r)\,(\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2) \tag{XIII.99b}$$

$$V_3(r)\,(\boldsymbol{L} \cdot \boldsymbol{S}) \tag{XIII.99c}$$

$$V_4(r)\left[3\frac{(\boldsymbol{\sigma}_1 \cdot \boldsymbol{r})\,(\boldsymbol{\sigma}_2 \cdot \boldsymbol{r})}{r^2} - \boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2\right]. \tag{XIII.99d}$$

The spin-dependent operators appearing in the last three expressions are written in their traditional form. They can be expressed differently. Thus, squaring both sides of (XIII.97) and using the identity

$$\boldsymbol{\sigma}_1^2 = \boldsymbol{\sigma}_2^2 = 3$$

we get

$$\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2 = 2\boldsymbol{S}^2 - 3, \tag{XIII.100}$$

and squaring both sides of eq. (XIII.98) we get

$$\boldsymbol{L} \cdot \boldsymbol{S} = \tfrac{1}{2}(\boldsymbol{J}^2 - \boldsymbol{L}^2 - \boldsymbol{S}^2). \tag{XIII.101}$$

Finally, from eq. (XIII.97)

$$\begin{aligned}(\boldsymbol{S} \cdot \boldsymbol{r})^2 &= \tfrac{1}{4}[(\boldsymbol{\sigma}_1 \cdot \boldsymbol{r}) + (\boldsymbol{\sigma}_2 \cdot \boldsymbol{r})]^2 = \tfrac{1}{4}[(\boldsymbol{\sigma}_1 \cdot \boldsymbol{r})^2 + (\boldsymbol{\sigma}_2 \cdot \boldsymbol{r})^2 + 2(\boldsymbol{\sigma}_1 \cdot \boldsymbol{r})\,(\boldsymbol{\sigma}_2 \cdot \boldsymbol{r})] \\ &= \tfrac{1}{2}[(\boldsymbol{\sigma}_1 \cdot \boldsymbol{r})\,(\boldsymbol{\sigma}_2 \cdot \boldsymbol{r}) + r^2].\end{aligned}$$

Therefore

$$(\boldsymbol{\sigma}_1 \cdot \boldsymbol{r})\,(\boldsymbol{\sigma}_2 \cdot \boldsymbol{r}) = 2(\boldsymbol{S} \cdot \boldsymbol{r})^2 - r^2,$$

and consequently

$$S_{12} \equiv 3\,\frac{(\boldsymbol{\sigma}_1 \cdot \boldsymbol{r})\,(\boldsymbol{\sigma}_2 \cdot \boldsymbol{r})}{r^2} - \boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2 \tag{XIII.102}$$

$$= 2\left[3\,\frac{(\boldsymbol{S} \cdot \boldsymbol{r})^2}{r^2} - \boldsymbol{S}^2\right]. \tag{XIII.102'}$$

The operator S_{12} is called the "tensor operator" and the interaction (XIII.99*d*) the "tensor force".

If V is a linear combination of interactions of type (XIII.99), the Hamiltonian will be invariant both under rotation and under reflection (in a reflection, $\boldsymbol{r}$ and $\boldsymbol{p}$ become respectively $-\boldsymbol{r}$ and $-\boldsymbol{p}$; spin operators remain unchanged). This second invariance property will be discussed later. For the present we mention only the following. Call P the operator which operating on $\psi(\boldsymbol{r})$ gives $\psi(-\boldsymbol{r})$. The eigenfunctions of P are functions of well-defined parity. Invariance under reflection signifies that $[H, P] = 0$. If the Hamiltonian has this property we may look for its eigenfunctions among the functions of well-defined parity.

Interactions (XIII.99) are arranged in order of decreasing symmetry.

The first is independent of spin.

The second commutes with $\boldsymbol{L}$ and $\boldsymbol{S}$ separately: it is invariant not only in an overall rotation but also in a rotation of the orbital variables alone and in a rotation of the spins alone. If V contains only terms of the form (XIII.99*a*) and (XIII.99*b*), then the eigenfunctions of H may be sought among the common eigenfunctions of $\boldsymbol{L}^2$, $\boldsymbol{S}^2$, L_z, S_z and the corresponding eigenvalues will have a rotational degeneracy of order $(2L+1)\,(2S+1)$ and will be independent of the eigenvalues of L_z and S_z.

If V also contains a term of the form (XIII.99c), then H will still commute with $\boldsymbol{L}^2$ and $\boldsymbol{S}^2$ but will cease to be separately invariant under rotations of space and rotations of the spins. Its eigenfunctions may be sought among the common eigenfunctions of $\boldsymbol{L}^2$, $\boldsymbol{S}^2$, $\boldsymbol{J}^2$ and J_z and its eigenvalues will have a rotational degeneracy of order $(2J+1)$ only.

Interaction (XIII.99d) is the least symmetrical of the four. It still commutes with $\boldsymbol{S}^2$ (from (102′), $[\boldsymbol{S}^2, S_{12}]=0$), but not with $\boldsymbol{L}^2$. If V contains a term of the form (XIII.99d), the eigenfunctions of H may be sought among the common eigenfunctions of P, $\boldsymbol{S}^2$, $\boldsymbol{J}^2$, and J_z.

V. ADDITION OF ANGULAR MOMENTA

24. The Addition Problem

In many problems, the Hamiltonian is invariant under rotation and therefore commutes with the components of the total angular momentum $\boldsymbol{J}$. We then seek the eigenfunctions of the Hamiltonian among the simultaneous eigenfunctions of $\boldsymbol{J}^2$ and J_z. It is therefore important to be able to enumerate, and to form, the vectors of angular momentum $(J\ M)$.

In the simple case of a spinless particle in a central field (Ch. IX) the total angular momentum is simply the orbital angular momentum $\boldsymbol{l}$, and the eigenfunctions of total angular momentum have the form $\chi(r)\, Y_l^m(\theta, \varphi)$. In the general case, $\boldsymbol{J}$ is a sum of individual angular momenta

$$\boldsymbol{J}=\sum_i \boldsymbol{J}_i,$$

i.e., of the orbital angular momenta and of the spins of the particles constituting the system. In general we know how to construct the eigenvectors of the individual angular momenta. Thus, in the case of the two-nucleon system considered in § 13 we have

$$\boldsymbol{J}=\boldsymbol{L}+\tfrac{1}{2}\boldsymbol{\sigma}_1+\tfrac{1}{2}\boldsymbol{\sigma}_2 \qquad \text{(XIII.103)}$$

and the eigenfunctions of the individual angular momenta are of the form $\psi(r)\, Y_l^m(\theta, \varphi)\, |\mu_1\rangle\, |\mu_2\rangle$, where μ_1 and μ_2 may take the values $+\frac{1}{2}$ or $-\frac{1}{2}$ according as nucleons 1 and 2 have their spins up or down respectively. The addition problem consists in forming linear combinations of these to obtain a complete set of eigenfunctions of the total angular momentum.

25. Addition Theorem for Two Angular Momenta

The simplest problem is that of adding *two* angular momenta. Suppose that

$$\boldsymbol{J}=\boldsymbol{j}_1+\boldsymbol{j}_2,$$

where $\boldsymbol{j}_1$ and $\boldsymbol{j}_2$ are respectively the angular momenta of the separate systems 1 and 2 which together form the system under study, and suppose that we have constructed a complete set of common eigenvectors

$$|\alpha\, j_1\, j_2\, m_1\, m_2\rangle \tag{XIII.104}$$

of $\boldsymbol{j}_1^2$, $\boldsymbol{j}_2^2$, j_{1z} and j_{2z}. α represents the additional quantum numbers necessary to specify the dynamical state completely or, if one prefers, the eigenvalues of the observables A which with $\boldsymbol{j}_1^2$, $\boldsymbol{j}_2^2$, j_{1z} and j_{2z} form a complete set of commuting observables, and which commute as well with the components of $\boldsymbol{j}_1$ and of $\boldsymbol{j}_2$. In addition we suppose that the vectors (XIII.104) form a standard basis with respect to the angular momenta 1 and 2. To each set of quantum numbers $(\alpha\, j_1\, j_2)$ there correspond as many vectors as there are distinct pairs $(m_1\, m_2)$; these vectors are obtained one from another by repeated application of $j_{1\pm}$ or $j_{2\pm}$, as set forth in §6, and span a subspace $\mathscr{E}(\alpha\, j_1\, j_2)$ of $(2j_1+1)\ (2j_2+1)$ dimensions.

Note that A, $\boldsymbol{j}_1^2$ and $\boldsymbol{j}_2^2$ commute with $\boldsymbol{J}$. We therefore look for the eigenvectors of $\boldsymbol{J}^2$ and J_z among the eigenvectors of these operators and in doing so each subspace $\mathscr{E}(\alpha\, j_1\, j_2)$ may be treated separately. We consider a particular $\mathscr{E}$ and to simplify notation we represent the vectors $|\alpha\, j_1\, j_2\, m_1\, m_2\rangle$ of this subspace by $|m_1\, m_2\rangle$ and the eigenvectors of total angular momentum situated in this space by $|J\, M\rangle$ (which supposes J and M suffice to define them, which we shall show to be the case).

In this paragraph, we shall determine the possible values of the pair $(J\ M)$ and their respective orders of degeneracy. The construction of the eigenvectors will be discussed in § 27.

The solution to our problem is based on the following two observations:

(a) Each vector $|m_1\, m_2\rangle$ is an eigenvector of J_z belonging to the eigenvalue

$$M=m_1+m_2.$$

Proof: since $J_z = j_{1z} + j_{2z}$, $J_z |m_1 m_2\rangle = (m_1 + m_2) |m_1 m_2\rangle$.

(b) To each value of J there corresponds a certain number $N(J)$ of linearly independent series of $(2J+1)$ eigenvectors of total angular momentum, the vectors of a given series being obtained one from another by repeated application of J_+ or J_-, and corresponding respectively to the $(2J+1)$ possible values for M: $-J$, $-J+1$, ..., $+J$.

It follows [1]) that if $n(M)$ denotes the order of degeneracy of the eigenvalue M,

$$n(M) = \sum_{J \geqslant |M|} N(J)$$

and consequently

$$N(J) = n(J) - n(J+1). \qquad \text{(XIII.105)}$$

To obtain $N(J)$, it is therefore sufficient to determine $n(M)$ for each possible value of M. From (a) it is seen that $n(M)$ is simply the number of pairs $(m_1 m_2)$ such that

$$M = m_1 + m_2.$$

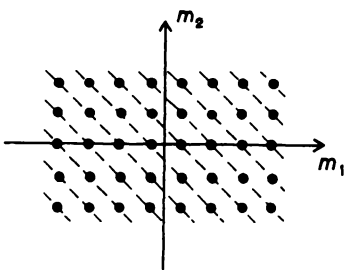


Fig. XIII.3. Finding the possible values of $M = m_1 + m_2$ and the frequency of their occurrence $n(M)$ ($j_1 = \frac{1}{2}$; $j_2 = 2$).

To find this number, it is convenient to use the diagram of Fig. XIII.3 in which each pair $(m_1 + m_2)$ is represented by a point of abscissa m_1 and ordinate m_2. $n(M)$ is the number of points situated on the diagonal $x + y = M$. In the case when $j_1 > j_2$, we find

$$n(M) = \begin{cases} 0 & \text{if } |M| > j_1 + j_2 \\ j_1 + j_2 + 1 - |M| & \text{if } j_1 + j_2 \geqslant |M| \geqslant |j_1 - j_2| \\ 2j_2 + 1 & \text{if } |j_1 - j_2| \geqslant |M| \geqslant 0. \end{cases}$$

---

[1]) The same argument has already been used in the three-dimensional harmonic oscillator problem (§ XII.15).

Substituting these numbers in (XIII.105), we find

$$N(J)=1 \text{ for } J=j_1+j_2, j_1+j_2-1, \ldots, |j_1-j_2|.$$

Whence the FUNDAMENTAL ADDITION THEOREM: *In the* $(2j_1+1)(2j_2+1)$*-dimensional space spanned by the vectors* $|\alpha\ j_1\ j_2\ m_1\ m_2\rangle$ (α, j_1, j_2 *fixed*; m_1, m_2 *variable*)

(i) *the possible values of* J *are* [1])

$$j_1+j_2, j_1+j_2-1, \ldots, |j_1-j_2|;$$

(ii) *To each of these values there corresponds one, and only one series of* $(2J+1)$ *eigenvectors* $|J\ M\rangle$ *of the total angular momentum.*

26. Applications and Examples

We first note an obvious consequence of the addition theorem: The total angular momentum resulting from the addition of any number of angular momenta is integral or half-integral according as the number of individual half-integral angular momenta is even or odd.

It will be seen that this property is verified in all the examples to follow.

As a first example, consider the *addition of two spins* $\frac{1}{2}$.

State space here has 4 dimensions. The total spin S can take two values: 0 and 1.

To $S=0$ there corresponds just one vector $|00\rangle$: we say that the spin is in the *singlet* state.

To $S=1$ there correspond the three vectors $|11\rangle$, $|10\rangle$, $|1-1\rangle$; these are the three vectors of the *triplet* state.

It is easy to write down the projectors P_0 and P_1 onto the singlet and triplet states respectively as functions of S^2 or of $\boldsymbol{\sigma}_1\cdot\boldsymbol{\sigma}_2$. Since $\mathbf{S}^2=S(S+1)$, $\mathbf{S}^2$ has the eigenvalues 0 in the singlet state and 2 in the triplet state, whence [cf. identity (XIII.100)]:

$$P_0=1-\tfrac{1}{2}\mathbf{S}^2=\tfrac{1}{4}(1-\boldsymbol{\sigma}_1\cdot\boldsymbol{\sigma}_2)$$
$$P_1=\tfrac{1}{2}\mathbf{S}^2=\tfrac{1}{4}(3+\boldsymbol{\sigma}_1\cdot\boldsymbol{\sigma}_2).$$

Note that

$$\boldsymbol{\sigma}_1\cdot\boldsymbol{\sigma}_2=P_1-3P_0. \qquad \text{(XIII.106)}$$

As a second example we consider *a particle of spin* $\frac{1}{2}$. Its orbital

[1]) In other words, J takes all values such that j_1+j_2+J is an integer with j_1, j_2 and J forming the sides of a triangle.

angular momentum l and its spin are coupled to give the total angular momentum j which can take the two values

$$j=l+\tfrac{1}{2}, \qquad j=l-\tfrac{1}{2},$$

except if $l=0$ (s-state) in which case j takes only the one value $j=\frac{1}{2}$. Thus j can take all the half-integral values from $\frac{1}{2}$ to ∞, and to each of them there correspond two terms (and two series of $(2j+1)$ vectors) of *opposite parity*.

As a final example we consider the *two-nucleon system* of § 23. Here we must couple three angular momenta – the orbital angular momentum and the two spins [eq. (XIII.98)]. We first couple the spins to give the total spin $\boldsymbol{S}$, which has the two possible values 0 and 1. $\boldsymbol{S}$ is then coupled with the angular momentum $\boldsymbol{L}$ of the relative coordinate, which may take all positive or zero integral values. To each pair of values (LS) there correspond $(2S+1)(2L+1)$ vectors that can be linearly combined to get the eigenvectors of total angular momentum. From the addition theorem, we have the following values for J:

$$\begin{array}{lll} \text{singlet state: } S=0 & J=L & \\ \text{triplet state: } S=1 & \begin{cases} J=L-1, L, L+1 & \text{if } L\neq 0 \\ J=1 & \text{if } L=0. \end{cases} \end{array}$$

We employ the following spectroscopic notation to denote the terms thus formed:

The value of L is given by a capital letter with the same convention as in § IX–6; it has a superscript on the left equal to $2S+1$ (the multiplicity of the total spin) and a subscript on the right equal to J. For example, $^3D_2$ is the term with $L=2$, triplet spin, and total angular momentum $J=2$. To each value of J there correspond 4 terms [and in all $4(2J+1)$ vectors] except for $J=0$ which only has 2. We list below the various terms corresponding to the first 4 values of J:

$J=0$		$^3P_0$		$^1S_0$
$J=1$	$^3S_1$	$^3D_1$	$^3P_1$	$^1P_1$
$J=2$	$^3P_2$	$^3F_2$	$^3D_2$	$^1D_2$
$J=3$	$^3D_3$	$^3G_3$	$^3F_3$	$^1F_3$

The same notation is currently used for the particle of spin $\frac{1}{2}$ considered above. In this case the orbital angular momentum is denoted by a small letter – capitals being reserved for the total angular momentum of a system of several particles – and the superscript on the left is simply omitted. The terms corresponding to the first four values of j are:

$j=$	$\frac{1}{2}$	$\frac{3}{2}$	$\frac{5}{2}$	$\frac{7}{2}$
	$s_{\frac{1}{2}} p_{\frac{1}{2}}$	$p_{\frac{3}{2}} d_{\frac{3}{2}}$	$d_{\frac{5}{2}} f_{\frac{5}{2}}$	$f_{\frac{7}{2}} g_{\frac{7}{2}}$

27. Eigenvectors of the Total Angular Momentum. Clebsch–Gordon Coefficients

To each pair $(J\ M)$ given by the addition theorem there corresponds an eigenvector $|\alpha\ j_1\ j_2\ J\ M\rangle$ of the total angular momentum. To define this vector unambiguously, we take it to be of norm 1, and we fix its phase by an appropriate convention to which we return below. Like the $|\alpha\ j_1\ j_2\ m_1\ m_2\rangle$, the vectors $|\alpha\ j_1\ j_2\ J\ M\rangle$ form an orthonormal basis in the subspace $\mathscr{E}(\alpha\ j_1\ j_2)$. We pass from the one basis to the other by the unitary transformation:

$$|\alpha\ j_1\ j_2\ J\ M\rangle = \sum_{m_1 m_2} |\alpha\ j_1\ j_2\ m_1\ m_2\rangle \langle \alpha\ j_1\ j_2\ m_1\ m_2\ |\alpha\ j_1\ j_2\ J\ M\rangle. \tag{XIII.107}$$

The coefficients of this transformation have an important property: *they are independent of* α, *depending only on the quantities* j_1, j_2, J, m_1, m_2, M. In the subspace $\mathscr{E}(\alpha\ j_1\ j_2)$ the $|\alpha\ j_1\ j_2\ m_1\ m_2\rangle$ are the basis vectors of a standard representation; a representation in which the components of $\boldsymbol{j}_1$ and $\boldsymbol{j}_2$ are represented by matrices independent of α [cf. eq. (XIII.28)]; it follows that the matrices representing $\boldsymbol{J}^2$ and J_z are also independent of α and that the components

$$\langle \alpha\ j_1\ j_2\ m_1\ m_2\ |\alpha\ j_1\ j_2\ J\ M\rangle$$

of their common eigenvectors have the same property. These therefore have a purely geometrical character and depend only upon the angular momenta in question and their orientation, and not upon the physical nature of the dynamical variables 1 and 2 from which the angular momenta are constructed. They are called *Clebsch–Gordon* (C.–G.) *coefficients* or vector-addition coefficients, and will be written

$\langle j_1\, j_2\, m_1\, m_2 | J\, M\rangle$. With this notation, (XIII.107) becomes

$$|\alpha\, j_1\, j_2\, J\, M\rangle = \sum_{m_1 m_2} |\alpha\, j_1\, j_2\, m_1\, m_2\rangle \langle j_1\, j_2\, m_1\, m_2 | J\, M\rangle. \qquad \text{(XIII.108)}$$

To completely define the C.–G. coefficients it remains to fix the phases of the vectors $|\alpha\, j_1\, j_2\, J\, M\rangle$. For the relative phases of the $(2J+1)$ vectors corresponding to a given J we adopt the "standard" convention of § 6. The $|\alpha\, j_1\, j_2\, J\, M\rangle$ are then defined up to a phase depending on J. This we specify by requiring that the component of $|\alpha\, j_1\, j_2\, J\, J\rangle$ along $|\alpha\, j_1\, j_2\, j_1\, J-j_1\rangle$ be real and positive

$$\langle j_1\, j_2\, j_1\, m_2 | J\, J\rangle \text{ real } \geqslant 0. \qquad \text{(XIII.109)}$$

Many properties of the C.–G. coefficients follow directly from their definition.

From the Addition Theorem, in order for $\langle j_1\, j_2\, m_1\, m_2 | J\, M\rangle$ not to be zero we must simultaneously have (*selection rules*)

$$m_1 + m_2 = M, \qquad |j_1 - j_2| \leqslant J \leqslant j_1 + j_2.$$

We show below that the C.–G. coefficients relating to a given value of J can all be obtained from the coefficient $\langle j_1\, j_2\, j_1\, J-j_1 | J\, J\rangle$ by applying recursion relations with real coefficients. Since this one is real, all of the others are real.

Moreover, since they are the coefficients of a unitary transformation, they obey the *orthogonality relations*

$$\sum_{m_1 m_2} \langle j_1\, j_2\, m_1\, m_2 | J\, M\rangle \langle j_1\, j_2\, m_1\, m_2 | J'\, M'\rangle = \delta_{JJ'}\, \delta_{MM'} \qquad \text{(XIII.110}a\text{)}$$

$$\sum_{JM} \langle j_1\, j_2\, m_1\, m_2 | J\, M\rangle \langle j_1\, j_2\, m_1'\, m_2' | J\, M\rangle = \delta_{m_1 m_1'}\, \delta_{m_2 m_2'}\,. \qquad \text{(XIII.110}b\text{)}$$

In the simplest cases we can determine the linear combinations (XIII.108) directly. For $J = j_1 + j_2$ and $M = J$, we have

$$|\alpha\, j_1\, j_2\, j_1 + j_2\, j_1 + j_2\rangle = |\alpha\, j_1\, j_2\, j_1\, j_2\rangle.$$

By repeated application of $J_- \equiv j_{1-} + j_{2-}$ to both sides of this equation we construct all the $|\alpha\, j_1\, j_2\, J\, M\rangle$ corresponding to $J = j_1 + j_2$. Following this, we construct all the vectors of the series $J = j_1 + j_2 - 1$, beginning with the one corresponding to $M = J$, which is unambiguously defined by the phase condition (XIII.109) and its property of being orthogonal to $|\alpha\, j_1\, j_2\, j_1 + j_2\, j_1 + j_2 - 1\rangle$, and forming all the others by repeated application of J_-. And so on.

In the case of the addition of two spins $\frac{1}{2}$, the eigenvectors of total spin are built up in this way from the eigenvectors $|++\rangle$, $|+-\rangle$, $|-+\rangle$ and $|--\rangle$ of the individual spins. The results are given in the following table:

	$S=1$	$S=0$
$M=1$	$\|11\rangle=\|++\rangle$	
$M=0$	$\|10\rangle=\frac{\|+-\rangle+\|-+\rangle}{\sqrt{2}}$	$\|00\rangle=\frac{\|+-\rangle-\|-+\rangle}{\sqrt{2}}$
$M=-1$	$\|1-1\rangle=\|--\rangle$.	

When larger angular momenta are to be added one must turn to more elaborate methods of calculation. Various recursion relations may be established [eq. (C.18–20)]. As an example, the application of J_+ or J_- to both sides of (XIII.108) gives [eqs. (C.19) and (C.18)]:

$$\begin{aligned}\sqrt{J(J+1)-M(M+1)}\,\langle j_1\, j_2\, m_1\, m_2\,|J\, M+1\rangle \\ =\sqrt{j_1(j_1+1)-m_1(m_1-1)}\,\langle j_1\, j_2\, m_1-1\, m_2\,|J\, M\rangle \\ +\sqrt{j_2(j_2+1)-m_2(m_2-1)}\,\langle j_1\, j_2\, m_1\, m_2-1\,|J\, M\rangle\end{aligned} \qquad \text{(XIII.111)}$$

$$\begin{aligned}\sqrt{J(J+1)-M(M-1)}\,\langle j_1\, j_2\, m_1\, m_2\,|J\, M-1\rangle \\ =\sqrt{j_1(j_1+1)-m_1(m_1+1)}\,\langle j_1\, j_2\, m_1+1\, m_2\,|J\, M\rangle \\ +\sqrt{j_2(j_2+1)-m_2(m_2+1)}\,\langle j_1\, j_2\, m_1\, m_2+1\,|J\, M\rangle.\end{aligned} \qquad \text{(XIII.112)}$$

When $M=J$ the left-hand side of (XIII.111) vanishes. It may then be used to give all of the coefficients $\langle j_1\, j_2\, m_1\, m_2\,|J\, J\rangle$ as multiples of one of them, $\langle j_1\, j_2\, j_1\, J-j_1\,|J\, J\rangle$ say. With the normalization condition for the vector $|\alpha\, j_1\, j_2\, J\, J\rangle$ ($\sum_{m_1 m_2}\langle j_1\, j_2\, m_1\, m_2\,|J\, J\rangle^2=1$) and the phase condition (XIII.109) they are completely determined. The other C.–G. coefficients may then all be derived by repeated application of the recursion relation (XIII.112). This method of calculating the C.–G. coefficients has been used by Racah to express them in the condensed form (C.21) [1]).

[1]) To carry this calculation through to the end, one must use the following identity, due to Racah:

$$\sum_s \frac{(a+s)!\,(b-s)!}{(c+s)!\,(d-s)!}=\frac{(a+b+1)!\,(a-c)!\,(b-d)!}{(c+d)!\,(a+b-c-d+1)!}$$

(a, b, c, d, integers such that $a\geqslant c\geqslant 0$, $b\geqslant d\geqslant 0$; s taking all integral values from $-c$ to $+d$).

In addition to the properties already mentioned, the C.–G. coefficients have important *symmetry properties*, which greatly facilitate tabulation. These are given, along with the principal properties of the C.–G. coefficients, in Appendix C (Section I), which also contains a table of the simpler coefficients.

28. Application: Two-Nucleon System

As an application of the addition of angular momenta to the treatment of rotationally invariant systems, we return to the two-nucleon system of § 23. We shall study the Schrödinger equation for various forms of spin dependent potentials. The discussion will be limited to potentials of type (XIII.99).

Suppose we have a potential of the form

$$V = V_1(r) + V_2(r)\,(\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2).$$

In this case the Hamiltonian commutes with $\boldsymbol{L}$ and $\boldsymbol{S}$ and the eigenfunctions are products of spin functions $|S\,\mu\rangle$ by functions of $\boldsymbol{r}$ of well-defined orbital angular momentum $(l\,m)$. From (XIII.106) we see that the potential is different according as $S=0$ or 1. The solving of the Schrödinger equation is thus equivalent to solving two Schrödinger equations for a spinless particle in a central potential, each corresponding to one of the two possible values of S. If $S=0$, the orbital part of the eigenfunction is that of a (spinless) particle in the potential $V_1 - 3V_2$; if $S=1$ that of a particle in the potential $V_1 + V_2$. The eigenvalue problem reduces to a radial equation for each pair of values $(L\,S)$.

If the potential is of the form

$$V = V_1(r) + V_2(r)\,(\boldsymbol{\sigma}_1 \cdot \boldsymbol{\sigma}_2) + V_3(r)\,(\boldsymbol{L} \cdot \boldsymbol{S}),$$

the Hamiltonian is no longer separately invariant under rotations of space and of the spins, but since it still commutes with $\boldsymbol{L}^2$ and $\boldsymbol{S}^2$ we can look for the simultaneous eigenfunctions of $\boldsymbol{L}^2$, $\boldsymbol{S}^2$, $\boldsymbol{J}^2$ and J_z. To each set $(L\,S\,J)$ there corresponds a type of function whose dependence on the angles θ, φ and on the spin variables is completely determined, and which is explicitly given with the aid of the C.–G. coefficients as follows:

$$\begin{aligned} \Psi^M_{LSJ} &\equiv F(r)\,\mathcal{Y}^M_{LSJ} \\ \mathcal{Y}^M_{L\,l} &= \sum Y_L{}^m(\theta, \varphi)\,|S\,\mu\rangle\,\langle L\,S\,m\,\mu\,|J\,M\rangle. \end{aligned} \tag{XIII.113}$$

Thus the three functions of the state $^1P_1$ have the "angular dependence" $\mathcal{Y}^M_{101} = Y_1^M |00\rangle$ $(M = 0, \pm 1)$. The five functions of the state $^3D_2$:

$$\mathcal{Y}^M_{212} = \sum_{m\mu} Y_2^m |1\,\mu\rangle \langle 21\,m\,\mu\,|2\,M\rangle \qquad (M = 0, \pm 1, \pm 2).$$

Following (XIII.100) and (XIII.101), the application of the Hamiltonian to a function of this type gives

$$H\Psi^M_{LSJ} = \left[-\frac{\hbar^2}{M_0}\frac{1}{r}\frac{\mathrm{d}^2}{\mathrm{d}r^2}r + \frac{\hbar^2}{M_0}\frac{L(L+1)}{r^2} + V_{LSJ}\right]\Psi^M_{LSJ},$$

with

$$V_{LSJ}(r) = V_1(r) + [2S(S+1) - 3]\,V_2(r) + \tfrac{1}{2}[J(J+1) - L(L+1) - S(S+1)]\,V_3(r).$$

The problem of solving the Schrödinger equation thus reduces to solving the radial equation

$$\left[-\frac{\hbar^2}{M_0}\frac{1}{r}\frac{\mathrm{d}^2}{\mathrm{d}r^2}r + \frac{\hbar^2}{M_0}\frac{L(L+1)}{r^2} + V_{LSJ}(r)\right] F(r) = E\,F(r).$$

Thus we have a problem resembling that of a (spinless) particle in a central potential, with the sole difference that the "effective central potential" $V_{LSJ}(r)$ differs from one triplet $(L\,S\,J)$ to another [1]).

As a final example we consider a potential of the form

$$V = V_C(r) + V_T(r)\,S_{12}.$$

Due to the presence of the "tensor" force, the Hamiltonian no longer commutes with $\boldsymbol{L}^2$, but it still commutes with $\boldsymbol{S}^2$, and also with the "parity" operator P introduced in § 23. Consequently we can look for the eigenfunctions of H among the simultaneous eigenfunctions of P, $\boldsymbol{S}^2$, $\boldsymbol{J}^2$, J_z; that is, among the functions of well-defined total angular momentum $(J\,M)$ having a well-defined parity and a well-defined value of S.

---

[1]) For $S = 0$, the spin-orbit force vanishes and we have, for any L and J,

$$V_{LOL} = V_1 - 3V_2.$$

Similarly, if $S = 1$ and $L = J$ the "effective potential" does not depend on L:

$$V_{L1L} = V_1 + V_2 - V_3.$$

If $S=0$, then necessarily $L=J$ (therefore $P=(-)^J$) and the eigenfunction is necessarily of the form $F(r)\,\mathscr{Y}^M_{J0J}$. Again, since $\boldsymbol{S}\,|00\rangle=0$, from (XIII.102′)

$$S_{12}\,\mathscr{Y}^M_{J0J} \equiv S_{12}\,Y^M_J(\theta,\varphi)\,|00\rangle=0.$$

Consequently, $F(r)$ satisfies the radial equation for a particle of angular momentum J in the potential $V_C(r)$.

If $S=1$ and $P=(-)^J$, necessarily $L=J$ and the "angular dependence" of the eigenfunction is, as before, completely defined:

$$\Psi^M_{J1J}=F(r)\,\mathscr{Y}^M_{J1J}.$$

One can show (Problem XIII.11) that $S_{12}\,\mathscr{Y}^M_{J1J}=2\mathscr{Y}^M_{J1J}$. Thus $F(r)$ satisfies the radial equation for a particle of angular momentum J in the potential $V_C(r)+2V_T(r)$.

If $S=1$ and $P=(-)^{J+1}$, the only possible values of L are $J+1$ and $J-1$ (unless $J=0$ in which case there is just the one value $L=1$) so that the eigenfunction is of the form

$$\Psi \equiv F_{J-1}(r)\,\mathscr{Y}_- + F_{J+1}(r)\,\mathscr{Y}_+,$$

where, to simplify the writing, we have put $\mathscr{Y}_\pm \equiv \mathscr{Y}^M_{J\pm1J}$. Now S_{12} acting on $\mathscr{Y}_+$ or on $\mathscr{Y}_-$ gives a combination of these two functions (Problem 11); hence $(H-E)\,\Psi$ is also a linear combination of these two functions, the coefficients being functions of r. Writing $(H-E)\,\Psi=0$ amounts to making these two coefficients vanish and gives a system of two second-order differential equations in $F_{J-1}(r)$ and $F_{J+1}(r)$.

As an example we write the system of coupled radial equations for $J=1$. This case is encountered in the study of the deuteron. The wave function is a mixture of the states $^3S_1$ and $^3D_1$, and may be put in the form

$$\Psi \equiv \frac{1}{r}u_S(r)\,\mathscr{Y}^M_{011}+\frac{1}{r}u_D(r)\,\mathscr{Y}^M_{211}.$$

Since (Problem XIII.11)

$$\begin{aligned} &S_{12}\,\mathscr{Y}^M_{011}=\sqrt{8}\,\mathscr{Y}_{211} \qquad && S_{12}\,\mathscr{Y}^M_{211}=\sqrt{8}\,\mathscr{Y}^M_{011}-2\mathscr{Y}^M_{211},\\ &\boldsymbol{L}^2\,\mathscr{Y}^M_{011}=0 && \boldsymbol{L}^2\,\mathscr{Y}^M_{211}=6\mathscr{Y}^M_{211} \end{aligned} \qquad \text{(XIII.114)}$$

and since

$$H \equiv -\frac{\hbar^2}{M_0}\frac{1}{r}\frac{\mathrm{d}^2}{\mathrm{d}r^2}r+\frac{\hbar^2\boldsymbol{L}^2}{M_0r^2}+V_C(r)+V_T(r)\,S_{12},$$

the equation $(H-E)\,\Psi=0$ is equivalent to

$$\left[\frac{\hbar^2}{M_0}\frac{\mathrm{d}^2}{\mathrm{d}r^2}+E-V_C(r)\right]u_S=\sqrt{8}\,V_T(r)\,u_D$$
$$\left[\frac{\hbar^2}{M_0}\left(\frac{\mathrm{d}^2}{\mathrm{d}r^2}-\frac{6}{r^2}\right)+E+2\,V_T(r)-V_C(r)\right]u_D=\sqrt{8}\,V_T(r)\,u_S. \tag{XIII.115}$$

29. Addition of Three or More Angular Momenta. Racah Coefficients. "3*sj*" Symbols

The two-nucleon system studied in § 28 provides an example of a system where the total angular momentum is the sum of three individual angular momenta [eq. (XIII.103)]. We were able to study this specially simple example without having recourse to elaborate techniques. We now examine the addition of three angular momenta in the general case.

We suppose our system to be made up of three distinct systems 1, 2, and 3 with angular momenta $\boldsymbol{j}_1$, $\boldsymbol{j}_2$, and $\boldsymbol{j}_3$ respectively. The total angular momentum is then

$$\boldsymbol{J}=\boldsymbol{j}_1+\boldsymbol{j}_2+\boldsymbol{j}_3.$$

The addition problem consists in forming the eigenvectors of the total angular momentum in the space spanned by the $(2j_1+1)\,(2j_2+1)\,(2j_3+1)$ eigenvectors

$$|\alpha\; j_1\; j_2\; j_3\; m_1\; m_2\; m_3\rangle$$

of individual angular momenta corresponding to well-defined values of the quantum numbers α, j_1, j_2 and j_3. α, defined here as in § 25, plays no role in what follows and will simply be omitted. There are several ways of constructing the vectors of angular momentum $(J\;M)$.

(i) We can couple $\boldsymbol{j}_1$ and $\boldsymbol{j}_2$ (Fig. XIII.4*a*) to form the angular momentum $\boldsymbol{J}_{12}=\boldsymbol{j}_1+\boldsymbol{j}_2$, and then couple $\boldsymbol{J}_{12}$ and $\boldsymbol{j}_3$ to form $\boldsymbol{J}$. We obtain in this way the eigenvectors

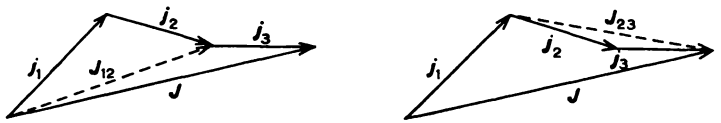
$$|(j_1\; j_2)\,J_{12},\; j_3;\; J\;M\rangle$$
$$=\sum_{\substack{m_1\;m_2\\ M_{12}m_3}}|j_1\; j_2\; j_3\; m_1\; m_2\; m_3\rangle\,\langle j_1\; j_2\; m_1\; m_2\,|J_{12}\;M_{12}\rangle\,\langle J_{12}\; j_3\; M_{12}\; m_3\,|J\;M\rangle \tag{XIII.116}$$

common to $\boldsymbol{j}_1^2$, $\boldsymbol{j}_2^2$, $\boldsymbol{j}_3^2$, $\boldsymbol{J}_{12}^2$, $\boldsymbol{J}^2$ and J_z;

(ii) We can couple j_2 and j_3 (Fig. XIII.4*b*) to form the angular momentum $J_{23}=j_2+j_3$, and then couple j_1 and J_{23} to form J. We obtain in this way the eigenvectors

$$|j_1, (j_2\, j_3)\, J_{23}; J\, M\rangle$$
$$= \sum_{\substack{m_2\, m_3 \\ m_1\, M_{23}}} |j_1\, j_2\, j_3\, m_1\, m_2\, m_3\rangle \langle j_2\, j_3\, m_2\, m_3\, |J_{23}\, M_{23}\rangle \langle j_1\, J_{23}\, m_1\, M_{23}\, |J\, M\rangle \tag{XIII.117}$$

common to j_1^2, j_2^2, j_3^2, J_{23}^2, J^2 and J_z;



a) $j_1+j_2=J_{12}$, $J_{12}+j_3=J$ *b)* $j_2+j_3=J_{23}$, $j_1+J_{23}=J$

Fig. XIII.4. Different ways of coupling three angular momenta.

(iii) We can couple j_1 and j_3 to form J_{13} and then J_{13} and j_2 to form J.

Thus we have the choice between three different sets of basis vectors for the total angular momentum. In many problems it is useful to be able to pass from one set to another. The transformation involved is a unitary transformation. We have for example

$$|j_1, (j_2\, j_3)\, J_{23}; J\, M\rangle$$
$$= \sum_{J_{12}} |\, (j_1\, j_2)\, J_{12}, j_3; J\, M\rangle \langle (j_1\, j_2)\, J_{12}, j_3\, J\, |j_1, (j_2\, j_3)\, J_{23}\, J\rangle. \tag{XIII.118}$$

Obviously the coefficients of this unitary transformation are independent of α, for the same reason that the C.–G. coefficients are independent of α. Applying J_+ or J_- to both sides of (XIII.118) it is easily seen that they are also independent of M, and so depend only on the six angular momenta j_1, j_2, j_3, J_{12}, J_{23} and J.

Rather than directly make use of these coefficients it is more convenient to use either the Racah W coefficients or the Wigner "6j" symbols, which are defined as follows:

$$\langle (j_1\, j_2) J_{12}, j_3 J\, |j_1, (j_2\, j_3) J_{23} J\rangle = \sqrt{(2J_{12}+1)\,(2J_{23}+1)}\; W\,(j_1\, j_2\, J\, j_3; J_{12}\, J_{23})$$
$$= (-)^{j_1+j_2+j_3+J} \sqrt{(2J_{12}+1)(2J_{23}+1)} \begin{Bmatrix} j_1 j_2 J_{12} \\ j_3 J\, J_{23} \end{Bmatrix}.$$

It is clear, just from their definition, that the W coefficients are simply the sum over the "m" indices of four C.–G. coefficients. Except for the simplest arguments, the direct calculation of these coefficients is very tedious: it consists in calculating first a great many C.–G. coefficients, and then a complicated expression built up from these coefficients. Racah has succeeded in obtaining a workable expression for W [formula (C.36)]. There exist tables of the W coefficients for the most common arguments.

The "$6j$" symbols differ from the W only through their sign. They are interesting chiefly because of their important symmetry properties.

The main properties of the W and of the "$6j$" are assembled in Appendix C (Section II).

The treatment given above for the case of three angular momenta can be generalized to the addition of a greater number n of angular momenta:

$$\boldsymbol{J}=\boldsymbol{j}_1+\boldsymbol{j}_2+\ldots+\boldsymbol{j}_n. \tag{XIII.119}$$

By coupling two of the individual angular momenta: $\boldsymbol{j}_i+\boldsymbol{j}_k=\boldsymbol{J}_{ik}$, we reduce the problem to the addition of $(n-1)$ angular momenta, the vectors $\boldsymbol{j}_i$ and $\boldsymbol{j}_k$ being replaced in the right-hand side of (XIII.119) by their sum $\boldsymbol{J}_{ik}$. By repeating this operation we reduce the problem to the addition of $(n-2)$ angular momenta, and so on. We thus achieve the addition of the n angular momenta through the introduction of all $(n-2)$ intermediate angular momenta, and form in this way a set of basis vectors for the total angular momentum.

In this fashion one can form several distinct sets of basis vectors. We have seen that when $n=3$ there are three of them. It can be shown that in the general case there are $\frac{1}{2}n!$. One passes from one set to another by a unitary transformation (with real coefficients). One can easily convince oneself that the coefficients of that transformation depend neither upon α nor upon the eigenvalue M of the component J_z of $\boldsymbol{J}$. Thus they depend only on the quantum numbers $J, j_1, j_2, \ldots, j_n$ and on the two series of $(n-2)$ quantum numbers such as J_{ik} giving the length of the intermediate angular momenta characterising each of the sets of basis vectors: in all, $1+n+2(n-2)=3(n-1)$ quantum numbers "j". The coefficients may be put in the form of "$3(n-1)\ j$" symbols, generalizations of the "$6j$" symbols introduced in the addition of three angular momenta. The "$3(n-1)\ j$" symbols are sums over

the "m" indices of $2(n-1)$ C.–G. coefficients. The principal properties of the "$9j$" symbols (symbols for the addition of four angular momenta) are given in Appendix C (Section III).

VI. IRREDUCIBLE TENSOR OPERATORS [1])

30. Representation of Scalar Operators

If an observable is invariant under rotation, the subspace of each of its eigenvalues is invariant under rotation. This important property has already been mentioned in § 17. The observable treated there was the Hamiltonian, but clearly the property in question holds for any scalar observable.

More generally, even if not diagonal, a scalar observable is represented in a given standard representation by a particularly simple matrix, as will now be shown.

Let $|\tau\, J\, M\rangle$ be the basis vectors of a standard representation $\{J^2\, J_z\}$ (notation of § 6) and let S be a scalar operator (not necessarily an observable). By hypothesis

$$[\boldsymbol{J}, S] = 0.$$

It follows that the vector $S\,|\tau'\, J'\, M'\rangle$, like $|\tau'\, J'\, M'\rangle$, is a vector of angular momentum $(J'\, M')$ orthogonal to any vector of different angular momentum. Therefore, the matrix element $\langle\tau\, J\, M\,|S\,|\tau'\, J'\, M'\rangle$ is zero if $J \neq J'$ or $M \neq M'$. Furthermore, since J_+ commutes with S, when $J = J'$ and $M = M'$ we have

$$\begin{aligned}
&\langle\tau\, J\, M\,|S\,|\tau'\, J\, M\rangle \\
&\quad = [J\,(J+1) - M\,(M-1)]^{-\frac{1}{2}}\,\langle\tau\, J\, M\,|S\, J_+\,|\tau'\, J\, M-1\rangle \\
&\quad = [J\,(J+1) - M\,(M-1)]^{-\frac{1}{2}}\,\langle\tau\, J\, M\,|J_+ S\,|\tau'\, J\, M-1\rangle \\
&\quad = \langle\tau\, J\, M-1\,|S\,|\tau'\, J\, M-1\rangle.
\end{aligned}$$

This shows that the matrix element is independent of M. The two properties are summarized in the relation

$$\langle\tau\, J\, M\,|S\,|\tau'\, J'\, M'\rangle = \delta_{JJ'}\,\delta_{MM'}\,S^{(J)}_{\tau\tau'}, \qquad \text{(XIII.120)}$$

---

[1]) For a systematic account of the algebra of irreducible tensors and its applications in the theory of angular momentum in Quantum Mechanics, see U. Fano and G. Racah, *Irreducible Tensorial Sets* (Academic Press Inc., New York, 1959).

where $S^{(J)}_{\tau\tau'}$ is a quantity depending only on J, τ and τ'. In the case when S is an observable, the matrix $S^{(J)}_{\tau\tau'}$ is Hermitean and can be diagonalized.

31. Irreducible Tensor Operators. Definition

We now generalize (XIII.120) to a class of operators, called irreducible tensor operators, which are not rotationally invariant, but nevertheless have specially simple laws of transformation under rotation.

The concept of a tensor operator is a generalization of the concept of a vector operator.

We begin with the definition of a *tensor*. Suppose we have an n-dimensional space $\mathscr{E}_n$ such that in a rotation the vectors of $\mathscr{E}_n$ transform linearly into vectors of $\mathscr{E}_n$: with each rotation is associated a linear operator of $\mathscr{E}_n$. By definition, the vectors of $\mathscr{E}_n$ are n-component tensors. Thus the spinors and the vectors of ordinary space are respectively 2- and 3-component tensors; vectors of a subspace $\mathscr{E}(\tau J)$ of the type defined in § 6 are $(2J+1)$-component tensors, and kets of the ket-vector space of a quantum system are tensors of an infinite number of components.

If we choose a set of basis vectors in $\mathscr{E}_n$, each of the above defined tensors is represented by its n components, and the rotation is represented by the action of an $n \times n$ matrix on these n components. Thus the rotation of a vector of ordinary space, represented by its Cartesian coordinates, is given by the matrix $\mathscr{R}$ defined in § 10. Similarly, if we set up a standard representation in $\mathscr{E}(\tau J)$, any tensor $|u\rangle$ of $\mathscr{E}(\tau J)$ is represented by its $(2J+1)$ components $u_M \equiv \langle \tau J M | u \rangle$ and the components u'_M of its transform in the rotation $\mathscr{R}(\alpha\beta\gamma)$ are obtained by application of the rotation matrix $R^{(J)}(\alpha\beta\gamma)$ (§ 16):

$$u'_M = \sum_{M'} R^{(J)}_{MM'}(\alpha\beta\gamma)\, u_{M'}. \qquad \text{(XIII.121)}$$

As a further example we consider the nine quantities $V_i W_j$ $(i, j = 1, 2, 3)$ obtained by multiplying each component of the vector $\mathbf{V}$ by each component of the vector $\mathbf{W}$. These are the nine components of a tensor which we denote by $\mathbf{V} \otimes \mathbf{W}$. The components of its transform in a rotation $\mathscr{R}$ are given by

$$\begin{aligned}[\mathbf{V} \otimes \mathbf{W}]'_{ij} \equiv V_i'\, W_j' &= \mathscr{R}_{ik}\, \mathscr{R}_{jl}\, V_k\, V_l \\ &= \mathscr{R}_{ik}\, \mathscr{R}_{jl}\, [\mathbf{V} \otimes \mathbf{W}]_{kl}.\end{aligned}$$

Among the large variety of tensors that can be formed, the *irreducible tensors* enjoy a privileged position. By definition, a tensor

is irreducible, if the space $\mathscr{E}_n$ in which it is defined is irreducible with respect to rotations.

The vectors of ordinary space, the spinors, the vectors of a space $\mathscr{E}(\tau J)$, are irreducible tensors.

On the other hand, the tensor $\boldsymbol{V} \otimes \boldsymbol{W}$ is reducible. The nine-dimensional space in which it is defined is the direct sum of three irreducible invariant subspaces (with respect to rotations), having respectively 1, 3 and 5 dimensions. The projections of $\boldsymbol{V} \otimes \boldsymbol{W}$ onto each of these subspaces are therefore irreducible tensors; they are, to within a constant, the scalar product $\boldsymbol{V} \cdot \boldsymbol{W}$, the vector $\boldsymbol{V} \times \boldsymbol{W}$, and an irreducible 5-component tensor which transforms under rotation like the harmonic polynomials of second degree (cf. § B.10)[1]).

Similarly, the vectors of the $(2j_1+1)(2j_2+1)$-dimensional space of § 25 are reducible $(2j_1+1)(2j_2+1)$-component tensors whose decomposition into irreducible parts is given by the Addition Theorem.

One passes from the concept of a tensor to that of a tensor operator in the same way as one passes from the concept of a vector to that of a vector operator.

If n operators transform linearly one into another in a rotation like n linearly independent vectors of a space $\mathscr{E}_n$, then they are the components of an n-dimensional *tensor operator* [2]). A linear transformation of these n components will furnish n new operators, which may be regarded as the components of the same tensor operator in another representation. If the space $\mathscr{E}_n$ is irreducible the *tensor operator* is said to be *irreducible*.

Vector operators constitute a special class of irreducible tensor operators.

If $\boldsymbol{V}$ and $\boldsymbol{W}$ are two vector operators, the nine operators $V_i W_j$ are the components of a reducible tensor operator which is the direct sum of three irreducible tensor operators, the scalar $\boldsymbol{V} \cdot \boldsymbol{W}$, the vector $\boldsymbol{V} \times \boldsymbol{W}$ and the tensor operator defined, for example, by the five components given in note 1 below.

---

[1]) The rotation matrices obviously depend upon the representation chosen for this tensor. In the representation where its components are:

$$\tfrac{1}{2}(V_1 W_2 + V_2 W_1), \quad \tfrac{1}{2}(V_2 W_3 + V_3 W_2), \quad \tfrac{1}{2}(V_3 W_1 + V_1 W_3),$$
$$V_1 W_1 - V_2 W_2, \quad 2V_3 W_3 - V_1 W_1 - V_2 W_2.$$

These transform among themselves like the linearly independent polynomials

$$xy, \quad yz, \quad zx, \quad x^2 - y^2, \quad 2z^2 - x^2 - y^2.$$

[2]) Thus this transformation law is not the same as that for the *components of a vector* of $\mathscr{E}_n$ along its n basis vectors. In the same way the transformation law (XIII.54) for the components of a vector operator $\boldsymbol{K}$ is not the same as the transformation law (XIII.43) for the components of a vector $\boldsymbol{V}$ along the same set of basis vectors. Note in particular the difference between the law (XIII.122) below and the law (XIII.121).

A tensor operator is unambiguously defined by the law of transformation under rotation of its components in a given representation:

By definition, the $(2k+1)$ *operators* $T_q^{(k)}$ $(q=-k,-k+1,\ldots,+k)$ *are the standard components of an irreducible* k^{th}*-order tensor operator,* $\mathsf{T}^{(k)}$*, if they transform in a rotation according to the law*

$$RT_q^{(k)}R^{-1}=\sum_{q'}T_{q'}^{(k)}R_{q'q}^{(k)}. \qquad \text{(XIII.122)}$$

This transformation law is the same as the one for the basis vectors $|k\,q\rangle$ of a standard representation for a $(2k+1)$-dimensional space irreducible with respect to rotations:

$$R\,|k\,q\rangle=\sum_{q'}|k\,q'\rangle\,R_{q'q}^{(k)}.$$

If (XIII.122) is satisfied for any infinitesimal rotation it will be satisfied for any rotation whatever. For an infinitesimal rotation, R is given by (XIII.55), the matrices $R^{(k)}$ are easily obtained from their definition (XIII.64) and the law (XIII.122) is equivalent to the following commutation relations for the $T_q^{(k)}$ with the components of total angular momentum

$$[J_\pm, T_q^{(k)}]=\sqrt{k\,(k+1)-q\,(q\pm 1)}\,T_{q\pm 1}^{(k)} \qquad \text{(XIII.123}a\text{)}$$

$$[J_z, T_q^{(k)}]=q\,T_q^{(k)}. \qquad \text{(XIII.123}b\text{)}$$

Relations (XIII.123), which may be compared with (XIII.23–25), provide another definition of the irreducible tensor operator $\mathsf{T}^{(k)}$ (strictly equivalent to the one given above).

If the $T_q^{(k)}$ represent physical quantities, they are necessarily invariant in a rotation through 2π (cf. § 15), and k *is therefore an integer*. In what follows we shall consider only irreducible tensor operators of integral order.

It can easily be shown that the $(2k+1)$ operators

$$S_q^{(k)}\equiv(-)^q\,T_{-q}^{(k)\dagger}$$

satisfy relations (XIII.123) (Problem XIII.16), and are therefore the standard components of an irreducible tensor operator of order k, $\mathsf{S}^{(k)}$. By definition, $\mathsf{S}^{(k)}$ and $\mathsf{T}^{(k)}$ are *Hermitean conjugates one of the other*:

$$\mathsf{S}^{(k)}=\mathsf{T}^{(k)\dagger}$$

(since k is an integer it is clear that Hermitean conjugation is a reciprocal operation).

Scalars are irreducible tensor operators of the zeroth order. Vector operators are irreducible tensor operators of order 1: if K_x, K_y, K_z are the Cartesian components of a vector operator, its standard components are

$$K_1^{(1)} = -\frac{1}{\sqrt{2}}(K_x + iK_y), \quad K_0^{(1)} = K_z, \quad K_{-1}^{(1)} = \frac{1}{\sqrt{2}}(K_x - iK_y). \quad \text{(XIII.124)}$$

[It will be observed that the coefficients here differ from those of (XIII.94).]

The $(2k+1)$ spherical harmonics $Y_k^q(\theta, \varphi)$ $(q = -k, \ldots, +k)$ considered as operators, are the standard components of an irreducible tensor operator of order k, $\mathsf{Y}^{(k)}$.

32. Representation of Irreducible Tensor Operators. Wigner-Eckart Theorem

The most important property of irreducible tensor operators is given by the WIGNER–ECKART THEOREM:

In a standard representation $\{J^2 J_z\}$ *whose basis vectors are denoted by* $|\tau J M\rangle$, *the matrix element* $\langle \tau J M | T_q^{(k)} | \tau' J' M' \rangle$ *of the* q^{th} *standard component of a given* k^{th}*-order irreducible tensor operator,* $\mathsf{T}^{(k)}$, *is equal to the product of the Clebsch–Gordon coefficient*

$$\langle J' k M' q | J M \rangle$$

by a quantity independent of M, M' *and* q.

Thus:

$$\langle \tau J M | T_q^{(k)} | \tau' J' M' \rangle = \frac{1}{\sqrt{2J+1}} \langle \tau J \| \mathsf{T}^{(k)} \| \tau' J' \rangle \langle J' k M' q | J M \rangle, \quad \text{(XIII.125)}$$

where $\langle \tau J \| \mathsf{T}^{(k)} \| \tau' J' \rangle$ is a quantity called the *reduced matrix element*, which varies from one tensor operator to another and depends on the indices τ, J and τ', J' (the factor $1/\sqrt{2J+1}$ is introduced simply for convenience).

To prove this theorem we consider the $(2k+1)(2J'+1)$ vectors

$$T_q^{(k)} | \tau' J' M' \rangle \qquad (q = -k, \ldots, +k; \; M' = -J', \ldots, +J'),$$

and the following linear combinations of these vectors:

$$|\sigma J'' M''\rangle = \sum_{M'q} T_q^{(k)} |\tau' J' M'\rangle \langle J' k M' q | J'' M''\rangle.$$

Application of the orthogonality relations for the C.–G. coefficients [eq. (XIII.110*b*)] gives

$$T_q^{(k)} |\tau' J' M'\rangle = \sum_{J''M''} |\sigma J'' M''\rangle \langle J' k M' q | J'' M''\rangle. \qquad \text{(XIII.126)}$$

Note that the vectors $T_q^{(k)} |\tau' J' M'\rangle$ are not necessarily linearly independent, and that therefore certain of the $|\sigma J'' M''\rangle$ may vanish.

From (XIII.123*a*) and (XIII.124):

$$\begin{aligned} J_+ T_q^{(k)} |\tau' J' M'\rangle &= [J_+, T_q^{(k)}] |\tau' J' M'\rangle + T_q^{(k)} J_+ |\tau' J' M'\rangle \\ &= \sqrt{k(k+1) - q(q+1)}\, T_{q+1}^{(k)} |\tau' J' M'\rangle \\ &\quad + \sqrt{J'(J'+1) - M'(M'+1)}\, T_q^{(k)} |\tau' J' M'+1\rangle, \end{aligned}$$

and consequently

$$\begin{aligned} J_+ |\sigma J'' M''\rangle &= \sum_{M'q} T_q^{(k)} |\tau' J' M'\rangle \{\sqrt{k(k+1) - q(q-1)} \langle J' k M' q-1 | J'' M''\rangle \\ &\quad + \sqrt{J'(J'+1) - M'(M'-1)} \langle J' k M'-1 q | J'' M''\rangle\}. \end{aligned}$$

From the recursion relation (XIII.111) for the Clebsch–Gordon coefficients, the bracket on the right-hand side is equal to

$$\sqrt{J''(J''+1) - M''(M''+1)} \langle J' k M' q | J'' M''+1\rangle,$$

so that we have the vector $|\sigma J'' M''+1\rangle$ appearing there:

$$J_+ |\sigma J'' M''\rangle = \sqrt{J''(J''+1) - M''(M''+1)}\, |\sigma J'' M''+1\rangle.$$

We can show in the same way that

$$\begin{aligned} J_- |\sigma J'' M''\rangle &= \sqrt{J''(J''+1) - M''(M''-1)}\, |\sigma J'' M-1\rangle \\ J_z |\sigma J'' M''\rangle &= M'' |\sigma J'' M''\rangle. \end{aligned}$$

From these three relations, we can conclude that the $(2J''+1)$ vectors $|\sigma J'' M''\rangle$ corresponding to the same value of J'',

(i) are either all zero, or

(ii) are (un-normalized) eigenfunctions of angular momentum $(J'' M'')$ obtained one from another in accordance with the standard convention.

Therefore, the scalar products $\langle \tau J M | \sigma J'' M'' \rangle$ are all zero, except those, if they exist, for which $J''=J$ and $M''=M$, i.e. the $(2J+1)$ products $\langle \tau J M | \sigma J M \rangle$, and these are independent of M.

The theorem then follows, for by (XIII.126) the matrix element $\langle \tau J M | T_q^{(k)} | \tau' J' M' \rangle$ is given by

$$\langle \tau J M | T_q^{(k)} | \tau' J' M' \rangle = \sum_{J'' M''} \langle \tau J M | \sigma J'' M'' \rangle \langle J' k M' q | J'' M'' \rangle.$$

Q.E.D.

Among the most important consequences of the Wigner–Eckart theorem are the *selection rules* for the operator $T_q^{(k)}$.

For the matrix element $\langle \tau J M | T_q^{(k)} | \tau' J' M' \rangle$ *to be non-zero we must simultaneously have*:

$$q = M - M' \qquad \text{(XIII.127)}$$

$$|J - J'| \leqslant k \leqslant J + J'. \qquad \text{(XIII.128)}$$

These two rules are an immediate consequence of the presence of the Clebsch–Gordon coefficient in the right-hand side of (XIII.125). In practice, it is the second of these rules that is the most useful. It is usually expressed in the form of the following corollary [1]).

If $A^{(k)}$ *is a component of a* k^{th}*-order irreducible tensor operator, the matrix element of* $A^{(k)}$ *between two vectors of angular momentum* J, J' *are necessarily zero if* k *violates the inequalities*: $|J - J'| \leqslant k \leqslant J + J'$.

33. Applications

The Wigner–Eckart theorem has many applications in atomic and nuclear physics, notably in β-decay, in electromagnetic radiation and in angular correlation problems in general.

Let us take the emission of electromagnetic radiation from a nucleus (γ-emission) as an example. Suppose that in going from the excited state $\mathcal{N}^*$ to its ground state $\mathcal{N}$, an atom emits a γ-ray:

$$\mathcal{N}^* \rightarrow \mathcal{N} + \gamma.$$

Let J and J' be the spins (i.e. total angular momenta) of the nuclei $\mathcal{N}$ and $\mathcal{N}^*$ respectively. In the theory of γ-emission it is shown that

[1]) The component $A^{(k)}$ appearing in this corollary is not necessarily a standard component; any other, i.e. any linear combination of standard components, has the same property.

the probability *amplitude* for emission of a γ-ray of polarization ν in the direction $\Omega = (\theta, \varphi)$ is proportional to the matrix element

$$\langle \tau\, J\, M \,|H\,(\Omega, \nu)\,|\tau'\, J'\, M'\rangle$$

of a certain operator $H\,(\Omega, \nu)$ between the state vectors for the initial and final states [1]). $H\,(\Omega, \nu)$ may be expanded into spherical harmonics. The details will not be gone into here [2]); it takes the form of a series of irreducible tensor operators of two types (of opposite parities): the electric and magnetic multipole moments. The electric 2^l-pole moment $\mathsf{Q}^{(l)}$ is an irreducible tensor operator of order l and of parity $(-)^l$; the magnetic 2^l-pole moment $\mathsf{M}^{(l)}$ is an irreducible tensor operator of order l and of parity $(-)^{l+1}$. Among the more familiar multipole moments we have:

(i) the magnetic moment (in the ordinary sense of the term) which is the magnetic dipole moment $\mathsf{M}^{(1)}$;

(ii) the quadrupole moment (in the ordinary sense of the term), which is the electric quadrupole moment $\mathsf{Q}^{(2)}$.

In accordance with the selection rules for tensor operators, the only non-zero contributions are from moments with multipolarity within the limits

$$|J - J'| \leqslant l \leqslant J + J' \qquad \text{(XIII.129)}$$

(there is also a parity selection rule that we do not consider here). Also, the contributions of the components $Q_m{}^{(l)}$ of a moment satisfying (XIII.129) are proportional, by the Wigner–Eckart theorem, to the Clebsch–Gordon coefficient $\langle J'\, l\, M'\, m\,|J\, M\rangle$; to obtain them we need only to know the proportionality coefficient, i.e. the reduced matrix element $\langle \tau\, J\,\|\mathsf{Q}^{(l)}\|\tau'\, J'\rangle$.

Thus we have the transition probability once the reduced matrix elements for the multipole moments obeying the selection rules are known. In practice the multipole expansion converges rapidly and significant contributions come only from the one or two multipoles of lowest order.

The even multipole moments ($\mathsf{M}^{(1)}$, $\mathsf{Q}^{(2)}$, ...) are also involved where one calculates the level shifts in an atom or nucleus in a static

---

[1]) Cf. § XXI.31.

[2]) Cf. J. Blatt and V. Weisskopf, *Theoretical Nuclear Physics* (Wiley, New York, 1952), Ch. XII and Appendix B.

electromagnetic field. Thus, the coupling of an atom with a constant magnetic field permits a measurement of its magnetic moment, and its coupling with an inhomogeneous electric field permits a measurement of its quadrupole moment. Actually what one measures is the average value of these operators for the state of the nucleus considered, that is, the matrix elements

$$\langle \tau\, J\, M\, |M_m^{(1)}\, |\tau\, J\, M'\rangle, \qquad \langle \tau\, J\, M\, |Q_m^{(2)}\, |\tau\, J\, M'\rangle, \text{ etc.}$$

or, what amounts to the same, the reduced diagonal matrix elements

$$\langle \tau\, J\, \|\mathsf{M}^{(1)}\, \|\tau\, J\rangle, \qquad \langle \tau\, J\, \|\mathsf{Q}^{(2)}\, \|\tau\, J\rangle, \text{ etc.}$$

Note that the magnetic moment necessarily vanishes if $J=0$, the quadrupole moment if $J=0$ or $\frac{1}{2}$. More generally, the 2^l-pole moment of a nucleus of spin J necessarily vanishes if $2J<l$.

EXERCISES AND PROBLEMS

1. Starting from relations (XIII.24) and (XIII.25) between the $(2j+1)$ vectors

$$|\tau\, j\, \mu\rangle \qquad (\mu=-j, -j+1, \ldots, +j),$$

prove relations (XIII.26) and (XIII.27).

2. Show that in any representation where J_x and J_z are real matrices (therefore symmetrical), J_y is a pure imaginary matrix (therefore antisymmetrical).
[N.B. Standard representations fall into this category.]

3. Show that if any operator commutes with *two* of the components of an angular momentum vector, it commutes with the third.

4. Let $\boldsymbol{l}$ be the orbital angular momentum of a particle, θ and φ its polar angles and P the "parity operator". P is the operator effecting a reflection in the origin; its action on a function $F(\theta, \varphi)$ is defined by:

$$PF(\theta, \varphi) = F(\pi - \theta, \varphi + \pi).$$

Show that $[P, \boldsymbol{l}] = 0$, and from this that the spherical harmonics have a well-defined parity depending only on the quantum number l. Determine it.

5. Let $\boldsymbol{r}$, $\boldsymbol{r}'$ be two vectors in ordinary space, $\Omega \equiv (\theta, \varphi)$, $\Omega' \equiv (\theta', \varphi')$ their respective polar angles, $\boldsymbol{l}$, $\boldsymbol{l}'$ the corresponding angular momentum operators; let α $(0 \leqslant \alpha \leqslant \pi)$ be the angle between them: $\boldsymbol{r}\cdot\boldsymbol{r}' = rr' \cos\alpha$. The Legendre polynomial $P_l(\cos\alpha)$ is, through $\cos\alpha$, a function of the polar angles of $\boldsymbol{r}$ and $\boldsymbol{r}'$. Show that it satisfies the partial differential equations

$$\boldsymbol{l}^2\, P_l(\cos\alpha) = \boldsymbol{l}'^2\, P_l(\cos\alpha) = l(l+1)\, P_l(\cos\alpha)$$
$$(l_i + l_i')\, P_l(\cos\alpha) = 0 \qquad (i = x, y \text{ or } z).$$

Deduce from these the Addition Theorem:

$$\frac{2l+1}{4\pi} P_l(\cos\alpha) \equiv Y_l^0(\alpha)\, Y_l^0(0) = \sum_{m=-l}^{+l} (-)^m\, Y_l^m(\Omega)\, Y_l^{-m}(\Omega').$$

6. Let $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w}$ be three unit vectors forming a right-handed Cartesian system. Show that the infinitesimal rotation

$$\hat{\mathscr{R}} \equiv \mathscr{R}_v^{-1}(\varepsilon)\, \mathscr{R}_u^{-1}(\varepsilon)\, \mathscr{R}_v(\varepsilon)\, \mathscr{R}_u(\varepsilon)$$

(notation of section III) differs from $\mathscr{R}_w(-\varepsilon^2)$ only by terms of higher order than ε^2. Starting from the defining equation (XIII.58), calculate directly the infinitesimal rotation operator $\hat{R}$ up to order ε^2 and verify the commutation relation $[J_u, J_v] = \mathrm{i} J_w$.

7. With the aid of the commutation relations (XIII.56), verify that the scalar product of two vector operators, $\boldsymbol{A}$ and $\boldsymbol{B}$, $\boldsymbol{A}\cdot\boldsymbol{B} \equiv A_x B_x + A_y B_y + A_z B_z$, commutes with the components of total angular momentum.

8. Show that

$$\exp(-\mathrm{i}\beta J_y) = \exp(\tfrac{1}{2}\mathrm{i}\pi J_x) \exp(-\mathrm{i}\beta J_z) \exp(-\tfrac{1}{2}\mathrm{i}\pi J_x).$$

Deduce from this that the matrix elements $\langle JM \mid \exp(-\mathrm{i}\beta J_y) \mid JM' \rangle$ are polynomials of degree $2J$ with respect to the variables $\sin\frac{1}{2}\beta$ and $\cos\frac{1}{2}\beta$.

9. Prove the identity $(\boldsymbol{\sigma}\cdot\boldsymbol{A})(\boldsymbol{\sigma}\cdot\boldsymbol{B}) = (\boldsymbol{A}\cdot\boldsymbol{B}) + \mathrm{i}\boldsymbol{\sigma}\cdot(\boldsymbol{A}\times\boldsymbol{B})$.
[$\boldsymbol{\sigma} \equiv (\sigma_x\ \sigma_y\ \sigma_z)$ Pauli matrices, $\boldsymbol{A}$ and $\boldsymbol{B}$ vector operators commuting with $\boldsymbol{\sigma}$ but not necessarily with each other.]

10. Let $\boldsymbol{s}$ be the intrinsic angular momentum of a particle of spin 1 ($\boldsymbol{s}^2 = s(s+1) = 2$).

(i) Show that for any component $s_u \equiv (\boldsymbol{s}\cdot\boldsymbol{u})$ one has

$$s_u^3 = s_u \qquad \exp(-\mathrm{i}\varphi s_u) = 1 - \mathrm{i}\sin\varphi s_u - (1-\cos\varphi)\, s_u^2,$$

and give an explicit expression for the rotation matrix $R^{(1)}(\alpha\beta\gamma)$.

(ii) Let $|z\rangle$ be a vector of norm 1 such that $s_z|z\rangle = 0$, and let $|x\rangle$ and $|y\rangle$ be the vectors obtained from it by rotations of $+\frac{1}{2}\pi$ about Oy and $-\frac{1}{2}\pi$ about Ox respectively. Prove the following relations, as well as those resulting from circular permutation of x, y and z:

$$s_x|x\rangle = 0 \qquad s_x|y\rangle = \mathrm{i}|z\rangle \qquad s_x^2|y\rangle = |y\rangle$$
$$s_x|z\rangle = -\mathrm{i}|y\rangle \qquad s_x^2|z\rangle = |z\rangle.$$

Use these to show that $|x\rangle$, $|y\rangle$, $|z\rangle$ form an orthonormal basis and that the matrices representing s_x, s_y and s_z in that basis are those given in § 21.

(iii) Show that $\langle i \mid R(\alpha\beta\gamma) \mid j \rangle = \mathscr{R}_{ij}(\alpha\beta\gamma)$ $\quad (i, j = x, y \text{ or } z)$
(notation of § 10 and 14).

11. Let $\boldsymbol{S}$ be the total spin of a system of two nucleons. Show that the operator $Q \equiv (\boldsymbol{S}\cdot\boldsymbol{r})^2/r^2$ is a projector. Show that the tensor operator $S_{12} \equiv 2[3Q - \boldsymbol{S}^2]$ satisfies the identity: $S_{12}{}^2 = 4\boldsymbol{S}^2 - 2S_{12}$ and that its only possible eigenvalues are 0, 2, and -4. Determine the action of Q, and that of S_{12}, on the functions of the angular variables and spins $\mathscr{Y}^M_{LSJ}$ introduced in § 28.

[If one adopts the condensed notation

$$\mathscr{Y}^{(0)} \equiv \mathscr{Y}^M_{J0J},\ \mathscr{Y}_0{}^{(1)} \equiv \mathscr{Y}^M_{J1J},\ \mathscr{Y}_\pm{}^{(1)} \equiv \mathscr{Y}^M_{J\pm 1\,1J},$$

one finds

$$Q\,\mathscr{Y}^{(0)} = 0 \qquad (2J+1)\,Q\,\mathscr{Y}_+{}^{(1)} = J\,\mathscr{Y}_+{}^{(1)} + \sqrt{J(J+1)}\,\mathscr{Y}_-{}^{(1)}$$

$$Q\,\mathscr{Y}_0{}^{(1)} = \mathscr{Y}_0{}^{(1)} \qquad (2J+1)\,Q\,\mathscr{Y}_-{}^{(1)} = \sqrt{J(J+1)}\,\mathscr{Y}_+{}^{(1)} + (J+1)\,\mathscr{Y}_-{}^{(1)}.\ ]$$

12. Consider a spin $\frac{1}{2}$ particle. Show that in the space of states of a given orbital angular momentum l, the operators

$$\frac{l+1+\boldsymbol{l}\cdot\boldsymbol{\sigma}}{2l+1} \quad \text{and} \quad \frac{l-\boldsymbol{l}\cdot\boldsymbol{\sigma}}{2l+1}$$

are projectors onto the states of total angular momentum $j = l+\frac{1}{2}$ and $j = l-\frac{1}{2}$ respectively.

13. We add two equal angular momenta $j_1 = j_2 = j$. Without using the symmetry properties of the C.–G. coefficients show that in a permutation of m_1 and m_2 the eigenfunctions of the total angular momentum are either symmetrical (invariant) or antisymmetrical (multiplied by -1), and that this symmetry property depends only on J. Show that they are symmetrical or antisymmetrical according as $(-)^{2j+J}$ is equal to $+1$ or -1.

14. We denote by $\boldsymbol{J}^2\{A\}$ the following function of the operator A and the components of angular momentum:

$$\boldsymbol{J}^2\{A\} \equiv [J_x, [J_x, A]] + [J_y, [J_y, A]] + [J_z, [J_z, A]].$$

Show that if $\mathsf{T}^{(k)}$ is a kth-order irreducible tensor operator, its components verify the relation

$$\boldsymbol{J}^2\{T_q{}^{(k)}\} = k\,(k+1)\,T_q{}^{(k)}.$$

15. Let $a_r, a_r{}^\dagger$ $(r = 1, 2)$ be the annihilation and creation operators of a two-dimensional, isotropic, harmonic oscillator:

$$[a_r, a_s] = [a_r{}^\dagger, a_s{}^\dagger] = 0 \qquad [a_r, a_s{}^\dagger] = \delta_{rs}.$$

We write

$$S = \tfrac{1}{2}[a_1{}^\dagger a_1 + a_2{}^\dagger a_2]$$

$$J_1 = \tfrac{1}{2}[a_2{}^\dagger a_1 + a_1{}^\dagger a_2] \qquad J_2 = \tfrac{1}{2}\mathrm{i}\,[a_2{}^\dagger a_1 - a_1{}^\dagger a_2] \qquad J_3 = \tfrac{1}{2}[a_1{}^\dagger a_1 - a_2{}^\dagger a_2].$$

J_1, J_2, J_3 may be considered as the Cartesian coordinates of a certain vector operator $\boldsymbol{J}$.

[N.B. Following the notations of § XII.14, $L = 2J_2$.]

(i) Show that the components of $\boldsymbol{J}$ verify the commutation relations $\boldsymbol{J} \times \boldsymbol{J} = \mathrm{i}\boldsymbol{J}$ characteristic of an angular momentum, and that one has:

$$\boldsymbol{J}^2 = S(S+1) \qquad \text{(therefore } [S, \boldsymbol{J}] = 0\text{)}.$$

(ii) $\boldsymbol{J}$ will henceforth be considered to be the angular momentum of the system, and we denote the eigenvalues of $\boldsymbol{J}^2$ and J_3 by $j(j+1)$ and m respectively. Show that $\boldsymbol{J}^2$ and J_3 form a complete set of commuting observables, and that j may take all *integral* and *half-integral* values $\geqslant 0$ i.e.

$$j = 0, \tfrac{1}{2}, 1, \tfrac{3}{2}, 2, \ldots, \infty.$$

Show that the vectors $[(j+m)!(j-m)!]^{-\frac{1}{2}} a_1^{\dagger j+m} a_2^{\dagger j-m} \mid 0\rangle$ form the basis of a standard representation $\{\boldsymbol{J}^2, J_3\}$.

(iii) Show that $a_1^\dagger$ and $a_2^\dagger$ are respectively the $+\frac{1}{2}$ and $-\frac{1}{2}$ components of an irreducible tensor operator of order $\frac{1}{2}$, and that it follows that the expressions $R\, a_r^\dagger\, R^{-1}$ $(r = 1, 2)$, where R denotes the rotation operator

$$R = \exp(-\mathrm{i}\alpha J_3) \exp(-\mathrm{i}\beta J_2) \exp(-\mathrm{i}\gamma J_3),$$

are linear combinations of $a_1^\dagger$ and $a_2^\dagger$. Determine the coefficients.

(iv) Use the preceding results to demonstrate the Wigner formula (C.72) and the principal properties of the $R^{(j)}$ matrices set forth in Appendix C (except for the formulas of composition and reduction).

16. Show that if the $(2k+1)$ operators $T_q^{(k)}$ $(q = -k, \ldots, +k)$ verify the commutation relations (XIII.123), the $(2k+1)$ operators $S_q^{(k)} \equiv (-)^q T_{-q}^{(k)\dagger}$ have the same property.

17. Show that the integral

$$\int Y_{l_1}^{m_1}(\theta, \varphi)\, Y_{l_2}^{m_2}(\theta, \varphi)\, Y_{l_3}^{m_3}(\theta, \varphi)\, \mathrm{d}\Omega$$

is proportional to $(-)^{m_3}\langle l_1\, l_2\, m_1\, m_2 \mid l_3 - m_3\rangle$, the coefficient of proportionality being independent of m_1, m_2 and m_3. Determine this coefficient. [Use the addition theorem of Problem XIII.5.]

18. Show that the "tensor" operator

$$S_{12} \equiv 2\left[3\frac{(\boldsymbol{S}\cdot\boldsymbol{r})^2}{r^2} - \boldsymbol{S}^2\right]$$

considered as a function of $\boldsymbol{r}$ depends only on the angles θ and φ, and this through *spherical harmonics of order* 2.

[We find

$$S_{12} = \left(\frac{24\pi}{5}\right)^{\frac{1}{2}} \{S_-{}^2\, Y_2{}^2 - (S_-\, S_z + S_z\, S_-)\, Y_2{}^1 + \sqrt{\frac{2}{3}}(3S_z{}^2 - \boldsymbol{S}^2)\, Y_2{}^0 + (S_+\, S_z + S_z\, S_+)\, Y_2{}^{-1} + S_+{}^2\, Y_2{}^{-2}\}$$

$$= \left(\frac{24\pi}{5}\right)^{\frac{1}{2}} (\boldsymbol{S}^{(2)} \cdot \boldsymbol{Y}^{(2)}).$$

S_{12} is the scalar product [in the sense of definition (C.87)] of the second-order irreducible tensor operators $\mathbf{S}^{(2)}$ and $\mathbf{Y}^{(2)}$, which depend respectively on the spin and orbital variables.]

19. Let K_u denote the component of a vector operator $\boldsymbol{K}$ in a given direction, J_u the component of the total angular momentum $\boldsymbol{J}$ in the same direction, and let $|\tau Ja\rangle$, $|\tau Jb\rangle$ be two ket vectors belonging to the *same* subspace $\mathscr{E}(\tau J)$ (definition of § 16). Show that:

$$\langle \tau\, J\, a\, |K_u\, |\tau\, J\, b\rangle = \langle \tau\, J\, a\, |J_u\, |\tau\, J\, b\rangle \frac{\langle \mathbf{J}\cdot\mathbf{K}\rangle}{J\,(J+1)},$$

where $\langle \boldsymbol{J}\cdot\boldsymbol{K}\rangle$ denotes the average value of the scalar operator $\boldsymbol{J}\cdot\boldsymbol{K}$ in this subspace:

$$\langle \mathbf{J}\cdot\mathbf{K}\rangle = \langle \tau\, J\, a\, |\mathbf{J}\cdot\mathbf{K}\, |\tau\, J\, a\rangle$$

[i.e. the elements of the matrices of $\boldsymbol{K}$ in $\mathscr{E}(\tau J)$ are the same as those of its "projection" $\boldsymbol{J}(\boldsymbol{J}\cdot\boldsymbol{K})/J(J+1)$.]

SYSTEMS OF IDENTICAL PARTICLES. PAULI EXCLUSION PRINCIPLE

1. Identical Particles in Quantum Theory

Two particles are identical if their physical properties are exactly the same, this precluding the possibility of an observation that could distinguish between them. In Classical Mechanics this property of *indistinguishability of identical particles* plays a secondary role; in Quantum Mechanics it raises a serious problem.

As a definite example, let us consider the collision of two identical particles, and the extent to which their identity can affect the theory.

If the system obeys the laws of Classical Mechanics, its dynamical state is defined at any time by giving $\xi^{(1)} \equiv (\boldsymbol{r}^{(1)}, \boldsymbol{p}^{(1)})$, the position and momentum of particle 1, and $\xi^{(2)} \equiv (\boldsymbol{r}^{(2)}, \boldsymbol{p}^{(2)})$, the position and momentum of particle 2. The evolution of the system is determined by a certain Hamiltonian function depending on these 12 variables:

$$H(\xi^{(1)}, \xi^{(2)}) \equiv H(\boldsymbol{r}^{(1)}, \boldsymbol{p}^{(1)}, \boldsymbol{r}^{(2)}, \boldsymbol{p}^{(2)}).$$

If we have a potential $V(r)$ depending only on the distance between the two particles, and if m is their mass, then

$$H(\xi^{(1)}, \xi^{(2)}) = \frac{\boldsymbol{p}^{(1)2}}{2m} + \frac{\boldsymbol{p}^{(2)2}}{2m} + V(|\boldsymbol{r}^{(1)} - \boldsymbol{r}^{(2)}|). \tag{XIV.1}$$

The particles being identical, there can be no modification of the dynamical properties of the system when they are permuted, i.e. when we ascribe the dynamical state of particle 1 to particle 2, and *vice versa*. In particular the function H is invariant in such a permutation:

$$H(\xi', \xi'') = H(\xi'', \xi'). \tag{XIV.2}$$

On the other hand, the state of the system at any time can only be known to within a permutation of the variables 1 and 2; an observation of the system at a given time shows that one particle is in a certain state ξ' and the other in another state ξ'', but does not permit to determine which particle is in which state. This would seem to constitute a difficulty but, as we shall see, it is only an apparent one.

Suppose that at time t_0 we have one of the particles in state ξ_0' and the other in state ξ_0''. There are two possibilities; either it is particle 1 that occupies ξ_0' or it is particle 2. These correspond, however, to one and the same physical situation, for since H has the symmetry property (XIV.2), the laws of motion $\xi'(t)$ and $\xi''(t)$ of the particles in ξ_0' and ξ_0'' respectively at time t_0 are the same in the two cases. It is a simple matter of convention whether the particle starting from ξ_0' be called 1 and the one starting from ξ_0'' called 2, or inversely.

It is not so simple if the two-particle system is governed by the laws of Quantum Mechanics. The beginning of the foregoing analysis can be repeated word for word. Here again, the identity of the two particles finds its expression in the invariance of the Hamiltonian under permutation of their dynamical variables [eq. (XIV.2)], and more generally in the fact that any physical observable is invariant in such a permutation. As in Classical Mechanics, there results a certain ambiguity in the determination of the state of the system, but the ambiguity is now more profound, and its consequences more serious.

Let us suppose that an observation made on the system before the collision shows that one of the particles is in the state $\psi_0'(\boldsymbol{r})$, and that the other is in the state $\psi_0''(\boldsymbol{r})$ [1]. In practice, these functions represent wave packets localized in different regions of space; it follows that the functions

$$\begin{aligned} \psi_0(\boldsymbol{r}^{(1)}, \boldsymbol{r}^{(2)}) &\equiv \psi_0'(\boldsymbol{r}^{(1)})\, \psi_0''(\boldsymbol{r}^{(2)}) \\ \bar{\psi}_0(\boldsymbol{r}^{(1)}, \boldsymbol{r}^{(2)}) &\equiv \psi_0''(\boldsymbol{r}^{(1)})\, \psi_0'(\boldsymbol{r}^{(2)}) \end{aligned} \qquad \text{(XIV.3)}$$

are linearly independent. This initial observation does not permit to decide whether the system is in state ψ_0 or in state $\bar{\psi}_0$. More precisely, the observation consists in simultaneously measuring a certain set of compatible variables, and the functions ψ_0 and $\bar{\psi}_0$ are both eigenfunctions corresponding to the set of values given by that measuring operation. Since any linear combination of these functions $\lambda\psi_0 + \mu\bar{\psi}_0$ has the same property, the initial observation does not permit to decide which of these linear combinations represents the initial state of the system. We say that there is an *exchange degeneracy*.

We shall now examine the evolution of the system in the course

---

[1] We suppose that the two identical particles are spinless, thus sharpening the parallel between Classical and Quantum Theory.

of time. Let $\psi(\boldsymbol{r}^{(1)}, \boldsymbol{r}^{(2)}, t)$ and $\bar{\psi}(\boldsymbol{r}^{(1)}, \boldsymbol{r}^{(2)}, t)$ denote the solutions of the Schrödinger equation corresponding respectively to the initial conditions ψ_0 and $\bar{\psi}_0$. Since the Hamiltonian has the symmetry property (XIV.2), they are obtained from one another by permutation of the arguments $\boldsymbol{r}^{(1)}$ and $\boldsymbol{r}^{(2)}$. It is convenient to introduce functions that are symmetrical and antisymmetrical in this permutation:

$$\psi^{(\mathrm{S})} = \frac{1}{\sqrt{2}}(\psi + \bar{\psi}) \qquad \psi^{(\mathrm{A})} = \frac{1}{\sqrt{2}}(\psi - \bar{\psi}).$$

These are the solutions of the Schrödinger equation corresponding respectively to the initial conditions:

$$\psi_0{}^{(\mathrm{S})} = \frac{1}{\sqrt{2}}(\psi_0 + \bar{\psi}_0) \qquad \psi_0{}^{(\mathrm{A})} = \frac{1}{\sqrt{2}}(\psi_0 - \bar{\psi}_0).$$

If initially the system is in the state

$$\Psi_0 = \alpha\psi_0{}^{(\mathrm{A})} + \beta\psi_0{}^{(\mathrm{S})} \qquad (|\alpha|^2 + |\beta|^2 = 1),$$

at a later time t it will be in the state

$$\Psi(\boldsymbol{r}^{(1)}, \boldsymbol{r}^{(2)}, t) = \alpha\psi^{(\mathrm{A})} + \beta\psi^{(\mathrm{S})}. \tag{XIV.4}$$

The probability density $P(\boldsymbol{r}', \boldsymbol{r}'')$ of finding one of the particles at $\boldsymbol{r}'$ and the other at $\boldsymbol{r}''$ is [1])

$$P(\boldsymbol{r}', \boldsymbol{r}'') = |\Psi(\boldsymbol{r}', \boldsymbol{r}'')|^2 + |\Psi(\boldsymbol{r}'', \boldsymbol{r}')|^2 \tag{XIV.5}$$

$$= 2[|\alpha|^2\,|\psi^{(\mathrm{A})}(\boldsymbol{r}', \boldsymbol{r}'')|^2 + |\beta|^2\,|\psi^{(\mathrm{S})}(\boldsymbol{r}', \boldsymbol{r}'')|^2]. \tag{XIV.6}$$

For this expression to be independent of α and β we must have

$$|\psi^{(\mathrm{A})}(\boldsymbol{r}', \boldsymbol{r}'')| = |\psi^{(\mathrm{S})}(\boldsymbol{r}', \boldsymbol{r}'')|.$$

This equality holds for all $\boldsymbol{r}'$ and $\boldsymbol{r}''$ so long as the two particles have not entered into interaction and the two wave packets $\psi'(\boldsymbol{r})$ and $\psi''(\boldsymbol{r})$ do not overlap. It ceases to be true, in general, when one or the other of these conditions is not fulfilled. One can easily convince oneself by considering a few particular cases. Suppose, for example, that the two particles do not interact ($V(|\boldsymbol{r}^{(1)} - \boldsymbol{r}^{(2)}|) = 0$), and that they are moving freely towards one another. ψ will then be a product

---

[1]) To obtain the second line we use the symmetry properties of the functions $\psi^{(\mathrm{A})}$ and $\psi^{(\mathrm{S})}$:

$$\Psi(\boldsymbol{r}'', \boldsymbol{r}') = -\alpha\psi^{(\mathrm{A})}(\boldsymbol{r}', \boldsymbol{r}'') + \beta\psi^{(\mathrm{S})}(\boldsymbol{r}', \boldsymbol{r}'').$$

of two free wave packets: $\psi = \psi'(\boldsymbol{r}^{(1)}, t)\, \psi''(\boldsymbol{r}^{(2)}, t)$. During a certain period of time the wave packets will overlap, i.e. there exists a region of space where the functions $\psi'(\boldsymbol{r})$ and $\psi''(\boldsymbol{r})$ are both different from zero; if $\boldsymbol{r}$ is a point in that region,

$$|\psi^{(S)}(\boldsymbol{r}, \boldsymbol{r})| = \sqrt{2}\, |\psi'(\boldsymbol{r})\, \psi''(\boldsymbol{r})| \neq 0,$$

while

$$|\psi^{(A)}(\boldsymbol{r}, \boldsymbol{r})| = 0 \tag{XIV.7}$$

for any $\boldsymbol{r}$. Another interesting example is provided by the problem of scattering in a central potential. It is described in § 9, where we show that the scattering amplitude for $\psi^{(A)}$ is a superposition of odd-order spherical partial waves, and that of $\psi^{(S)}$ a superposition of even-order spherical partial waves; in general, these two amplitudes will not have the same absolute value, and consequently the differential cross section will depend crucially on the ratio $|\alpha|^2/|\beta|^2$.

Thus the existence of an exchange degeneracy is a source of real difficulty since it prevents us from making a precise theoretical prediction for the statistical distribution of results of measurements to be made on the system after the collision.

This difficulty is overcome by the introduction of the following *symmetrization postulate*, which fixes once and for all the coefficients α and β of the linear combination (XIV.3), and thus lends itself easily to experimental verification.

The dynamical states of a system of two identical particles are necessarily either all symmetrical ($\alpha = 0, \beta = 1$) *or all antisymmetrical* ($\alpha = 1, \beta = 0$) *in the permutation of the two particles.*

One or the other of these two possibilities actually occurs, depending on the nature of the particles considered. This postulate is easily extended to systems of any number of identical particles. It will be given in its general form and its principal consequences analysed in section I of this chapter. Section II is devoted to applications.

In the course of this work we shall be using certain elementary properties of permutations. These are given in § 14 of Appendix D [1]).

---

[1]) The elements of group theory brought together in Appendix D are not necessary for the understanding of the present chapter. The contents of § 14 of that Appendix form a whole independent of the rest; the only ideas of group theory used there are certain definitions relating to groups (group, class, invariant subgroup, etc.) which are set forth in § 2 of the Appendix.

I. SYMMETRIZATION POSTULATE

2. Similar Particles and the Symmetrical Representation

Consider a system of N particles. The dynamical variables describing the ith particle are functions of its position $\boldsymbol{r}^{(i)}$, of its momentum $\boldsymbol{p}^{(i)}$ and of its spin $\boldsymbol{s}^{(i)}$. These three vectors will henceforth collectively be denoted by $\xi^{(i)}$. Given the magnitude of the spin of the ith particle, we can construct the space of its dynamical states, $\mathscr{F}^{(i)}$. The space $\mathscr{E}$ of the dynamical states of the whole system is the tensor product

$$\mathscr{E} = \mathscr{F}^{(1)} \otimes \mathscr{F}^{(2)} \otimes \ldots \otimes \mathscr{F}^{(N)}. \tag{XIV.8}$$

By definition, two of these particles are *similar* if they have the same spin (similarity does not necessarily imply identity). The observables and state vectors of the one will then be in one-to-one correspondence with the observables and state vectors of the other, and consequently it is possible to exchange two similar particles. More generally, if n particles are similar, the $n!$ permutations of these n particles are well-defined operations, to each of which there corresponds a certain operator of the space $\mathscr{E}$. We now proceed to construct these permutation operators, and to simplify the discussion we suppose $n = N$.

Consider one of these N similar particles. ξ is the set of its fundamental observables, and $\mathscr{F}$ the space of its state vectors. Let q be a complete set of commuting observables of $\mathscr{F}$, and $|q_\varkappa\rangle$ the eigenvectors of a basis for q, with corresponding eigenvalues $q_\varkappa$ (the index or set of indices $\varkappa$ serves to distinguish the eigenvalues one from another). Then

$$\langle q_\varkappa | q_\mu \rangle = \delta_{\varkappa\mu}. \tag{XIV.9}$$

q might, for example, be the three components x, y, z of $\boldsymbol{r}$, taken together with the component of spin in the z direction, s_z. Each particle a ($a = 1, 2, \ldots, N$) of the system has its own set of commuting observables $q^{(a)}$. Clearly the set $Q \equiv (q^{(1)}, q^{(2)}, \ldots, q^{(N)})$ is a complete set of commuting observables for the space $\mathscr{E}$, and the vectors

$$|q_\alpha{}^{(1)}\, q_\beta{}^{(2)} \ldots q_\nu{}^{(N)}\rangle \equiv |q_\alpha\rangle^{(1)} |q_\beta\rangle^{(2)} \ldots |q_\nu\rangle^{(N)} \tag{XIV.10}$$

formed by taking the tensor product of the basis vectors of the spaces $\mathscr{F}^{(1)}, \mathscr{F}^{(2)}, \ldots, \mathscr{F}^{(N)}$ form the basis for a certain representation of the

vectors and operators of $\mathscr{E}$ – the $\{Q\}$ representation. We shall call a representation of this type a *symmetrical representation.*

3. Permutation Operators

The state represented by vector (XIV.10) is obviously the one in which particle 1 is in state $|q_\alpha\rangle$, particle 2 in state $|q_\beta\rangle$, ..., particle N in state $|q_\nu\rangle$. A permutation of the N particles modifies their distribution among $|q_\alpha\rangle$, $|q_\beta\rangle$, ..., $|q_\nu\rangle$, and therefore replaces (XIV.10) by another, in general different [1]), vector of the basis $\{Q\}$. Thus such an operation establishes a one-to-one correspondence between the vectors of this orthonormal basis, and consequently defines a certain linear and unitary operator in state-vector space. In this way we associate with each permutation of the N particles a *permutation operator* P having the unitary property:

$$PP^\dagger = P^\dagger P = 1. \tag{XIV.11}$$

Thus, in the transposition (1 2) – interchange of particles 1 and 2 – the vector (XIV.10) transforms into a vector representing a state in which particle 1 is in $|q_\beta\rangle$, particle 2 in $|q_\alpha\rangle$, and the others in the same states as before; the associated permutation operator $P_{(12)}$ is defined by the equation

$$P_{(12)}\,|q_\alpha^{(1)}\,q_\beta^{(2)}\,q_\gamma^{(3)}\,\ldots\,q_\nu^{(N)}\rangle = |q_\beta^{(1)}\,q_\alpha^{(2)}\,q_\gamma^{(3)}\,\ldots\,q_\nu^{(N)}\rangle.$$

To simplify the writing, we continue our study of permutations for the case $N=3$. The principles enunciated will of course hold for any N. As an example let us define the operator $P_{(123)}$ associated with the permutation (1 2 3) in which 1 takes the place of 3, 2 that of 1, and 3 that of 2:

$$P_{(123)}\,|q_\alpha^{(1)}\,q_\beta^{(2)}\,q_\gamma^{(3)}\rangle = |q_\gamma^{(1)}\,q_\alpha^{(2)}\,q_\beta^{(3)}\rangle. \tag{XIV.12}$$

If $|\psi\rangle$ is a vector of $\mathscr{E}$,

$$\begin{aligned} P_{(123)}\,|\psi\rangle &= \sum_{\alpha\beta\gamma} P_{(123)}\,|q_\alpha^{(1)}\,q_\beta^{(2)}\,q_\gamma^{(3)}\rangle\,\langle q_\alpha^{(1)}\,q_\beta^{(2)}\,q_\gamma^{(3)}\,|\psi\rangle \\ &= \sum_{\alpha\beta\gamma} |q_\gamma^{(1)}\,q_\alpha^{(2)}\,q_\beta^{(3)}\rangle\,\langle q_\alpha^{(1)}\,q_\beta^{(2)}\,q_\gamma^{(3)}\,|\psi\rangle. \end{aligned}$$

---

[1]) When the N individual states $|q_\alpha\rangle, \ldots, |q_\nu\rangle$ are not all different, certain permutations will leave vector (XIV.10) unchanged; if they are all identical, (XIV.10) will be unchanged by any permutation.

Re-naming the dummy indices in the last line, this becomes

$$P_{(123)}\,|\psi\rangle = \sum_{\alpha\beta\gamma} |q_\alpha^{(1)}\, q_\beta^{(2)}\, q_\gamma^{(3)}\rangle \langle q_\beta^{(1)}\, q_\gamma^{(2)}\, q_\alpha^{(3)}\,|\psi\rangle. \quad \text{(XIV.13)}$$

If $\psi(q_\alpha\, q_\beta\, q_\gamma)$ is the wave function of $|\psi\rangle$ in the $\{Q\}$ representation ($\psi(q_\alpha\, q_\beta\, q_\gamma)$ is the probability amplitude for finding particle 1 in $|q_\alpha\rangle$, 2 in $|q_\beta\rangle$, 3 in $|q_\gamma\rangle$), then the wave function of $P_{(123)}\,|\psi\rangle$ is the function

$$P_{(123)}\,\psi(q_\alpha\, q_\beta\, q_\gamma) = \psi(q_\beta\, q_\gamma\, q_\alpha) \quad \text{(XIV.14)}$$

obtained *by applying to the arguments of* $\psi(q_\alpha\, q_\beta\, q_\gamma)$ *the inverse of the permutation* (1 2 3).

The law for the transformation of vectors in a permutation is especially simple for vectors of the form

$$|u^{(1)}\, v^{(2)}\, w^{(3)}\rangle \equiv |u\rangle^{(1)}\, |v\rangle^{(2)}\, |w\rangle^{(3)};$$

it is the same as the law of transformation of the basis vectors of the $\{Q\}$ representation. The vector resulting from the action of P on such a vector is just the one obtained by applying the permutation p to the particles 1, 2, 3, in the individual states $|u\rangle$, $|v\rangle$, $|w\rangle$, i.e. [eq. (XIV.12)]

$$P_{(123)}\,|u^{(1)}\, v^{(2)}\, w^{(3)}\rangle = |w^{(1)}\, u^{(2)}\, v^{(3)}\rangle.$$

The demonstration is simple; equation (XIV.13) here reads

$$P_{(123)}\,|u^{(1)}\, v^{(2)}\, w^{(3)}\rangle = \sum_{\alpha\beta\gamma} |q_\alpha^{(1)}\, q_\beta^{(2)}\, q_\gamma^{(3)}\rangle \langle q_\alpha\,|w\rangle \langle q_\beta\,|u\rangle \langle q_\gamma\,|v\rangle$$

and the right-hand side is just the expansion of $|w^{(1)}\, u^{(2)}\, v^{(3)}\rangle$ in terms of the basis vectors of $\{Q\}$.

This property, verified here for a particular case, is a general one (Problem XIV.1). In consequence, the operator P associated with a given permutation is independent of the particular symmetrical representation chosen to define it.

Just as the action of P on a vector gives the vector obtained from the latter by the permutation p, the transformation of a given operator $F\,(\xi^{(1)}, \xi^{(2)}, \ldots, \xi^{(N)})$ of $\mathscr{E}$ by the unitary operator P gives the operator resulting from the application of p to the arguments of F. If we put

$$p = \begin{pmatrix} 1 & 2 & \ldots & N \\ \alpha_1 & \alpha_2 & \ldots & \alpha_N \end{pmatrix},$$

we have

$$PF(\xi^{(1)}, \xi^{(2)}, \ldots, \xi^{(N)})\, P^\dagger = F(\xi^{(\alpha_1)}, \xi^{(\alpha_2)}, \ldots, \xi^{(\alpha_N)}). \quad \text{(XIV.15)}$$

In particular

$$P_{(123)}\, F(\xi^{(1)}, \xi^{(2)}, \xi^{(3)})\, P^\dagger_{(123)} = F(\xi^{(2)}, \xi^{(3)}, \xi^{(1)}) \quad \text{(XIV.16)}$$

($\xi^{(2)}$ takes the place of $\xi^{(1)}$, $\xi^{(3)}$ that of $\xi^{(2)}$ and $\xi^{(1)}$ that of $\xi^{(3)}$).

To establish this rule it is sufficient to show that it holds when F is any one of the fundamental observables of the system. Consider one of these observables; one can always construct a symmetrical representation in which it is diagonal. Suppose, for example, that it is one of the $q^{(i)}$ of the $\{Q\}$ representation considered above. To show that $Pq^{(i)}\, P^\dagger = q^{(\alpha_i)}$, it suffices to show that the two sides have the same action on all of the basis vectors of $\{Q\}$. The demonstration is not difficult and we shall content ourselves here with carrying it through for the special case $N=3$, $i=1$, $p=(1\,2\,3)$:

$$\begin{aligned} P_{(123)}\, q^{(1)}\, P^\dagger_{(123)}\, |q_\alpha{}^{(1)}\, q_\beta{}^{(2)}\, q_\gamma{}^{(3)}\rangle &= P_{(123)}\, q^{(1)}\, |q_\beta{}^{(1)}\, q_\gamma{}^{(2)}\, q_\alpha{}^{(3)}\rangle \\ &= q_\beta\, P_{(123)}\, |q_\beta{}^{(1)}\, q_\gamma{}^{(2)}\, q_\alpha{}^{(3)}\rangle \\ &= q_\beta\, |q_\alpha{}^{(1)}\, q_\beta{}^{(2)}\, q_\gamma{}^{(3)}\rangle = q^{(2)}\, |q_\alpha{}^{(1)}\, q_\beta{}^{(2)}\, q_\gamma{}^{(3)}\rangle. \end{aligned}$$

As applied to observables, this definition of permutations answers well to one's intuitive idea of permutations of observables: the observable $PBP^\dagger$ resulting from the application of the permutation p to an observable B has the same eigenvalue spectrum as B, and the eigenvectors of $PBP^\dagger$ are obtained from the eigenvectors of B corresponding to the same eigenvalue by application of the permutation operator P.

In particular, B is invariant under permutations of the N particles if $PBP^\dagger = B$ for each of the $N!$ permutations of these particles, i.e. if

$$[P, B] = 0$$

for any P. If such is the case, we say that B is *symmetrical* with respect to the N particles.

4. Algebra of Permutation Operators. Symmetrizers and Antisymmetrizers

Two successive permutations p', p'' are equivalent to a single permutation $p = p''\, p'$. From the definition of a permutation operator

it is clear that the same relation holds for the associated operators P, P', P'':

$$P = P'' P'. \tag{XIV.17}$$

Thus the permutation operators obey the same algebraic relations as the permutations to which they correspond [1]).

In particular, any P can be expressed as a product of transpositions. In general there are many such products for the same P, but all of these products will contain either an even or an odd number of transpositions. The *parity* of the permutation is $+$ or $-$ according as this number is even or odd, and is denoted $(-)^P$. If P, P' and P'' are related by XIV.17, then clearly $(-)^p = (-)^{p'+p''}$.

Certain permutations, notably transpositions, are equal to their own inverse. In such a case [cf. eq. (XIV.11)], the associated operator is an observable, whose possible eigenvalues are ± 1.

As an example consider the transposition (ij):

$$P_{(ij)}{}^2 = 1. \tag{XIV.18}$$

The eigenvectors of the eigenvalue $+1$ are invariant in the transposition (ij); they are by definition *symmetrical* in i and j. The projector onto the space of all vectors symmetrical in i and j is the *symmetrizing operator*

$$S_{[ij]} = \tfrac{1}{2}(1 + P_{(ij)}). \tag{XIV.19}$$

The eigenvectors of the eigenvalue -1 change sign in the transposition (ij); they are by definition *antisymmetrical* in i and j; the projector onto the space of vectors antisymmetrical in i and j is the *antisymmetrizing operator*

$$A_{[ij]} = \tfrac{1}{2}(1 - P_{(ij)}). \tag{XIV.20}$$

---

[1]) In particular, $P_{(123)} = P_{(12)} P_{(23)}$. Let us verify that the successive application of $P_{(23)}$ and $P_{(12)}$ to the wave function $\psi(q_\alpha\, q_\beta\, q_\gamma)$ does give back the result (XIV.14). Put

$$\psi_1(q_\alpha\, q_\beta\, q_\gamma) \equiv P_{(23)}\psi(q_\alpha\, q_\beta\, q_\gamma);$$

then

$$\psi_1(q_\alpha\, q_\beta\, q_\gamma) = \psi(q_\alpha\, q_\gamma\, q_\beta)$$

and

$$P_{(123)}\,\psi(q_\alpha\, q_\beta\, q_\gamma) = P_{(12)}\,\psi_1(q_\alpha\, q_\beta\, q_\gamma) = \psi_1(q_\beta\, q_\alpha\, q_\gamma) = \psi(q_\beta\, q_\gamma\, q_\alpha).$$

The same result is obtained by *first* applying the permutation (1 2), and *then* the permutation (2 3), to the arguments of the function $\psi(q_\alpha\, q_\beta\, q_\gamma)$.

We obviously have

$$\begin{aligned} &S_{[ij]}+A_{[ij]}=1 \qquad && P_{(ij)}\,S_{[ij]}=S_{[ij]}\;P_{(ij)}=S_{[ij]} \\ &S_{[ij]}-A_{[ij]}=P_{(ij)} && P_{(ij)}\,A_{[ij]}=A_{[ij]}\,P_{(ij)}=-A_{[ij]}. \end{aligned} \tag{XIV.21}$$

Any vector is the sum of a vector that is antisymmetrical in i and j and a vector that is symmetrical in i and j; a decomposition of this type was used in discussing the collision of two identical particles in § 1.

We next consider the extension of the notion of symmetry and antisymmetry of dynamical states, which so far has been defined only for the permutation $P_{(ij)}$, to the more general case of the $N!$ permutations P.

We take a vector $|u\rangle$ of $\mathscr{E}$ and denote by $\mathscr{E}_u$ the subspace spanned by $|u\rangle$ and all the vectors that can be obtained from it by permutation. The number of dimensions of $\mathscr{E}_u$ is equal or inferior to $N!$ according as the $N!$ vectors $P|u\rangle$ are linearly independent or not.

An extreme case is when the $P|u\rangle$ all represent the same state:

$$P|u\rangle = c_P|u\rangle \tag{XIV.22}$$

for any p. There are restricting conditions on the constants c_P. If P is a transposition we know that c_P can take only the values ± 1. Further, since any transposition (ij) is equal to the product $(1i)\ (2j)\ (1\,2)\ (2j)\ (1i)$,

$$P_{(ij)} = P_{(1i)}\,P_{(2j)}\,P_{(12)}\,P_{(2j)}\,P_{(1i)}$$

and therefore

$$c_{(ij)} = c_{(1i)}{}^2\,c_{(2j)}{}^2\,c_{(12)} = c_{(12)}. \tag{XIV.23}$$

Hence the constant c is the same for all the transpositions: either $c_{\mathrm{tr}}=+1$ or $c_{\mathrm{tr}}=-1$, and since any permutation p is a product of transpositions, the corresponding constant c_P is a certain power of c_{tr}, this power being even or odd according as p is even or odd.

Thus we can conclude that equation (XIV.22) is fulfilled only in the two following cases:

$$\text{(a) for any } p,\ c_P=1 \qquad P|u\rangle=|u\rangle, \tag{XIV.24}$$

$$\text{(b) for any } p,\ c_P=(-)^p \qquad P|u\rangle=(-)^p\,|u\rangle; \tag{XIV.25}$$

vector $|u\rangle$ is said to be *symmetrical* or *antisymmetrical* with respect to permutations of the N particles according as it is (a) or (b) that applies.

The symmetrical vectors form a subspace $\mathscr{E}^{(S)}$ of $\mathscr{E}$, and the antisymmetrical vectors a second subspace $\mathscr{E}^{(A)}$ of $\mathscr{E}$, orthogonal to $\mathscr{E}^{(S)}$. We shall now show that the corresponding projectors are respectively given by

$$S = \frac{1}{N!} \sum_P P \qquad A = \frac{1}{N!} \sum_P (-)^p P \tag{XIV.26}$$

($\sum_P$ being extended over the $N!$ possible permutations). Consider the sequence obtained by arranging the permutations in some arbitrary order. Multiplication of each element on the right or on the left by a permutation operator P_1 can only modify their order in the sequence. Hence

$$P_1 S = S P_1 = S \qquad P_1 A = A P_1 = (-)^{p_1} A. \tag{XIV.27}$$

Replacing each element P by its inverse $P^\dagger$ also merely changes the order of the elements, and since the inverse of a permutation has the same parity as the permutation itself, we have

$$S = S^\dagger \qquad A = A^\dagger. \tag{XIV.28}$$

From equation (XIV.27) and definitions (XIV.26) we easily deduce the relations

$$S^2 = S \qquad A^2 = A \tag{XIV.29}$$

and

$$SA = AS = 0. \tag{XIV.30}$$

Relations (XIV.28–30) show that S and A are orthogonal projectors. Now if $|u\rangle$ is in $\mathscr{E}^{(S)}$, then from (XIV.24),

$$S|u\rangle = \frac{1}{N!} \sum_P P|u\rangle = \left(\frac{1}{N!} \sum_P\right)|u\rangle = |u\rangle$$

and, conversely, if $|\rangle$ is an arbitrary vector, then from (XIV.27),

$$PS|\rangle = S|\rangle;$$

hence S is indeed the projector onto $\mathscr{E}^{(S)}$. One can similarly show that A is the projector onto $\mathscr{E}^{(A)}$.

When $N=3$, S and A are explicitly given by

$$S = \tfrac{1}{6}(1 + P_{(12)} + P_{(23)} + P_{(31)} + P_{(123)} + P_{(321)})$$
$$A = \tfrac{1}{6}(1 - P_{(12)} - P_{(23)} - P_{(31)} + P_{(123)} + P_{(321)}).$$

As this example clearly shows, $S + A \neq 1$ if $N > 2$. In fact $S + A$

projects onto the space of states invariant under an even permutation of the N particles, a subspace of $\mathscr{E}$ when $N>2$.

Returning to the space $\mathscr{E}_u$ defined above, we deduce from eq. (XIV.27) that for any P,

$$SP|u\rangle = S|u\rangle \qquad AP|u\rangle = (-)^p A|u\rangle.$$

Thus the vectors $P|u\rangle$ spanning $\mathscr{E}_u$ all have the same projection on $\mathscr{E}^{(S)}$, and to within a sign the same projection on $\mathscr{E}^{(A)}$. Therefore, according as $S|u\rangle$ is different to zero or not, $\mathscr{E}_u$ *contains one, and only one symmetrical vector or none whatever*; and similarly, according as $A|u\rangle$ is different to zero or not, $\mathscr{E}_u$ *contains one and only one antisymmetrical vector or none whatever* [1]).

5. Identical Particles and the Symmetrization Postulate

If the N particles of the system considered above are not only similar but identical, then none of the dynamical properties of the system are modified by a permutation of these particles. From this invariance property, important deductions can be made concerning the law of motion and the observables of the system.

If $|\psi_0\rangle$ is the state of the system at the initial time t_0, its state at a later time t is obtained by application of the evolution operator $U(t, t_0)$: $|\psi_t\rangle = U(t, t_0)\,|\psi\rangle$. If the initial state is $P|\psi_0\rangle$, the evolution of the system is exactly the same except for the permutation P and at time t it will be in the state $P|\psi_t\rangle$. Consequently

$$U(t, t_0)\,P|\psi_0\rangle = P|\psi_t\rangle = PU(t, t_0)|\psi_0\rangle$$

and, since this must hold whatever $|\psi_0\rangle$,

$$[P, U(t, t_0)] = 0. \qquad \text{(XIV.31)}$$

If H is the Hamiltonian of the system, $U(t, t_0)$ is the solution of the Schrödinger equation

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t} U(t, t_0) = HU(t, t_0)$$

defined by the initial condition: $U(t_0, t_0) = 1$. From (XIV.31), and

[1]) If the $N!$ vectors $P|u\rangle$ are linearly independent, the particular linear combinations of these vectors $S|a\rangle$ and $A|u\rangle$ are certainly non-zero; in this case, $\mathscr{E}_u$ contains one symmetrical vector and one antisymmetrical vector.

from the relation obtained by differentiating it with respect to t, we obtain

$$[P, H] = 0. \qquad \text{(XIV.32)}$$

Conversely, if H and P commute, $U(t, t_0)$ and its transform under permutation $PUP^\dagger$ are equal since they satisfy the same Schrödinger equation with the same initial condition; and therefore U and P commute. Thus a necessary and sufficient condition that the equation of motion be invariant under permutation is that H satisfy (XIV.32) for all P.

Consider next a physical observable [1]) B of the system and let $|u\rangle$ be an eigenvector of B corresponding to the eigenvalue b. If the system is in the state $|u\rangle$, then a measurement of B is certain to give the result b, and if the system is in the state $P|u\rangle$ obtained by applying the permutation P to $|u\rangle$, a measurement of B must give the same result:

$$BP|u\rangle = bP|u\rangle$$

and this for any permutation P. In other words, each vector of the space $\mathscr{E}_u$, generated by applying all possible permutations of the N particles to $|u\rangle$, must also be an eigenvector of B corresponding to the same eigenvalue b (exchange degeneracy). A necessary and sufficient condition for this to be true for all eigenvalues of B is that (cf. § VII.15)

$$[B, P] = 0 \qquad \text{(XIV.33)}$$

for any P.

Thus, the N particles are identical if the Hamiltonian H and all the physical observables of the system are symmetrical with respect to these particles.

In consequence, if we should wish to determine the state of the system by a simultaneous measurement of the variables q of each individual particle, such a determination can at best be made only to within an exchange degeneracy [2]); we can affirm that of the total

---

[1]) The meaning of the expression *physical observable* has been given in § XIII.15. The present analysis of invariance under permutation may be compared with the analysis of the invariance of physical observables under "rotation through 2π" given in that paragraph.

[2]) $q^{(1)}, q^{(2)}, \ldots, q^{(N)}$ form a complete set of commuting observables in the space $\mathscr{E}$, but only symmetrical functions of these can be physical observables,

number of particles N there are n_1 particles in the state $|q_1\rangle$, n_2 in the state $|q_2\rangle$, ..., $n_\varkappa$ in the state $|q_\varkappa\rangle$, ... $(n_1+n_2+\ldots+n_\varkappa+\ldots=N)$, but the identity of the particles in each of these states will remain undetermined. There are $(N!/n_1!\,n_2!\ldots n_\varkappa!\ldots)$ basis vectors of the $\{Q\}$ representation having the desired property. Let $\mathscr{E}(n_1\,n_2\ldots n_\varkappa\ldots)$ be the space spanned by these vectors (generated by the application of the $N!$ permutations to any one of them). The state of the system is certainly represented by one of the vectors of this space, but the above-mentioned observation does not permit to decide which one. However, as we have seen in the example of § 1, the predictions of the theory depend upon which one, and this ambiguity is therefore a source of real difficulty. It is removed by the introduction of the SYMMETRIZATION POSTULATE:

The states of a system containing N identical particles are necessarily either all symmetrical or all antisymmetrical with respect to permutations of the N particles.

Which of these prescriptions is to be applied depends upon the nature of the identical particles. Particles with symmetrical states are called *bosons*, those with antisymmetrical states, *fermions*. (The reasons for these names will become clear later on.) *Experiment shows* that the elementary particles of spin $\frac{1}{2}$ occurring in nature (electrons, protons, neutrons, etc.,) are fermions, while those of integral spin (photons, π mesons, etc.) are bosons.

The above-defined space $\mathscr{E}(n_1\,n_2\ldots n_\varkappa)$ has no more than one symmetrical vector and no more than one antisymmetrical vector. Thus the symmetrization postulate completely removes the exchange degeneracy. It remains to show that it does not come into conflict with the fundamental postulates of Quantum Mechanics concerning the motion of physical systems and the measurement of physical quantities.

Let us consider the boson case (fermions may be treated in the same way). In the preceding paragraph we defined the projector onto

---

and these do not form a complete set in $\mathscr{E}$. The symmetrization postulate, to be introduced further along, will consist in limiting the state vector space to a certain subspace of $\mathscr{E}$ in which the physical observables in question form a complete set (and in which, therefore, the exchange degeneracy is completely removed).

the symmetrical states, S. It is a particular combination of permutation operators [eq. (XIV.26)], and therefore commutes with the evolution operator $U(t, t_0)$ for the system

$$[S, U(t, t_0)] = 0 \tag{XIV.34}$$

and with the physical observables

$$[S, B] = 0. \tag{XIV.35}$$

Eq. (XIV.34) shows that if the system is initially in a symmetrical state it will remain in a symmetrical state so long as it is left undisturbed. Let us now suppose that an ideal measurement of B is performed. From eq. (XIV.35), S and B have at least one common set of basis vectors; if the state of the system is symmetrical, the expansion of the state vector in terms of this basis contains only symmetrical eigenvectors of B. It is therefore certain that the operation of measuring B will leave the system in a symmetrical state.

6. Bosons and Bosc-Einstein Statistics

Consider a system of N bosons. The states of the system span the subspace $\mathscr{E}^{(S)}$ of $\mathscr{E}$. We can form a basis in $\mathscr{E}^{(S)}$ from the vectors of the $\{Q\}$ representation by proceeding in the following way.

In each subspace $\mathscr{E}(n_1\, n_2 \ldots n_\varkappa \ldots)$ we can form one, and only one, normalized symmetrical vector (defined to within a phase), the vector

$$\left[\frac{N!}{n_1!\, n_2! \ldots n_\varkappa! \ldots}\right]^{\frac{1}{2}} S|q_1{}^{n_1}\, q_2{}^{n_2} \ldots q_\varkappa{}^{n_\varkappa} \ldots\rangle \tag{XIV.36}$$

$|q_1{}^{n_1}\, q_2{}^{n_2} \ldots q_\varkappa{}^{n_\varkappa} \ldots\rangle$ being the basis vector of $\{Q\}$ for which the first n_1 particles are in state $|q_1\rangle$, the following n_2 in state $|q_2\rangle$, ..., the following $n_\varkappa$ in state $|q_\varkappa\rangle$, ...; S is the symmetrizing operator defined above [eq. (XIV.26)]; the constant in brackets is a normalization constant ($0! = 1$). The proof is as follows. The interchange of two particles occupying the same state leaves $|q_1{}^{n_1}\, q_2{}^{n_2} \ldots q_\varkappa{}^{n_\varkappa} \ldots\rangle$ unchanged, while the interchange of two particles occupying different states gives another of the basis vectors of $\{Q\}$. More generally, this vector is invariant under any of the $\prod n_\varkappa! \equiv n_1!\, n_2! \ldots n_\varkappa! \ldots$ permutations not changing the distribution of the N particles among the individual states $|q_1\rangle$, $|q_2\rangle$, ... $|q_\varkappa\rangle$, ...; any other permutation changes it into another of the basis vectors of $\{Q\}$. Applying each of the $N!$

permutations to $|q_1^{n_1} q_2^{n_2} \ldots q_\varkappa^{n_\varkappa} \ldots\rangle$ we form the $[N!/\prod n_\varkappa!]$ basis vectors of $\mathscr{E}(n_1 n_2 \ldots n_\varkappa \ldots)$, each of them being obtained $(\prod n_\varkappa!)$ times. The vector (XIV.36), being equal to $[\prod n_\varkappa!/N!]^{\frac{1}{2}}$ multiplied by the sum of these basis vectors, is therefore symmetrical and normalized to unity. Q.E.D.

Thus, to each sequence $n_1, n_2, \ldots, n_\varkappa, \ldots$ of non-negative integers such that

$$n_1 + n_2 + \ldots + n_\varkappa + \ldots = N,$$

there corresponds one and only one symmetrical state of the system, represented by the vector (XIV.36). The set of vectors so formed constitutes an orthonormal basis in $\mathscr{E}^{(S)}$.

Next we show that *a boson gas obeys Bose–Einstein statistics.* By boson gas is meant a system formed of a very large number N of bosons having mutual interactions weak enough to be neglected in a first approximation. The Hamiltonian H of the system can then be written as a sum of N individual Hamiltonians:

$$H = h^{(1)} + h^{(2)} + \ldots + h^{(i)} + \ldots + h^{(N)}. \qquad \text{(XIV.37)}$$

According to the Boltzmann theory, thermodynamical equilibrium is realized when the system is in the most probable "macroscopic state". A given "macroscopic state" is in fact a set of quantum states (or "microscopic states") sufficiently close to one another that it is impossible to differentiate between them at the macroscopic level. According to the ergodic hypothesis, microscopic states of the same energy are all equally probable. The probability of a given macroscopic state is therefore proportional to the number of distinct microscopic states composing it. The determination of the thermodynamical equilibrium of the system depends essentially on this number. Suppose that h is included in the set q of dynamical variables defining the representation $\{Q\}$. Each distribution

$$n_1, n_2, \ldots, n_\varkappa, \ldots$$

of the N particles among the different possible individual states

$$|q_1\rangle, |q_2\rangle, \ldots, |q_\varkappa\rangle, \ldots$$

defines one, and only one, microscopic state of the system [represented by (XIV.36)]. This is just the hypothesis of Bose–Einstein statistics,

where the particles are supposed indistinguishable and where in consequence states of the system differing only in the identity of the particles occupying the various individual states count as one and the same microscopic state. In Maxwell–Boltzmann statistics, on the other hand, each particle is supposed distinguishable at the microscopic level, and the $[N!/\prod n_\varkappa!]$ states of the system corresponding to the same distribution $n_1, n_2, \ldots, n_\varkappa, \ldots$ are counted as separate microscopic states.

Important remark – The general expression for the density operator representing the state of a system in thermodynamical equilibrium

$$\varrho = \mathrm{e}^{-H/kT}/\mathrm{Tr}\,\mathrm{e}^{-H/kT} \qquad \text{(XIV.38)}$$

and the expressions derived from it in § VIII.25 remain valid here. The profound difference introduced by the symmetrization postulate comes from the fact that ϱ is now an operator of $\mathscr{E}^{(S)}$ rather than of $\mathscr{E}$, and the different manipulations of quantum statistics, in particular the calculation of traces, must now be made in this more restricted space. Thus, when H is of the form (XIV.37), ϱ considered as an operator of the space $\mathscr{E}$ is a tensor product of operators defined in each of the individual spaces $\mathscr{F}^{(1)}, \mathscr{F}^{(2)} \ldots, \mathscr{F}^{(N)}$:

$$\varrho = \prod_{i=1}^{N} [\mathrm{e}^{-h^{(i)}/kT}/\mathrm{Tr}_i\,\mathrm{e}^{-h^{(i)}/kT}].$$

But this factorisation loses all significance if ϱ is an operator of $\mathscr{E}^{(S)}$, for the N factors taken separately are not operators of $\mathscr{E}^{(S)}$.

7. Fermions and Fermi-Dirac Statistics. Exclusion Principle

An analysis similar to the foregoing can be made of the N-fermion system. Its states span the subspace $\mathscr{E}^{(A)}$ of $\mathscr{E}$.

Starting from the $\{Q\}$ representation we obtain a complete set of orthonormal antisymmetrical vectors by taking a normalized antisymmetrical vector – if it exists – in each subspace $\mathscr{E}(n_1\, n_2 \ldots n_\varkappa \ldots)$. A necessary and sufficient condition that such a vector exist is that $A|q_1{}^{n_1}\, q_2{}^{n_2} \ldots q_\varkappa{}^{n_\varkappa} \ldots\rangle$ be different to zero. Let us suppose that we have at least one of the integers $n_1, n_2, \ldots, n_\varkappa, \ldots$ greater than 1. In the state represented by $|q_1{}^{n_1}\, q_2{}^{n_2} \ldots q_\varkappa{}^{n_\varkappa} \ldots\rangle$ at least two particles, the ith and the jth say, occupy the same individual state, so that

this vector is symmetrical in the interchange of these two particles, i.e.

$$|q_1^{n_1}\, q_2^{n_2} \ldots q_\varkappa^{n_\varkappa} \ldots\rangle = \tfrac{1}{2}(1+P_{(ij)})\, |q_1^{n_1}\, q_2^{n_2} \ldots q_\varkappa^{n_\varkappa} \ldots\rangle.$$

But, from (XIV.27)

$$A(1+P_{(ij)}) = 0,$$

so that

$$A|q_1^{n_1}\, q_2^{n_2} \ldots q_\varkappa^{n_\varkappa} \ldots\rangle = 0. \qquad \text{(XIV.39)}$$

In other words, *two fermions cannot occupy the same individual quantum state.* This is the *Pauli exclusion principle* [1]).

Suppose now that at most one particle is in any individual state ($n_\varkappa = 0$ or 1). The vector

$$A|q_1^{n_1}\, q_2^{n_2} \ldots q_\varkappa^{n_\varkappa} \ldots\rangle \equiv \frac{1}{N!} \sum_P (-)^p\, P|q_1^{n_1}\, q_2^{n_2} \ldots q_\varkappa^{n_\varkappa} \ldots\rangle$$

is the sum of $N!$ mutually orthogonal vectors and therefore is certainly non-zero. Its norm is $(1/N!)$. If $|q_\alpha\rangle, |q_\beta\rangle, \ldots, |q_\nu\rangle$ are the N individual occupied states, the corresponding antisymmetrical state is represented by the vector of norm 1: $\sqrt{N!}\, A\, |q_\alpha^{(1)}\, q_\beta^{(2)} \ldots q_\nu^{(N)}\rangle$. This vector can be written in the form of an $N \times N$ determinant (*the Slater determinant*)

$$\sqrt{N!}\, A|q_\alpha^{(1)}\, q_\beta^{(2)} \ldots q_\nu^{(N)}\rangle \equiv \frac{1}{\sqrt{N!}} \begin{vmatrix} |q_\alpha\rangle^{(1)} & |q_\alpha\rangle^{(2)} & \ldots & |q_\alpha\rangle^{(N)} \\ |q_\beta\rangle^{(1)} & |q_\beta\rangle^{(2)} & \ldots & |q_\beta\rangle^{(N)} \\ \vdots & & & \\ |q_\nu\rangle^{(1)} & |q_\nu\rangle^{(2)} & \ldots & |q_\nu\rangle^{(N)} \end{vmatrix} \qquad \text{(XIV.40)}$$

This identity can be directly verified from the definition of a determinant. Further, (XIV.40) is still valid when certain of the N individual states are identical, in which case two or more lines of the determinant are equal, giving back the Pauli exclusion principle.

Thus, to each set $|q_\alpha\rangle, |q_\beta\rangle, \ldots, |q_\nu\rangle$ of N *different* states chosen from among the individual states $|q_1\rangle, |q_2\rangle, \ldots, |q_\varkappa\rangle, \ldots$, there corresponds one, and only one, antisymmetrical state, which is represented by vector (XIV.40). The set of vectors so constructed constitutes an orthonormal basis in $\mathscr{E}^{(A)}$.

---

[1]) The exclusion principle was formulated by Pauli in 1925 as a general property of electrons permitting one to explain the structure of complex atoms (cf. § 12).

A fermion gas is governed by Fermi–Dirac statistics. The demonstration is analogous to the one for the boson case. Only the numbering of microscopic states is different. Each set of N different individual states defines one, and only one, microscopic state of the fermion system [represented by vector (XIV.40)]. This is just the hypothesis of Fermi–Dirac statistics, which states that the N particles are indistinguishable and that there can be no more than one of them in any individual state.

The note at the end of the preceding paragraph also applies to fermions. The density operator for a fermion system in thermodynamic equilibrium is given by equation (XIV.38); here of course, it is an operator in $\mathscr{E}^{(A)}$.

Thus, the differences between the three types of identical particle statistics lie in the definition of the state-vector space as indicated by the following table.

Statistics	Maxwell–Boltzmann	Bose–Einstein	Fermi–Dirac
Character of the particles	distinguishable	indistinguishable	indistinguishable + exclusion
State-vector space	$\mathscr{E}$	$\mathscr{E}^{(S)}$	$\mathscr{E}^{(A)}$

8. Is it always Necessary to Symmetrize the Wave Function

We consider a system of n identical particles. If the particles are electrons the state of the system will be represented by an antisymmetrical wave function. These are not, however, the only electrons in the universe. To ignore the others, and to treat this system as an entity distinct from the rest, supposes that the dynamical properties of the n electrons are not affected by the presence of the others. The question arises whether such a hypothesis is well founded, or whether the symmetrization postulate, in establishing a certain correlation between these n electrons and the others, renders it invalid.

In practice, the electrons of a system are all inside a certain spatial domain D, and the dynamical properties in which we are interested all correspond to measurements to be made inside this domain. It turns out that the other electrons may simply be ignored so long as

they remain outside D and so long as their interaction with the electrons of the system remains negligible. This is a general result and applies to bosons as well as to fermions. We shall prove it here for the special case of a system of two fermions.

If one ignores the existence of all other particles, the dynamical state of the two fermions is represented by a certain normalized antisymmetrical wave function, $\varphi(1, 2)$. 1 and 2 denote the coordinates and the spin components s_z of particles 1 and 2 respectively. In general, a given state of the system, χ say, is represented by a certain antisymmetrical wave function normalized to unity, $\chi(1, 2)$. If at a given time the system is in the state φ, its dynamical properties at that time are given by the set of probabilities

$$w = |\langle\chi|\varphi\rangle|^2. \tag{XIV.41}$$

In reality, the two fermions are part of a system of N fermions. Let us see if these dynamical properties are the same as those found if we correctly take into account the existence of the other $(N-2)$ fermions. Let $\Psi(3, 4, \ldots, N)$ be the normalized antisymmetrical wave function representing the dynamical state of the other $(N-2)$ fermions. If fermions 1 and 2 were not identical to fermions 3, 4, ..., N, the state of the whole system would be represented by

$$\varphi(1, 2)\, \Psi(3, 4, \ldots, N)$$

and would retain this factorisation property so long as the interaction between the first two fermions and the others remained negligible. Actually, the vector $|\Phi\rangle$, which correctly represents the state of the whole system, is proportional to the antisymmetrical vector $A|\varphi\Psi\rangle$, A being the antisymmetrizing operator for the N particles [definition (XIV.26)].

By hypothesis, the wave packets φ and Ψ do not overlap; more precisely, it is certain that the two fermions are inside a specified spatial domain D and that all the others are outside D. Moreover, we are concerned only with the dynamical properties of the two fermions inside D.

Denote by $\Theta(3, 4, \ldots, N)$ an antisymmetrical wave function of norm 1 that vanishes when any of the $(N-2)$ position vectors $\mathbf{r}^{(3)}, \ldots, \mathbf{r}^{(N)}$ are inside D. Θ represents the state of a system of $(N-2)$ fermions all situated outside D. By hypothesis, Ψ is a function of

this type. If the functions $\Theta_1, \Theta_2, \ldots, \Theta_i, \ldots$ form a complete orthonormal set of such functions, then

$$\Psi = \sum_i \Theta_i \langle \Theta_i | \Psi \rangle .$$

We denote by $\chi(12)$ any normalized antisymmetrical wave function of 1 and 2 having the converse property, i.e. the property that χ vanishes if either of the vectors $\boldsymbol{r}^{(1)}$, $\boldsymbol{r}^{(2)}$ is outside D. χ represents, therefore, two fermions inside D. By hypothesis, φ is such a function.

The permutations of the N particles may be put into two categories depending upon their action on the vector $|\chi\Theta\rangle$. Those of the first category, which we denote by F, can at most change the sign of $|\chi\Theta\rangle$; they are the $2!\,(N-2)!$ permutations which exchange 1 and 2 and/or exchange $3, 4, \ldots, N$ among themselves. Thus,

$$F|\chi\Theta\rangle = (-)^f |\chi\Theta\rangle .$$

All other permutations, which we denote by G, exchange at least one of the particles 1, 2 with one of the $(N-2)$ others; hence, $G|\chi\Theta\rangle$ represents a state in which at least one of the particles 1, 2 is definitely outside D; it is orthogonal to any vector of the type $|\chi\Theta\rangle$:

$$\langle \chi'\Theta' | G | \chi\Theta \rangle = 0.$$

We can now deduce the following identity:

$$\begin{aligned} \langle \chi'\Theta' | A | \chi\Theta \rangle &= \frac{1}{N!} \sum_P (-)^p \langle \chi'\Theta' | P | \chi\Theta \rangle \\ &= \frac{1}{N!} \sum_F (-)^f \langle \chi'\Theta' | F | \chi\Theta \rangle \qquad \text{(XIV.42)} \\ &= \frac{2!\,(N-2)!}{N!} \langle \chi'\Theta' | \chi\Theta \rangle . \end{aligned}$$

Notice that the norm of $A|\chi\Theta\rangle$ is $\langle \chi\Theta | A | \chi\Theta \rangle$, i.e. $2!(N-2)!/N!$.

We now wish to know the probability w that the two fermions inside D are in the state χ. If the $(N-2)$ other fermions were distinguishable from these two, the state of the system would be represented by $|\varphi\Psi\rangle$ and the probabilities in question would be:

$$\sum_i |\langle \chi\Theta_i | \varphi\Psi \rangle|^2 = |\langle \chi | \varphi \rangle|^2 \left(\sum_i |\langle \Theta_i | \Psi \rangle|^2 \right) = |\langle \chi | \varphi \rangle|^2 . \qquad \text{(XIV.43)}$$

Since the N fermions are identical, the state of the system is

$$|\Phi\rangle = \sqrt{\binom{2}{N}}\, A|\varphi\Psi\rangle \qquad \left(\binom{2}{N} = \frac{N!}{2!\,(N-2)!}\right)$$

and the sought-for probability is that of finding the system in any one of the states represented by the orthonormal antisymmetrical vectors

$$|X_i\rangle = \sqrt{\binom{2}{N}}\, A|\chi\Theta_i\rangle,$$

namely

$$w = \sum_i |\langle X_i|\Phi\rangle|^2 = \binom{2}{N}^2 \sum_i |\langle\chi\Theta_i|A|\varphi\Psi\rangle|^2.$$

From (XIV.42) and (XIV.43),

$$w = \sum |\langle\chi\Theta_i|\varphi\Psi\rangle|^2 = |\langle\chi|\varphi\rangle|^2,$$

which is just (XIV.41).

Thus we can just ignore the existence of the $(N-2)$ other fermions and still obtain the correct result.

II. APPLICATIONS

9. Collision of two Spinless Identical Particles

Here we resume the discussion of the collision problem considered in § 1.

Let $(\boldsymbol{R}, \boldsymbol{P})$ and $(\boldsymbol{r}, \boldsymbol{p})$ be the dynamical variables for the center of mass and for the relative particle respectively:

$$\boldsymbol{R} = \tfrac{1}{2}(\boldsymbol{r}^{(1)} + \boldsymbol{r}^{(2)}), \qquad \boldsymbol{r} = \boldsymbol{r}^{(1)} - \boldsymbol{r}^{(2)}. \tag{XIV.44}$$

The Hamiltonian is

$$H = \frac{\boldsymbol{P}^2}{4m} + \frac{\boldsymbol{p}^2}{m} + V(\boldsymbol{r}) \tag{XIV.45}$$

and the dynamical state of the system at any given time is represented by a certain wave function $\Psi(\boldsymbol{R}, \boldsymbol{r})$ depending on $\boldsymbol{R}$ and $\boldsymbol{r}$.

In the interchange of the two particles, $\boldsymbol{R}$ is invariant, and $\boldsymbol{r}$ goes

over into $-\boldsymbol{r}$. The wave function must satisfy the symmetrization postulate, so we have [1])

$$\Psi(\boldsymbol{R}, -\boldsymbol{r}) = \pm \Psi(\boldsymbol{R}, \boldsymbol{r}) \tag{XIV.46}$$

the upper sign corresponding to the case when the two particles are bosons, the lower to the case when they are fermions.

Let us first treat the same problem supposing the two particles distinguishable. The solution has been given in section I of Chapter X. We briefly recall it here, keeping the notation of § 4–7 of that chapter. Before collision ($t \ll 0$), the state of the system formed of the target particle and a given particle of the incident beam is characterized by the relative velocity $\boldsymbol{v} = \hbar\boldsymbol{k}/\frac{1}{2}m$ and by the impact parameter $\boldsymbol{b}$. It is represented by a product of the form $\Phi(\boldsymbol{R}, t)\ \varphi_{\boldsymbol{b}}(\boldsymbol{r}, t)$ where Φ and φ are two free wave packets of norm 1 [2]). If we work in the center-of-mass system, the group velocity of the wave $\Phi(\boldsymbol{R}, t)$ is zero and the wave $\varphi(\boldsymbol{r}, t)$ propagates with velocity $\boldsymbol{v}$. Under the conditions in which the cross sections are to be calculated, the spreading out of $\varphi(\boldsymbol{r}, t)$ is negligible; its form has been chosen once and for all and is given by [eq. (X.13)]

$$\varphi_{\boldsymbol{b}}(\boldsymbol{r}, t) = \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{b} - \mathrm{i}Et/\hbar}\, \chi(\boldsymbol{r} - \boldsymbol{v}t - \boldsymbol{b})\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}. \tag{XIV.47}$$

After the collision ($t \gg 0$) the wave function $\Psi_{\boldsymbol{b}}(\boldsymbol{R}, \boldsymbol{r}, t)$ defined by these initial conditions becomes

$$\Psi_{\boldsymbol{b}}(\boldsymbol{R}, \boldsymbol{r}, t) = \Phi(\boldsymbol{R}, t)\, [\varphi_{\boldsymbol{b}}(\boldsymbol{r}, t) + \psi_{\boldsymbol{b}}{}^{(d)}(\boldsymbol{r}, t)], \tag{XIV.48}$$

---

[1]) Permutation is an operation that differs from spatial reflection, in which $\boldsymbol{R}$ and $\boldsymbol{r}$ *both* change sign. In the present discussion, where the particles are spinless and where, as we shall see, the center of mass has no effective role, the two operations could easily be confused, as they have the same effect on the relative-particle wave function.

[2]) Rigorously speaking, the initial state should be represented by a product of the type (XIV.3), since there can be no correlation between the particles until they enter into collision. A product of the type $\Phi\varphi$ can be put in the form (XIV.3) only for particular types of waves Φ and φ. Example:

$$\Phi\varphi \propto \exp\, [\mathrm{i}(\boldsymbol{K}\cdot\boldsymbol{R} + \boldsymbol{k}\cdot\boldsymbol{r}) - a\, (4(\boldsymbol{R} - \boldsymbol{R}_0)^2 + (\boldsymbol{r} - \boldsymbol{r}_0)^2)].$$

However, the calculation of cross sections is independent of the particular form of the initial wave packets, and the error made in starting from a wave packet of the type $\Phi\varphi$ is negligible. This was implicitly assumed in the discussion of § X.7.

where $\psi_{\boldsymbol{b}}^{(d)}(\boldsymbol{r}, t)$ represents the divergent wave packet [eq. (X.20)]

$$\psi_{\boldsymbol{b}}^{(d)}(\boldsymbol{r}, t) \sim \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{b}-\mathrm{i}Et/\hbar}\,\chi[\boldsymbol{u}(r-vt)+\boldsymbol{s}-\boldsymbol{b}]\,f(\theta, \varphi)\,\frac{\mathrm{e}^{\mathrm{i}kr}}{r}. \quad \text{(XIV.49)}$$

Since we are working in the center-of-mass system, the wave packet $\Phi(\boldsymbol{R}, t)$ remains concentrated at the origin and the cross section $\sigma^{(1)}(\Omega)$ for the scattering of particle 1 in the direction $\Omega = (\theta, \varphi)$ is equal to the cross section for the scattering of the relative particle in the same direction, i.e. (§ X.6)

$$\sigma^{(1)}(\Omega) = |f(\theta, \varphi)|^2. \quad \text{(XIV.50)}$$

The cross section $\sigma^{(2)}(\Omega)$, for the scattering of particle 2 in the direction Ω is equal to the cross section for the scattering of the relative particle in the opposite direction, i.e.

$$\sigma^{(2)}(\Omega) = |f(\pi-\theta, \varphi+\pi)|^2. \quad \text{(XIV.51)}$$

When the two particles are identical, two important modifications of this treatment are required:

(a) the detectors no longer distinguish between particles 1 and 2, and consequently the cross section $\sigma(\Omega)$ must be redefined;

(b) the function $\Psi_{\boldsymbol{b}}$ must be properly symmetrized.

Modification (a) is not peculiar to Quantum Mechanics. Let us define $\sigma(\Omega)$ as the number of particles (1 and 2) emitted into the solid angle $(\Omega, \Omega+\mathrm{d}\Omega)$ per unit time and per unit incident flux, i.e. [cf. eq. (XIV.5)]

$$\sigma(\Omega) = \sigma^{(1)}(\Omega)+\sigma^{(2)}(\Omega). \quad \text{(XIV.52)}$$

Note that with this definition

$$\sigma^{(\mathrm{tot})} = \tfrac{1}{2}\int \sigma(\Omega)\,\mathrm{d}\Omega \quad \text{(XIV.53)}$$

if we retain the usual definition of the total cross section $\sigma^{(\mathrm{tot})}$, namely the number of particles eliminated from the incident beam per unit time and per unit incident flux.

Modification (b), on the other hand, is a specifically quantum effect. The properly symmetrized wave packet which reproduces the same initial conditions as $\Psi_{\boldsymbol{b}}(\boldsymbol{R}, \boldsymbol{r}, t)$ is, as will now be shown, given by the expression

$$\hat{\Psi}_{\boldsymbol{b}}(\boldsymbol{R}, \boldsymbol{r}, t) = \frac{1}{\sqrt{2}}\,[\Psi_{\boldsymbol{b}}(\boldsymbol{R}, \boldsymbol{r}, t) \pm \Psi_{\boldsymbol{b}}(\boldsymbol{R}, -\boldsymbol{r}, t)].$$

Before collision, i.e. for $t \ll 0$, this function $\hat{\Psi}_b$ takes the form $\Phi(\boldsymbol{R}, t)\,\hat{\varphi}_b(\boldsymbol{r}, t)$ where [eq. (XIV.47)]

$$\hat{\varphi}_b(\boldsymbol{r}, t) = e^{-i\boldsymbol{k}\cdot\boldsymbol{b} - iEt/\hbar}\frac{1}{\sqrt{2}}[\chi(\boldsymbol{r}-\boldsymbol{v}t-\boldsymbol{b})\,e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \pm \chi(-\boldsymbol{r}-\boldsymbol{v}t-\boldsymbol{b})\,e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}]$$

$\Phi\hat{\varphi}$ is thus the sum of two waves. The first represents the initial state of a collision in the center-of-mass system where particle 1 impinges upon particle 2 with the relative velocity $\boldsymbol{v}$ and the impact parameter $\boldsymbol{b}$; the second represents the state in which the roles of particles 1 and 2 are interchanged. Since these two waves do not overlap (since $t \ll 0$) and since both are of norm $\frac{1}{2}$, $\Phi\hat{\varphi}$ is a wave function of norm 1 representing, like $\Phi\varphi$, two particles impinging upon one another in their center-of-mass system with relative velocity $\boldsymbol{v}$ and impact parameter $\boldsymbol{b}$.

After collision $\hat{\Psi}_b$, like Ψ_b, takes the form of a sum of two terms [cf. eq. (XIV.48)]. The first, $\Phi\hat{\varphi}$, represents the transmitted wave and does not enter into the calculation of the cross section; the second, $\Phi\hat{\psi}^{(d)}$, represents the scattered wave and is obtained from $\Phi\psi^{(d)}$ by effecting upon $\psi^{(d)}$ the same operation of symmetrization as was used to pass from Ψ_b to $\hat{\Psi}_b$:

$$\hat{\psi}_b{}^{(d)}(\boldsymbol{r}, t) = \frac{1}{\sqrt{2}}[\psi_b{}^{(d)}(\boldsymbol{r}, t) \pm \psi_b{}^{(d)}(-\boldsymbol{r}, t)].$$

Going back to the asymptotic expression (XIV.49), we see that the passage from $\Phi\psi^{(d)}$ to $\Phi\hat{\psi}^{(d)}$ is accomplished by replacing the scattering amplitude $f(\theta, \varphi)$ by the symmetrized amplitude

$$\hat{f}(\theta, \varphi) = \frac{1}{\sqrt{2}}[f(\theta, \varphi) \pm f(\pi-\theta, \varphi+\pi)].$$

From the expression for the scattered wave packets we deduce the scattering cross section in the same way as when the particles are distinguishable. This gives

$$\sigma^{(1)}(\Omega) = |\hat{f}(\theta, \varphi)|^2 = \tfrac{1}{2}\,|f(\theta, \varphi) \pm f(\pi-\theta, \varphi+\pi)|^2$$
$$\sigma^{(2)}(\Omega) = |\hat{f}(\pi-\theta, \varphi+\pi)|^2 = \sigma^{(1)}(\Omega)$$

whence, in accordance with definition (XIV.52),

$$\sigma(\Omega) = 2\,|\hat{f}(\theta, \varphi)|^2 = |f(\theta, \varphi) \pm f(\pi-\theta, \varphi+\pi)|^2. \qquad \text{(XIV.54)}$$

Recall that the scattering amplitude is the coefficient of the outgoing wave in the stationary solution of the Schrödinger equation

$$\frac{\hbar^2}{m}(\triangle + k^2)\,\psi(\boldsymbol{r}) = V(\boldsymbol{r})\,\psi(\boldsymbol{r})$$

that has the asymptotic form

$$e^{i\boldsymbol{k}\cdot\boldsymbol{r}} + f(\theta, \varphi)\,\frac{e^{ikr}}{r}.$$

The symmetrized scattering amplitude $\hat{f}(\theta, \varphi)$ is equal to $\sqrt{2}$ times the even part of $f(\theta, \varphi)$ or to $\sqrt{2}$ times the odd part of $f(\theta, \varphi)$, according as the particles are bosons or fermions. If V is a central potential, the even part of $f(\theta)$ is the sum of contributions from partial waves of even order, the odd part the sum of contributions of partial waves of odd order. At energies sufficiently low that the only important contribution is from the s wave, two (spinless) fermions are hardly scattered by one another at all, while two bosons have a scattering cross section four times greater than if they were distinguishable (the scattering amplitude is multiplied by $\sqrt{2}$ and the cross section is equal to twice the squared modulus of this symmetrized amplitude; the combination of these two effects gives the factor 4).

10. Collision of two Protons

The treatment given in the preceding paragraph can easily be extended to the collision of two identical particles with spin. As an example we consider the collision of two protons, assuming a central interaction potential. The total spin is a constant of the motion, but the interaction may differ in the singlet and triplet states; let $f_s(\theta)$ and $f_t(\theta)$ denote the corresponding (un-symmetrized) scattering amplitudes.

Since protons are fermions, the wave function describing a two-proton system must be antisymmetrical in the interchange of the two particles. If this function represents a triplet state, it will be symmetrical in the exchange of the spins, hence antisymmetrical in the exchange of $\boldsymbol{r}^{(1)}$ and $\boldsymbol{r}^{(2)}$, and the symmetrized scattering amplitude is

$$\hat{f}_t(\theta) = \frac{1}{\sqrt{2}}[f_t(\theta) - f_t(\pi - \theta)].$$

The cross section for mutual scattering of protons in the triplet state is therefore

$$\sigma_t(\Omega) = 2\,|\hat{f}_t(\theta)|^2 = |f_t(\theta) - f_t(\pi-\theta)|^2.$$

If this function represents a singlet state, it will be antisymmetrical in the exchange of the spins, hence symmetrical in the exchange of the spatial coordinate, and the symmetrized scattering amplitude is

$$\hat{f}_s(\theta) = \frac{1}{\sqrt{2}}\,[f_s(\theta) + f_s(\pi-\theta)].$$

The cross section for mutual scattering of two protons in the singlet state is therefore

$$\sigma_s(\Omega) = 2\,|\hat{f}_s(\theta)|^2 = |f_s(\theta) + f_s(\pi-\theta)|^2.$$

If the target and the incoming beam are both composed of unpolarized protons, in each collision there will be a random distribution of the spins of the incident and struck particles. Since the space of the triplet states has three dimensions, and that of singlet states one, the total spin of the initial state has a probability $\frac{3}{4}$ of being triplet and $\frac{1}{4}$ of being singlet; therefore

$$\begin{aligned}\sigma(\Omega) &= \tfrac{3}{4}\sigma_t(\Omega) + \tfrac{1}{4}\sigma_s(\Omega) \\ &= \tfrac{3}{4}\,|f_t(\theta) - f_t(\pi-\theta)|^2 + \tfrac{1}{4}\,|f_s(\theta) + f_s(\pi-\theta)|^2.\end{aligned} \tag{XIV.55}$$

If in addition the forces are independent of spin

$$f_s(\theta) = f_t(\theta) = f(\theta),$$

and we find:

$$\sigma(\Omega) = |f(\theta)|^2 + |f(\pi-\theta)|^2 - \tfrac{1}{2}\,[f^*(\theta)\,f(\pi-\theta) + f(\theta)\,f^*(\pi-\theta)]. \tag{XIV.56}$$

In particular, if we are at sufficiently low energies that nuclear forces can be neglected and $V(r)$ treated as a repulsive Coulomb potential e^2/r, the amplitude $f(\theta)$ will be given by formula (XI.33) and the scattering cross section by the *Mott formula*

$$\sigma(\Omega) = \left(\frac{e^2}{4E}\right)^2 \left[\sin^{-4}\frac{\theta}{2} + \cos^{-4}\frac{\theta}{2} - \sin^{-2}\frac{\theta}{2}\cos^{-2}\frac{\theta}{2}\cos\left(\frac{e^2}{\hbar v}\ln\left(\tan^2\frac{\theta}{2}\right)\right)\right], \tag{XIV.57}$$

in which E is the energy in the center-of-mass system and v the

relative velocity of the two protons. It is instructive to compare this with the Rutherford formula [eq. (XI.36)].

A classical calculation would give only the first two terms of the Mott formula:

$$\sigma_{\mathrm{cl}}(\Omega) = \left(\frac{e^2}{4E}\right)^2 \left[\sin^{-4}\frac{\theta}{2} + \cos^{-4}\frac{\theta}{2}\right].$$

The third term is due to the purely quantum effect of interference between the scattering amplitudes $f(\theta)$ and $f(\pi-\theta)$. When $e^2/\hbar v \gg 1$ it oscillates rapidly about 0, the more rapidly as we move away from the angle $\theta = \frac{1}{2}\pi$ (in either direction). In the limit $\hbar \to 0$, these oscillations, while maintaining the same amplitude, become increasingly rapid, so that the average value of $\sigma(\Omega)$, taken over a small but non-zero solid angle $\varrho(\Omega)$, tends to the cross section $\sigma_{\mathrm{cl}}(\Omega)$.

11. Statistics of Atomic Nuclei

An atomic nucleus may in many problems be treated as a particle endowed with a certain spin j.

In the domain of atomic physics this is an excellent approximation. An atomic nucleus is an assembly of $\mathcal{N}$ nucleons: Z protons and N neutrons ($\mathcal{N} = Z+N$), whose dynamical variables are functions of the fundamental variables of its $\mathcal{N}$ constituents. Let $\boldsymbol{R}$ and $\boldsymbol{P}$ be the position and momentum of the center of mass (cf. § IX.13) and let ϱ denote the ensemble of the internal variables – nucleon spins and relative positions and momenta of nucleons with respect to one another. Among these we have notably the total angular momentum $\boldsymbol{j}$ of the internal variables, the so-called *spin of the nucleus*: $\boldsymbol{j}$ is the sum of the $\mathcal{N}$ individual nucleon spins and the $(\mathcal{N}-1)$ relative orbital angular momenta; according to the law for the addition of angular momenta, j may take integral or half-integral values according as $\mathcal{N}$ is even or odd.

In the absence of an external field, the motion of the center of mass separates out, and the Hamiltonian of the system becomes the sum of two terms: the kinetic energy of the center of mass $\boldsymbol{P}^2/2M$ (M is the total mass of the nucleus) and the internal energy $h(\varrho)$ comprising the kinetic energy of the nucleons and their energy of mutual interaction. h has a certain number of bound states. We denote by ε_0 the eigenvalue corresponding to the ground state. Since the forces between nucleons are invariant under rotation, $h(\varrho)$ commutes with the three components of $\boldsymbol{j}$ (cf. Ch. XIII). Each eigenvalue of its discrete spectrum corresponds to a definite value j of the spin and has a degeneracy of order $(2j+1)$. Henceforth,

j will denote the spin of the ground state, μ the possible values of j_z ($\mu = -j, -j+1, \ldots, +j$) and χ_μ the eigenfunction for the ground state of angular momentum $(j\mu)$:

$$h(\varrho)\,\chi_\mu = \varepsilon_0\,\chi_\mu \qquad j_z\,\chi_\mu = \mu\,\chi_\mu.$$

The $(2j+1)$ vectors χ_μ differ only in their orientation. They can be obtained one from another by repeated application of j_+ and j_-, and form a possible basis for the representation of the vector operator $\boldsymbol{j}$. Nuclear radii are of the order of 10^{-13} to 10^{-12} cm, and consequently the average inter-nucleon distance in a nucleus in the state χ_μ is a distance of that order. So long as the nucleus remains in its ground state the wave function of the system is of the form $\sum_\mu \psi(\boldsymbol{R}, \mu)\,\chi_\mu$ and we are able to treat it as a particle of spin j whose wave function in the $\{\boldsymbol{R}, j_z\}$ representation is $\psi(\boldsymbol{R}, \mu)$ since $h(\varrho)$ can be replaced in the Hamiltonian by the constant ε_0 and $\psi(\boldsymbol{R}, \mu)$ evolves in time like the wave function of a free particle of mass M (to within a constant phase factor without physical significance).

If the nucleus is placed in an external field, for example an electric field with potential $V(\boldsymbol{r})$, the Hamiltonian will include, in addition to the terms already mentioned, terms for the interaction of the Z protons with the external field. It may now be only approximately treated as a particle of spin j, for the Hamiltonian no longer commutes with $h(\varrho)$. It remains, however, an excellent approximation if the field varies little over distances of the order of a nuclear radius, since one can then replace the value of the field at each proton by its value at the center of mass; the external interaction then becomes $ZeV(\boldsymbol{R})$ and the Hamiltonian has eigenstates of the form $\sum_\mu \psi(\boldsymbol{R}, \mu)\,\chi_\mu$, where $\psi(\boldsymbol{R}, \mu)$ is a stationary state of a particle of mass M, spin j and charge Ze in the electric potential [1]) $V(\boldsymbol{R})$.

Consider a system of n nuclei, all in their ground state. It is a perfectly justifiable approximation to treat each as a particle of given spin so long as they remain sufficiently far apart, as is the case in a molecule or a solid. If we denote by $\boldsymbol{R}^{(i)}$ the position of the ith nucleus and by $\mu^{(i)}$ the z-component of its spin, the wave function of the system is a certain function of the $\boldsymbol{R}^{(i)}$ and the $\mu^{(i)}$, whose motion is determined by a certain Hamiltonian that depends on the variables $\boldsymbol{R}^{(i)}$, $\boldsymbol{P}^{(i)}$ and $\boldsymbol{j}^{(i)}$.

---

[1]) The first corrections to this approximation preserve this picture of the nucleus; they consist in attributing to it a quadrupole moment (if $j \geqslant 1$) which can be expressed as a function of the components of $\boldsymbol{j}$, thus leading to a coupling of the nuclear spin with the external field.

Let us examine to what extent this simplified description will be affected by the symmetrization postulate. We may expect that symmetrization will be involved only when some of the nuclei are identical. This can be proved by an argument analogous to the one of § 8. Two different nuclei are distinguishable particles in spite of the fact that their constituents, neutrons and protons, are identical.

Let us therefore consider a system of two identical nuclei of spin j. We describe the state of the system by the wave function

$$\psi(\mathbf{R}^{(1)}, \mu^{(1)}; \mathbf{R}^{(2)}, \mu^{(2)}).$$

The interchange of the two nuclei is the operation P defined by

$$P\psi = \psi(\mathbf{R}^{(2)}, \mu^{(2)}; \mathbf{R}^{(1)}, \mu^{(1)}).$$

In actual fact the system contains $2Z$ protons and $2N$ neutrons and its dynamical state is obtained by antisymmetrizing with respect to the protons and with respect to the neutrons the vector

$$\Psi = \sum_{\mu_1\mu_2} \psi(\mathbf{R}^{(1)}, \mu_1; \mathbf{R}^{(2)}, \mu_2)\, \chi^{(1)}_{\mu_1}\, \chi^{(2)}_{\mu_2}.$$

The operation P consists in interchanging the protons and the neutrons of the first nucleus with the protons and the neutrons of the second, in all $\mathcal{N}$ elementary interchanges; since in each of these the correctly antisymmetrized state vector changes sign, we can deduce that

$$P\Psi = (-)^{\mathcal{N}}\, \Psi,$$

and that therefore

$$P\psi = (-)^{\mathcal{N}}\, \psi.$$

This may easily be generalized to a system containing more than two identical nuclei. The wave function must be symmetrical or antisymmetrical with respect to permutations of the identical nuclei, according to whether these nuclei contain an even or an odd number of nucleons.

In other words, atomic nuclei are:

(a) *bosons* if they contain an *even* number of nucleons;

(b) *fermions* if they contain an *odd* number of nucleons.

These differences of statistics make a spectacular appearance in many phenomena where purely nuclear effects would seem to be unimportant. Such is the case with the remarkable properties of

liquid helium (He^4) at very low temperatures. He^4 obeys Bose–Einstein statistics; the isotope He^3 obeys Fermi–Dirac statistics and behaves at these temperatures in a fundamentally different way.

Another example, to which we shall presently return (§ XVIII.17), is provided by the band spectrum of homonuclear diatomic molecules.

12. Complex Atoms. Central Field Approximation

The Pauli principle has a profound effect on the spectra of complex atoms.

In the absence of an external field, the Hamiltonian of an atom is independent of the spins of its Z electrons and is given by equation (XIII.71). To be exact, one should add terms for the spin-orbit interaction, but these can be neglected for the purpose of the present discussion. Except for the special case of the hydrogen atom ($Z=1$), the eigenvalue problem for such a Hamiltonian cannot be exactly solved.

For determining the stationary states of the atom, we often make use of the independent-particle or *central-field approximation*, according to which each electron moves independently of the others in a central potential $V(r)$ representing the attraction of the nucleus and the averaged repulsive effect of the other electrons. This last-mentioned effect clearly depends on the dynamical state of the electrons and thus a single potential $V(r)$ cannot, even approximately, account for the whole spectrum of the atom. However, if we restrict our study to the ground and first excited states, $V(r)$ can be fixed once and for all, and the more judicious the choice, the better the approximation. The overall effect of the electrons is to put up a screen about the nuclear Coulomb field, the effect becoming more appreciable as we move away from the nucleus: $V(r)$, which is of the form $-Ze^2/r$ near the origin, increasingly deviates from this pure Coulomb form with increasing r to become $-e^2/r$ in the asymptotic region. These semi-quantitative considerations suffice for the present. Further on we discuss two methods for systematically determining $V(r)$, the methods of Thomas–Fermi (§ 13), and Hartree–Fock (Ch. XVIII).

In the central field approximation, the Hamiltonian is written

$$H = h^{(1)} + h^{(2)} + \ldots + h^{(Z)}, \tag{XIV.58}$$

where

$$h = \frac{\mathbf{p}^2}{2m} + V(r).$$

The eigenvectors of H are $Z \times Z$ Slater determinants which can be built up from a set of basis vectors for h. The eigenvalue of H corresponding to a given Slater determinant is equal to the sum of the energies of the Z individual states that appear in it. Thus the eigenvalue problem for H is easily solved once we have the solution of the eigenvalue problem for the individual Hamiltonian h.

h is the Hamiltonian of a particle of spin $\frac{1}{2}$ in a spin-independent central potential. The solution of the corresponding eigenvalue problem for a spinless particle has been given in Chapter IX. The only effect of the spin is to double the degeneracy of each level. h, $\mathbf{l}^2$, l_z and s_z constitute a complete set of commuting observables whose basis vectors $|n\, l\, m_l\, m_s\rangle$ are labelled with the four quantum numbers n, l, m_l, m_s: the spin quantum number, m_s, can take the values $\pm \frac{1}{2}$; n, the principle quantum number, has the same definition as in the hydrogen atom problem. (The number of nodes of the radial function is $n-l-1$.) Since the energy of each state, ε_{nl}, depends only on n and l, each individual level is $2(2l+1)$-fold degenerate.

The order of the levels ε_{nl} in the energy spectrum does not depend crucially on the form of the potential $V(r)$. For a given l they are in order of increasing n. If $V(r)$ was simply the Coulomb field of the nucleus, all the levels corresponding to a given value of n ($l=0, 1, \ldots, n-1$) would coincide (cf. Fig. XI.1, p. 418). The screening effect of the other electrons results in a raising of these levels that varies as the average distance of the electron from the nucleus, and which therefore increases with n and l. If we limit our study to the ground and first excited states, the order of succession of individual levels is nearly the same for all atoms, namely

1s	2s	2p	3s	3p	[4s, 3d]	4p	[5s, 4d]	5p	[6s, 4f, 5d]	6p	[7s, 5f, 6d]
2	2	6	2	6	2+10	6	2+10	6	2+14+10	6	2+14+10.

The bracketed levels nearly coincide and their order can vary from one atom to another. The number under each term is the degeneracy of the corresponding level.

To each set of Z different individual states there corresponds a Slater determinant, and therefore a stationary state of the atom. The energy of that state, being the sum of the energies of its constituent individual states, depends only on the number of electrons occupying each of the levels ε_{nl}. The specification of the occupation number of each individual level defines a *configuration*: with this

definition, states belonging to the same configuration have the same energy.

Let ν_i be the number of electrons occupying the level ε_i. ν_i is at most equal to the degeneracy g_i of the level. If $\nu_i = g_i$, the electrons form a *closed shell*; if $\nu_i < g_i$, the shell is *incomplete*. There are

$$g_i!/\nu_i!\,(g_i-\nu_i)!$$

ways to distribute the ν_i electrons among the g_i individual states of the level, Therefore, nearly all configurations are degenerate – the only ones that are not are those formed exclusively of closed shells.

The ground state configuration is formed by putting the Z electrons in the lowest energy levels. They are thus distributed among a certain number h of shells, the first $(h-1)$ being filled, the last one generally not, except for very special values of Z ($Z=2$, 4, 10, 12, 18 etc.).

Consider the carbon atom for example ($Z=6$). In the ground state, shells 1s and 2s are closed, and the two remaining electrons are in the 2p shell. As there are $\binom{2}{6}=15$ ways of distributing these 2 electrons among the 6 states available in that shell, the ground state of the carbon atom (in the independent particle approximation) is 15-fold degenerate. In the case of neon ($Z=10$), the ground state is made up of the three closed shells 1s, 2s and 2p, and therefore is non-degenerate.

The electrons in the lowest shells are also those nearest the nucleus. The chemical properties of atoms are practically independent of the motion of these electrons. At the low energies involved in chemical reaction, the interactions between atoms depend essentially, if not exclusively, on the motion of the electrons in the outer shells, hence, two atoms having outside shells of similar electronic structure will have similar chemical properties. Thus the outside shell of each of the rare-air gases (Ne, A, Kr, Xe) is a closed p shell; that of each of the halogens (F, Cl, Br, I), a p shell lacking but one electron; and that of the alkalies, (Na, K, Rb, Cs), an s shell containing one electron with a closed p shell just below. Quite generally, the position of each chemical element in the Periodic Table can easily be predicted if we know the order of succession of the electronic shells [1]).

---

[1]) For further details of the quantum mechanical explanation of the chemical properties of atoms, see: M. Born, *Atomic Physics*, 6th ed. (Blackie, Glasgow, 1957); L. Pauling and E. B. Wilson, *Introduction to Quantum Mechanics* (McGraw-Hill, New York, 1935). On the general theory of atomic spectra, see: E. U.

13. The Thomas–Fermi Model of the Atom

When $Z \gg 1$, one can determine the ground state potential by applying a semi-classical method due to Thomas and Fermi.

Let $\varrho(r)$ be the probability density for finding an electron in the volume element $(r, r+dr)$ when the atom is in its ground state. We suppose this function to be spherically symmetrical. It satisfies the normalization condition

$$4\pi \int_0^\infty \varrho(r) r^2 \, dr = Z. \tag{XIV.59}$$

The Z electrons form about the nucleus a cloud of negative electricity of average density $-e\varrho(r)$. The charges in the atom thus give rise to an average electrical potential $\Phi(r)$ whose sources are:

(i) the point charge of the nucleus, situated at the origin and equal to Ze;

(ii) a continuous distribution of electricity of density $-e\varrho(r)$.

In mathematical language, Φ is the solution of the Poisson equation

$$\Delta\Phi \equiv \frac{1}{r}\left(\frac{d^2}{dr^2} r\right)\Phi = 4\pi e\varrho \tag{XIV.60}$$

that has the following behavior at the origin:

$$\lim_{r\to 0} r\Phi = Ze. \tag{XIV.61}$$

In the limit $Z \gg 1$, the electric field due to an electron is small compared to that of all the others, and we can represent the potential acting on each electron in the independent particle approximation by $-e\Phi(r)$.

The ground state of the atom is the state in which the Z electrons occupy the Z lowest quantum states of a particle of mass m in the field $-e\Phi$. The density $\varrho(r)$ is thus the sum of the densities $|\psi|^2$ of the Z lowest levels. This implies that a functional relation exists between $\varrho(r)$ and the potential $-e\Phi$. To determine this relation, we turn to the following "*semi-classical*" *approximation.*

In the classical limit, the number of stationary states in the energy

---

Condon and G. H. Shortly, *The Theory of Atomic Spectra*, 4th ed. (Cambridge University Press, 1957); G. Racah, Phys. Rev. **62** (1942) 438 and **63** (1943) 367.

band $(\varepsilon, \varepsilon+\delta\varepsilon)$ is proportional to the volume occupied by this band in the phase space of the corresponding classical particle; the proportionality factor, $2/h^3$, is double that given in § VI.11, owing to the electron having two spin states. When the Z lowest quantum states are occupied, the energy distribution of the electrons in the atom is thus the same as that of a statistical mixture of Z classical electrons having a density in phase space $n(\boldsymbol{r}, \boldsymbol{p})$ equal to

$$n(\boldsymbol{r}, \boldsymbol{p}) = \begin{cases} \dfrac{2}{h^3} & \text{if} \quad \varepsilon \equiv \dfrac{p^2}{2m} - e\Phi < \varepsilon_0 \\ 0 & \text{if} \quad \varepsilon > \varepsilon_0. \end{cases} \tag{XIV.62}$$

ε_0 is the energy of the highest level occupied. Since the zero of energy may be arbitrarily fixed, we put $\varepsilon_0 = 0$.

In the spirit of the classical approximation, we assume that the electrons have the same spatial distribution as this classical statistical mixture:

$$\varrho(r) = \int n(\boldsymbol{r}, \boldsymbol{p})\, \mathrm{d}\boldsymbol{p} = \frac{2}{h^3} \int_{\varepsilon<0} \mathrm{d}\boldsymbol{p}.$$

Replacing ε by its expression in terms of $\boldsymbol{p}$ and Φ, we find, after a simple integration,

$$\varrho(r) = \begin{cases} \dfrac{8\pi}{3h^3} (2me\Phi)^{3/2} & \text{if} \quad \Phi > 0 \\ 0 & \text{if} \quad \Phi < 0\,. \end{cases} \tag{XIV.63}$$

Substitution of (XIV.63) into the right-hand side of (XIV.60) gives a second-order differential equation for Φ. With (XIV.59) and (XIV.61) the function is thereby completely defined. Relations (XIV.61), (XIV.63), (XIV.59), and (XIV.62) are the fundamental relations of the Thomas–Fermi model.

To deduce Φ and ϱ from these relations, it is convenient to make the following changes of variable and function:

$$r = Z^{-1/3} b x \qquad \Phi = \frac{Ze}{r} \chi, \tag{XIV.64}$$

where

$$b = \frac{1}{2} \left(\frac{3\pi}{4}\right)^{2/3} \frac{\hbar^2}{me^2} \simeq 0.5 \times 10^{-8} \text{ cm}. \tag{XIV.65}$$

From (XIV.63) we get ϱ as a function of the dimensionless quantities χ and x:

$$\varrho = \begin{cases} \dfrac{Z^2}{4\pi b^3}\left(\dfrac{\chi}{x}\right)^{3/2} & \text{if} \quad \chi > 0 \\ 0 & \text{if} \quad \chi < 0. \end{cases} \tag{XIV.66}$$

The fundamental equation (XIV.60) is equivalent to the equation

$$\frac{\mathrm{d}^2\chi}{\mathrm{d}x^2} = \begin{cases} x^{-1/2}\,\chi^{3/2} & \text{if} \quad \chi > 0 \\ 0 & \text{if} \quad \chi < 0. \end{cases} \tag{XIV.67}$$

Condition (XIV.61) gives $\chi(0) = 1$. From eq. (XIV.67), $\chi(x)$ obviously has at most one zero in the interval $(0, \infty)$; if x_0 is the position of this zero, χ will be positive in the interval $(0, x_0)$ and negative in the interval (x_0, ∞). Hence, taking into account (XIV.64), (XIV.66) and (XIV.67), (XIV.59) becomes

$$1 = \int_0^{x_0} \sqrt{x}\,\chi^{3/2}\,\mathrm{d}x = \int_0^{x_0} x\,\chi''\,\mathrm{d}x = x\,\chi' - \chi\Big|_0^{x_0} = x_0\,\chi'(x_0) + 1.$$

This condition requires the derivative χ' to vanish at the same point as does χ itself, which means that the point in question must be at infinity.

Thus, $\chi(x)$ is the solution of

$$\chi'' = x^{-1/2}\,\chi^{3/2} \tag{XIV.68}$$

that satisfies the conditions

$$\chi(0) = 1 \qquad \chi(\infty) = 0. \tag{XIV.69}$$

$\chi(x)$ must be found by numerical integration, and is given in Fig. XIV.1. Knowing $\chi(x)$, we can find $\varrho(r)$ and $\Phi(r)$.

For the classical approximation to be justified, a large majority of the Z individual electron states must be in the "large quantum number region" i.e. $Z \gg 1$. For a given atom, the electron density $\varrho(r)$ and the electrostatic potential $\Phi(r)$ given by the Thomas–Fermi model are those which obtain in the limit when the quantum of action $\hbar$ and the charge $(-e)$ for each electron become infinitesimal while the number of electrons Z becomes infinite, the characteristic

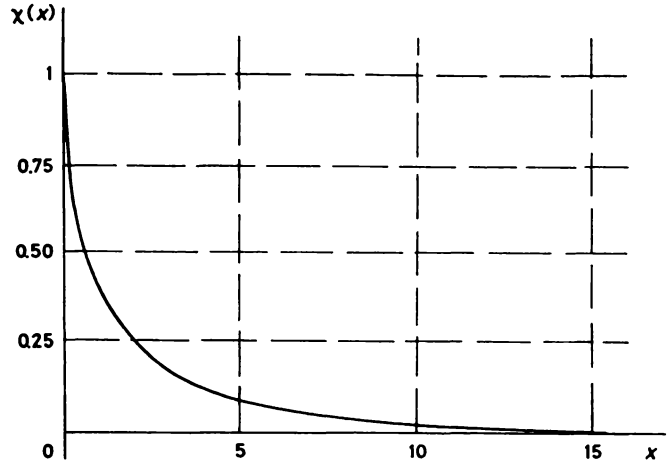


Fig. XIV.1. The Thomas–Fermi $\chi(x)$ function

$$\chi'' = x^{-1/2}\,\chi^{3/2}, \qquad \chi(0) = 1, \qquad \chi(\infty) = 0$$

for any value of Z, using eqs. (XIV.64–66).

length $\hbar^2/me^2$ and the total charge of the electronic cloud $(-Ze)$ remaining constant [1]).

The Thomas–Fermi model permits an estimation of the atomic radius. This quantity needs to be defined, since the electronic density becomes zero only at infinity and therefore the atom is not an object occupying a well-defined region in space. By atomic radius we shall mean the radius $R(\alpha)$ of the sphere centered at the origin and containing a given fraction $(1-\alpha)$ of the Z atomic electrons. Thus, according to this definition,

$$(1-\alpha)\,Z = 4\pi \int_0^R \varrho(r)\, r^2\, \mathrm{d}r.$$

We put

$$R(\alpha) = Z^{-1/3}\, b\, X(\alpha).$$

---

[1]) The Thomas–Fermi method over-estimates the electron density near the origin ($r \lesssim b/Z$), and also in the asymptotic region ($r \gg b$); at the origin $\varrho(r)$ diverges like $r^{-3/2}$ instead of remaining finite, at infinity it goes to zero like r^{-6}, rather than exponentially.

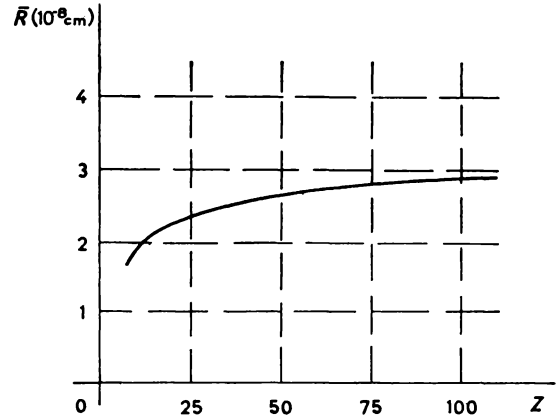


Fig. XIV.2. Z dependence of the radius $\bar{R}$ of the Thomas–Fermi atom:

$$4\pi \int_0^{\bar{R}} \varrho(R)\, r^2\, \mathrm{d}r = Z - 1.$$

A brief calculation, with use of (XIV.64), (XIV.66), and (XIV.68), results in the following equation for X:

$$\chi(X) - X\,\chi'(X) = \alpha.$$

This equation must be numerically solved.

If we adopt the same value of α for all atoms, X will be the same for them all, and the atomic radius will be proportional to $Z^{-1/3}$.

For $\alpha = 1/Z$, the corresponding radius

$$\bar{R} \equiv R\left(\frac{1}{Z}\right) = Z^{-1/3}\, b\, X\left(\frac{1}{Z}\right)$$

is the radius of the sphere containing all but one of the electrons. Figure XIV.2 gives $\bar{R}$ as a function of Z; observe that it is practically independent of Z ($\bar{R} \simeq 2\text{–}3 \times 10^{-8}$ cm).

14. Nucleon Systems and Isotopic Spin

Let us consider a system of Z protons and N neutrons. The motion of the system will be governed by a certain Hamiltonian H. Neutrons and protons are similar particles, and to distinguish between them we shall number the protons from 1 to Z and the neutrons from $Z+1$ to $Z+N \equiv \mathcal{N}$; A_Z will denote the antisymmetrizing operator

for the first Z particles, A_N that for the N last; we keep the notation of § 2–4. Let

$$\varphi(q^{(1)}, q^{(2)}, \ldots, q^{(Z)}, q^{(Z+1)}, \ldots, q^{(\mathcal{N})})$$

be a wave function representing a possible state of the system in the symmetrical representation $\{Q\}$. φ is subject to the antisymmetrization conditions

$$A_N \varphi = \varphi \qquad A_Z \varphi = \varphi. \tag{XIV.70}$$

There exists another formalism, strictly equivalent to the foregoing, to describe this system of Z protons and N neutrons. In this second formalism, the neutron and the proton are no longer considered as two different particles, but as two different states of the same particle, the *nucleon*. The system in question is described as a system of $\mathcal{N}$ nucleons, of which Z are in the proton state and N in the neutron state; the $\mathcal{N}$ nucleons are then identical fermions and the states of the system are antisymmetrical with respect to permutations of these $\mathcal{N}$ identical fermions. The object of the present paragraph is to describe this new formalism and to prove its equivalence to the ordinary one.

To distinguish between the proton state and the neutron state, each nucleon must be assigned a new dynamical variable — its charge — admitting of two possible values; we shall denote the corresponding eigenstates by ϖ and ν, ϖ representing the proton state, ν the neutron state. The charge space of the nucleon is a two-dimensional space like spin-space. We can therefore define operators for it which are analogous to those introduced for spin-space and which have the same mathematical properties. Consider in particular the three Pauli matrices; in the representation where ϖ and ν are basis vectors, these matrices will represent a vector operator $\boldsymbol{\tau} \equiv (\tau_1, \tau_2, \tau_3)$ in charge space analogous to the vector $\boldsymbol{\sigma} \equiv (\sigma_x, \sigma_y, \sigma_z)$. The vector

$$\mathbf{t} = \tfrac{1}{2}\boldsymbol{\tau} \tag{XIV.71}$$

is the analogue of the spin $\mathbf{s}$, and is called the *isotopic spin* of the nucleon. We see that

$$t_3 \varpi = \tfrac{1}{2}\tau_3 \varpi = \tfrac{1}{2}\varpi \qquad t_3 \nu = \tfrac{1}{2}\tau_3 \nu = -\tfrac{1}{2}\nu.$$

The projectors Π_ϖ and Π_ν, onto the proton and neutron states respectively, are given by

$$\Pi_\varpi = \tfrac{1}{2}(1+\tau_3) \qquad \Pi_\nu = \tfrac{1}{2}(1-\tau_3) \tag{XIV.72}$$

and the operator representing the charge of the nucleon is given by

$$e\,\Pi_\varpi = \tfrac{1}{2}(1+\tau_3)\,e.$$

The product of the $\mathcal{N}$ individual charge spaces is the charge space $\mathscr{E}_C$ for the system of $\mathcal{N}$ nucleons. The total charge of the system is represented by the operator

$$C \equiv \sum_{i=1}^{\mathcal{N}} e\,\Pi_\varpi^{(i)} = e(\tfrac{1}{2}\mathcal{N} + T_3), \tag{XIV.73}$$

T_3 being the third component of the *total isotopic spin*

$$\boldsymbol{T} = \sum_{i=1}^{\mathcal{N}} \boldsymbol{t}^{(i)}. \tag{XIV.74}$$

We obtain an orthonormal basis in $\mathscr{E}_C$ by taking products of $\mathcal{N}$ ϖ- or ν-vectors. In particular the basis vector

$$\zeta = \varpi^{(1)}\,\varpi^{(2)} \ldots \varpi^{(Z)}\,\nu^{(Z+1)} \ldots \nu^{(\mathcal{N})}$$

represents the charge state in which the first Z particles are protons and the last N neutrons. In what follows we shall deal only with states of charge $C = Ze$, i.e. with

$$T_3 = \tfrac{1}{2}(Z-N).$$

We can construct $(\mathcal{N}!/Z!\,N!)$ basis vectors corresponding to this eigenvalue. A typical one of these,

$$\zeta_\alpha = \varpi^{(\alpha_1)}\,\varpi^{(\alpha_2)} \ldots \varpi^{(\alpha_Z)}\,\nu^{(\alpha_{Z+1})} \ldots \nu^{(\alpha_{\mathcal{N}})},$$

represents the state in which the Z particles $\alpha_1, \alpha_2, \ldots, \alpha_Z$ are protons, and the others neutrons.

The ket vectors for the system of $\mathcal{N}$ nucleons are the vectors of the space formed by taking the tensor product of $\mathscr{E}_C$ with the space $\mathscr{E}_0$ of the other dynamical variables. When we permute these $\mathcal{N}$ nucleons we effect the same permutation on charge variables and the other variables; if P_C represents a given permutation of the charges,

P_0 the same permutation of the other variables, the overall permutation will be represented by the operator

$$P = P_0 P_C.$$

The antisymmetrizing operator for the system of $\mathcal{N}$ nucleons is

$$A = \frac{1}{\mathcal{N}!} \sum_P (-)^p P = \frac{1}{\mathcal{N}!} \sum_P (-)^p P_0 P_C. \qquad \text{(XIV.75)}$$

The states of the system corresponding to a total charge Ze are those vectors Φ of the space $\mathscr{E}_C \otimes \mathscr{E}_0$ that meet the antisymmetrization condition

$$A\Phi = \Phi \qquad \text{(XIV.76)}$$

and satisfy the equation

$$T_3 \Phi = \tfrac{1}{2}(Z-N)\Phi. \qquad \text{(XIV.77)}$$

We shall now show that there is a one-to-one correspondence between the vectors Φ subject to conditions (XIV.76) and (XIV.77), and the vectors φ of $\mathscr{E}_0$ subject to conditions (XIV.70), namely

$$|\Phi\rangle = \sqrt{\frac{\mathcal{N}!}{N!\,Z!}}\, A|\varphi\rangle\,|\zeta\rangle \qquad \text{(XIV.78)}$$

$$|\varphi\rangle = \sqrt{\frac{\mathcal{N}!}{N!\,Z!}}\, \langle\zeta|\Phi\rangle, \qquad \text{(XIV.79)}$$

and that this one-to-one correspondence conserves the scalar product.

Consider a vector $|\varphi\rangle$ obeying conditions (XIV.70). The corresponding $|\Phi\rangle$ obtained from (XIV.78) obviously satisfies (XIV.76), and also (XIV.77) since T_3 commutes with A and $T_3|\zeta\rangle = \frac{1}{2}(Z-N)\,|\zeta\rangle$.

We next show that the partial scalar product (XIV.79) gives back the vector $|\varphi\rangle$. With the aid of (XIV.75) and (XIV.78),

$$\begin{aligned} \sqrt{\frac{\mathcal{N}!}{N!\,Z!}}\, \langle\zeta|\Phi\rangle &= \sqrt{\frac{\mathcal{N}!}{N!\,Z!}}\, \langle\zeta|A|\Phi\rangle \\ &= \frac{1}{N!\,Z!} \sum_P (-)^p\, P_0|\varphi\rangle\, \langle\zeta|P_C|\zeta\rangle. \end{aligned} \qquad \text{(XIV.80)}$$

Now the $\mathcal{N}!$ permutations can be put into two groups. Those of the first group, F, permute the first Z particles among themselves and/or the last N particles among themselves; the $(N!\,Z!)$ permutations of

F leave the vector $|\zeta\rangle$ unchanged and multiply the vector $|\varphi\rangle$ by $(-)^f$:

$$\langle\zeta|F_C|\zeta\rangle = 1 \qquad F_0|\varphi\rangle = (-)^f|\varphi\rangle.$$

Those of the second group, G, take the vector ζ over into another of the vectors ζ_α; hence

$$\langle\zeta|G_C|\zeta\rangle = 0.$$

Thus the sum on the right-hand side of (XIV.80) has $(N!\,Z!)$ non-zero terms, each of which is equal to $|\varphi\rangle$, giving the desired result (XIV.79).

We easily see that the correspondence (XIV.78) conserves the scalar product, for if $|\Phi\rangle$ and $|X\rangle$ correspond respectively to $|\varphi\rangle$ and $|\chi\rangle$, then since

$$|X\rangle = \sqrt{\frac{\mathcal{N}!}{N!\,Z!}}\,A|\chi\rangle\,|\zeta\rangle$$

and since $|\Phi\rangle$ obeys (XIV.76) and (XIV.78)

$$\langle X|\Phi\rangle = \sqrt{\frac{\mathcal{N}!}{N!\,Z!}}\,\langle\chi|\,\langle\zeta|\Phi\rangle = \langle\chi|\varphi\rangle.$$

It remains to show that the correspondence is of a one-to-one nature. Let $|\Phi\rangle$ be a vector subject to conditions (XIV.76) and (XIV.77) and $|\varphi\rangle$ the vector associated by (XIV.79). In view of (XIV.77), $|\Phi\rangle$ is a linear combination of the vectors $|\zeta_\alpha\rangle$, the coefficients being vectors of $\mathscr{E}_0$:

$$|\Phi\rangle = \sqrt{\frac{\mathcal{N}!}{N!\,Z!}}\sum_\alpha |\varphi_\alpha\rangle\,|\zeta_\alpha\rangle.$$

$|\varphi\rangle$ is the coefficient of ζ in the sum on the right-hand side. A permutation of type F acting on Φ gives

$$F|\Phi\rangle = \sqrt{\frac{\mathcal{N}!}{N!\,Z!}}\sum_\alpha (F_0|\varphi_\alpha\rangle)\,(F_C|\zeta_\alpha\rangle). \qquad \text{(XIV.81)}$$

also

$$F|\Phi\rangle = (-)^f|\Phi\rangle = \sqrt{\frac{\mathcal{N}!}{N!\,Z!}}\sum_\alpha (-)^f|\varphi_\alpha\rangle\,|\zeta_\gamma\rangle. \qquad \text{(XIV.82)}$$

Now the action of F_C on one of the $|\zeta_\alpha\rangle$ is to transform it into another of the $|\zeta_\alpha\rangle$ and in particular it leaves $|\zeta\rangle$ invariant. Therefore, the

coefficient of $|\zeta\rangle$ in the expansion of $F|\Phi\rangle$ is the coefficient of $F_C|\zeta\rangle$ on the right-hand side of (XIV.81). Equating this to the coefficient of $|\zeta\rangle$ in (XIV.82), we find

$$F_0|\varphi\rangle = (-)^f|\varphi\rangle.$$

which shows that the vector $|\varphi\rangle$ corresponding to $|\Phi\rangle$ does indeed have the antisymmetry properties (XIV.70). Q.E.D.

Any state of Z protons and N neutrons represented by a vector $|\varphi\rangle$ in the ordinary formalism, will be represented in the new formalism by the corresponding vector $|\Phi\rangle$. Since this is a one-to-one correspondence conserving the scalar product, the probability amplitudes calculated with the vectors of the new formalism are equal to the probability amplitudes calculated with those of the old – thus assuring the equivalence of the two formalisms. If a dynamical variable is represented in the ordinary formalism by a certain operator Q, it is represented in the new formalism by an operator $\hat{Q}$ of the space $\mathscr{E}_0 \otimes \mathscr{E}_C$; $\hat{Q}$ is symmetrical with respect to the $\mathscr{N}$ nucleons, it has the same eigenvalue spectrum as Q and its eigenstates can be obtained from those of Q by relation (XIV.78). $\hat{Q}$ is the symmetrical operator whose matrix elements satisfy the equation

$$\langle X|\hat{Q}|\Phi\rangle = \langle\chi|Q|\varphi\rangle. \qquad \text{(XIV.83)}$$

As an example we construct the operator $\hat{H}$, assuming two-body forces between the nucleons.

The Hamiltonian is

$$H = K + V \qquad \text{(XIV.84)}$$

where K is the total kinetic energy; if $k_p(\xi)$ and $k_n(\xi)$ represent the kinetic energy of a proton and neutron respectively, one has

$$K = K_p + K_n,$$

where

$$K_p = \sum_{i \leqslant Z} k_p^{(i)} \qquad K_n = \sum_{i<Z} k_n^{(i)}.$$

The total potential energy V has $\frac{1}{2}\mathscr{N}(\mathscr{N}-1)$ two-body interaction terms. There are $\frac{1}{2}Z(Z-1)$ proton-proton interaction terms

$$V_{pp} = \sum_{i<j\leqslant Z} v_{pp}^{(ij)},$$

$v_{pp}^{(ij)} \equiv v_{pp}(\xi^{(i)}, \xi^{(j)})$ denoting the interaction potential between protons

i and j; likewise there are ZN proton-neutron interaction terms

$$V_{\mathrm{pn}} = \sum_{i \leqslant Z} \sum_{j > Z} v_{\mathrm{pn}}^{(ij)},$$

and $\frac{1}{2}N(N-1)$ neutron-neutron interaction terms

$$V_{\mathrm{nn}} = \sum_{j > i > Z} v_{\mathrm{nn}}^{(ij)}.$$

The total potential energy is therefore

$$V = V_{\mathrm{pp}} + V_{\mathrm{pn}} + V_{\mathrm{nn}}.$$

The operator $\hat{K}$ corresponding to K is

$$\hat{K} = \sum_{i=1}^{\mathcal{N}} (k_{\mathrm{p}}^{(i)} \Pi_{\varpi}^{(i)} + k_{\mathrm{n}}^{(i)} \Pi_{\nu}^{(i)}). \tag{XIV.85}$$

It is indeed symmetrical and commutes notably with A. To show that it satisfies (XIV.83), we return to the definitions of the vectors $|\Phi\rangle$ and $|X\rangle$ and the projectors Π_{ϖ} and Π_{ν}:

$$\begin{aligned} \frac{N!\,Z!}{\mathcal{N}!} \langle X|\hat{K}|\Phi\rangle &= \langle \chi\zeta|A\hat{K}A|\varphi\zeta\rangle \\ &= \langle \chi\zeta|\hat{K}A|\varphi\zeta\rangle \\ &= \langle \chi\zeta|KA|\varphi\zeta\rangle \end{aligned}$$

from which

$$\langle X|\hat{K}|\Phi\rangle = \sqrt{\frac{\mathcal{N}!}{N!\,Z!}}\, \langle \chi\zeta|K|\Phi\rangle.$$

Since K does not operate on the charge variables, the partial scalar product $\langle\zeta|K|\Phi\rangle$ is equal to the result of K operating on $\langle\zeta|\Phi\rangle$; applying (XIV.79) we find

$$\langle X|\hat{K}|\Phi\rangle = \langle \chi|K|\varphi\rangle.$$

In the same way, the operator corresponding to V is

$$\hat{V} = \hat{V}_{\mathrm{pp}} + \hat{V}_{\mathrm{pn}} + \hat{V}_{\mathrm{nn}}, \tag{XIV.86}$$

where

$$\hat{V}_{\mathrm{pp}} = \sum_{i<j \leqslant \mathcal{N}} v_{\mathrm{pp}}^{(ij)} \Pi_{\varpi}^{(i)} \Pi_{\varpi}^{(j)} \tag{XIV.86a}$$

$$\hat{V}_{\mathrm{pn}} = \sum_{i<j \leqslant \mathcal{N}} v_{\mathrm{pn}}^{(ij)} [\Pi_{\varpi}^{(i)} \Pi_{\nu}^{(j)} + \Pi_{\nu}^{(i)} \Pi_{\varpi}^{(j)}] \tag{XIV.86b}$$

$$\hat{V}_{\mathrm{nn}} = \sum_{i<j \leqslant \mathcal{N}} v_{\mathrm{nn}}^{(ij)} \Pi_{\nu}^{(i)} \Pi_{\nu}^{(j)}. \tag{XIV.86c}$$

This can be verified by the same method as for $\hat{K}$.

The above theory for systems of neutrons and protons also applies to systems containing more than two types of fermions. Consider a system containing r types of similar but distinguishable fermions – n_1 fermions of type 1, n_2 fermions of type 2, ..., n_r fermions of type r. Rather than treat the different types of fermions as different particles, we may think of the r types as r separate states $|1\rangle, |2\rangle, \ldots, |r\rangle$ of the same fermion, and treat the system as a system of $n_1+n_2+\ldots+n_r$ fermions of this one type of which n_1 are in the state $|1\rangle$, n_2 in the state $|2\rangle$, ..., n_r in the state $|r\rangle$. The equivalence of this "isotopic" formalism to the ordinary one can be demonstrated using the same methods as we have used for nucleon systems. These remarks all apply to boson systems as well if we everywhere replace the operation of antisymmetrization by symmetrization (Problem XIV.9).

15. Utility of Isotopic Spin. Charge Independence

Consider a system of $\mathcal{N}$ nucleons. So long as we are dealing with definite numbers Z and N of protons and neutrons respectively, it is *a priori* simpler to treat the protons and the neutrons as different particles (the ordinary formalism), rather than as nucleons in different charge states (the isotopic-spin formalism).

On the other hand, when dealing with phenomena in which the charge of the nuclear system is not conserved, the isotopic-spin formalism is indispensable. The classic example is β-decay. In the decay of a β^--radioactive nucleus, the number of nucleons remains constant but one of the neutrons is transformed into a proton. In the theory of this phenomenon one treats the nucleus as a system of $\mathcal{N}$ nucleons in interaction with the quantized electron and neutrino fields. The theory of β decay is outside the scope of this book; we mention it only as an example.

This apart, the principal interest of the isotopic spin formalism is due to the fact that nuclear interactions are practically independent of nucleon charge.

The neutron and proton masses are equal to within less than 0.2 %, and therefore the kinetic energy of a nucleon is practically independent of its charge:

$$k_p = k_n = k. \qquad \text{(XIV.87)}$$

In the measure that the purely electromagnetic interactions may be neglected [1]) the nucleon-nucleon potential is also nearly independent of charge.

$$v_{pp} = v_{pn} = v_{nn} = v. \tag{XIV.88}$$

We shall examine the consequences of charge independence under the assumption that only two body forces act between the nucleons [2]).

In the ordinary formalism, the hypothesis of charge independence finds its expression in the following two properties of H:

(i) H does not depend on Z or N, but only on the total number of particles $\mathcal{N}$;

(ii) it is symmetrical with respect to permutations of these $\mathcal{N}$ particles, and not just with respect to permutations of the protons among themselves and/or the neutrons among themselves.

These properties are an immediate consequence of (XIV.87) and (XIV.88) for, if these hold, the Hamiltonian (XIV.84) takes the form

$$H = \sum_{i=1}^{\mathcal{N}} k^{(i)} + \sum_{i<j\leqslant\mathcal{N}} v^{(ij)}, \tag{XIV.89}$$

where $\sum_{i<j\leqslant\mathcal{N}}$ indicates, as in eqs. (XIV.86*a*, *b*, *c*), the summation over all of the $\frac{1}{2}\mathcal{N}(\mathcal{N}-1)$ pairs (ij) of the set of $\mathcal{N}$ particles.

In the isotopic-spin formalism the charge-independence hypothesis is reflected in the Hamiltonian $\hat{H}$ being independent of the charge variables and we have simply

$$\hat{H} = H, \tag{XIV.90}$$

H on the right-hand side being considered as an operator in the space $\mathscr{E}_0 \otimes \mathscr{E}_C$. This can be shown from the definition (XIV.83) of operators

---

[1]) Electromagnetic effects become predominant only for sufficiently large distances between the nucleons ($\gg 10^{-13}$ cm); at large distances v_{pn} and v_{nn} are practically zero, v_{pp} becomes simply the Coulomb repulsion of the two protons. For a discussion of the experimental evidence supporting charge independence, see Blatt and Weisskopf, *loc. cit.*, note 2, p. 576; D. H. Wilkinson, Phil. Mag. 1 (1956) 1031.

[2]) It is not necessary to make this restriction in order to show the equivalence of charge independence to the invariance property (XIV.91). It can be shown in the general case by application of the theorem of § 18, Appendix D.

in the isotopic spin formalism or by directly calculating expressions (XIV.85) and (XIV.86) for $\hat{K}$ and $\hat{V}$, assuming (XIV.87) and (XIV.88); the latter procedure gives for $\hat{V}$

$$\hat{V} = \sum_{i<j\leqslant\mathcal{N}} v^{(ij)} [\Pi_{\varpi}^{(i)} \Pi_{\varpi}^{(j)} + \Pi_{\varpi}^{(i)} \Pi_{\nu}^{(j)} + \Pi_{\nu}^{(i)} \Pi_{\varpi}^{(j)} + \Pi_{\nu}^{(i)} \Pi_{\nu}^{(j)}],$$

and since

$$\Pi_{\varpi}^{(i)} + \Pi_{\nu}^{(i)} = 1,$$
$$\hat{V} = \sum_{i<j\leqslant\mathcal{N}} v^{(ij)} = V.$$

Since $\hat{H}$ is independent of the charge variables, it commutes with all three components of the total isotopic spin:

$$[\hat{H}, \boldsymbol{T}] = 0. \qquad \text{(XIV.91)}$$

We shall now show that the converse is also true, i.e. if $\hat{H}$ satisfies (XIV.91), it can be written in the form of an operator independent of the charge variables.

$\hat{H}$ is a certain function of the ordinary variables ξ and the charge variables τ of each of the particles; in view of our hypothesis regarding the nature of the forces, this function will be of the form

$$\hat{H} = \sum_{i=1}^{\mathcal{N}} \hat{k}^{(i)} + \sum_{i<j\leqslant\mathcal{N}} \hat{v}^{(ij)},$$

where $\hat{k}^{(i)}$ depends only on the variables i, and $\hat{v}^{(ij)}$ only on the variables i and j. However, isotopic spin has the same mathematical properties as ordinary spin, and in particular the components of the vectors $\boldsymbol{\tau}^{(i)}$ have all the properties of the Pauli matrices; it results [cf. eq. (XIII.78–79)] that any function $\hat{H}$ of the components of $\boldsymbol{\tau}^{(1)}, \ldots, \boldsymbol{\tau}^{(\mathcal{N})}$ can always be expressed as a linear function of these operators. Also, since $\hat{H}$ commutes with $\boldsymbol{T}$, it is invariant with respect to rotations in $\mathscr{E}_C$ and is therefore a scalar function of the vectors $\boldsymbol{\tau}^{(1)}, \ldots, \boldsymbol{\tau}^{(\mathcal{N})}$. Thus $\hat{k}^{(i)}$ is a linear, scalar function of $\boldsymbol{\tau}^{(i)}$, hence independent of $\boldsymbol{\tau}^{(i)}$; $\hat{v}^{(ij)}$ is a linear scalar function of $\boldsymbol{\tau}^{(i)}$ and $\boldsymbol{\tau}^{(j)}$, hence necessarily of the form

$$\hat{v}^{(ij)} = a^{(ij)} + (\boldsymbol{\tau}^{(i)} \cdot \boldsymbol{\tau}^{(j)})\, b^{(ij)},$$

where $a^{(ij)}$ and $b^{(ij)}$ are functions of the orbital and spin variables alone. To complete the proof we need only to show that the action of $(\boldsymbol{\tau}^{(i)} \cdot \boldsymbol{\tau}^{(j)})$ on the *antisymmetrical* vectors of the space $\mathscr{E}_0 \otimes \mathscr{E}_C$

is the same as that of an operator acting on the variables of $\mathscr{E}_0$ alone. $(\boldsymbol{\tau}^{(i)} \cdot \boldsymbol{\tau}^{(j)})$ is related to the operator $P_C^{(ij)}$ representing the transposition (ij) in charge space by the Dirac identity

$$P_C^{(ij)} = \tfrac{1}{2}(1 + \boldsymbol{\tau}^{(i)} \cdot \boldsymbol{\tau}^{(j)}). \qquad \text{(XIV.92)}$$

To establish (XIV.92) we introduce the isotopic spin of the pair (ij)

$$\mathbf{t}_{ij} = \mathbf{t}^{(i)} + \mathbf{t}^{(j)}.$$

Any state in charge space is the sum of a state $|1\rangle$ for which $t_{ij} = 1$, and a state $|0\rangle$ for which $t_{ij} = 0$. Any triplet state is symmetrical in i and j; any singlet state antisymmetrical in i and j. Thus

$$\mathbf{t}_{ij}^2 |1\rangle = 2 |1\rangle \qquad \mathbf{t}_{ij}^2 |0\rangle = 0$$
$$P_C^{(ij)} |1\rangle = |1\rangle \qquad P_C^{(ij)} |0\rangle = -|0\rangle,$$

and consequently

$$P_C^{(ij)} = \mathbf{t}_{ij}^2 - 1.$$

From this we deduce (XIV.92) with the aid of the identity

$$\mathbf{t}_{ij}^2 = \tfrac{1}{2}(3 + \boldsymbol{\tau}^{(i)} \cdot \boldsymbol{\tau}^{(j)}).$$

It follows that

$$\boldsymbol{\tau}^{(i)} \cdot \boldsymbol{\tau}^{(j)} = 2P_C^{(ij)} - 1.$$

But in the space of the antisymmetrical vectors of $\mathscr{E}_0 \otimes \mathscr{E}_C$,

$$P^{(ij)} \equiv P_C^{(ij)} P_0^{(ij)} = -1.$$

Thus

$$P_C^{(ij)} = -P_0^{(ij)}$$

and the operator $\boldsymbol{\tau}^{(i)} \cdot \boldsymbol{\tau}^{(j)}$ may be replaced by

$$-2P_0^{(ij)} - 1,$$

which operates on the variables of ordinary space alone. Q.E.D.

The equivalence of charge-independence and rotational invariance in charge space is entirely general [1]). Since all the *mathematical* developments concerning rotations remain valid for the particular case of rotations in charge space (addition of isotopic spins, Wigner–Eckart theorem, selection rules, etc.), this equivalence provides a specially convenient method of taking into account the charge independence of nuclear forces.

---

[1]) Cf. footnote 2, p. 627.

EXERCISES AND PROBLEMS

1. The operator P associated with a permutation p of the N particles of a system of similar particles was defined in § 2 by its action on the basis vectors of a given symmetrical representation $\{Q\}$. Show that this definition does not depend on the representation chosen.

2. We denote by S_n, A_n the symmetrizing and antisymmetrizing operators for the particles 1, 2, ..., n and by S_{n-1}, A_{n-1} the symmetrizing and antisymmetrizing operators for the particles 1, 2, ..., $n-1$. Show that

$$S_n = \frac{1}{n}\left[1 + \sum_{i=1}^{n-1} P_{(in)}\right] S_{n-1} = \frac{1}{n} S_{n-1}\left[1 + \sum_{i=1}^{n-1} P_{(in)}\right]$$

$$A_n = \frac{1}{n}\left[1 - \sum_{i=1}^{n-1} P_{(in)}\right] A_{n-1} = \frac{1}{n} A_{n-1}\left[1 - \sum_{i=1}^{n-1} P_{(in)}\right].$$

3. Show that among the *symmetrical* observables of a system of N similar particles, there can be a complete set of commuting observables only if $N = 2$.

4. Find the proton-proton scattering cross section (in the center-of-mass system) when the incident proton is completely polarized in a given direction Ou (spin component along Ou equal to $+\frac{1}{2}$), and the target proton: (i) completely polarized in the same direction, (ii) completely polarized in the opposite direction (spin component along Ou equal to $-\frac{1}{2}$), (iii) non-polarized. Compare (XIV.56) with the expressions obtained in these three cases.

5. Expression (XIV.56) gives the scattering cross section for two fermions of spin $\frac{1}{2}$ having a spin-independent interaction. How must it be modified if the two identical particles are: (a) fermions of spin j; (b) bosons of spin j?

6. Calculate the average value $\langle r \rangle$ of the distance from the origin of each electron in the Thomas–Fermi model of the atom and compare with the numerical value of $\langle r \rangle$ for the hydrogen atom (use the integral $\int_0^\infty \chi(x)\mathrm{d}x \simeq 1.8$).

7. An atom of atomic number Z is p times ionized. Calculate the electronic density $\varrho(r)$ using the Thomas–Fermi model. Show that expression (XIV.66) for $\varrho(r)$ is still valid with the same definitions of b and x as in § 13 [eq. (XIV.64–65)] and that the function $\chi(x)$ figuring in that expression is the solution of equation (XIV.67) that vanishes at a certain point x_0 in the interval $(0, \infty)$ and satisfies the conditions

$$\chi(0) = 1 \qquad \chi'(x_0) = -p/Zx_0.$$

Examine the general behavior of the curve $\varrho(r)$ and of the electrostatic potential $\Phi(r)$.

8. We consider a system of Z electrons whose Hamiltonian is independent of the spins. Show that the energy spectrum for states having a definite value

M for the component S_z of total spin is the same (same positions and same degeneracy of levels) as that which we obtain if we treat the electrons of spin $\frac{1}{2}$ and those of spin $-\frac{1}{2}$ as different types of fermions and look for the states for which the $(\frac{1}{2}Z + M)$ first electrons have spin $\frac{1}{2}$ and the others spin $-\frac{1}{2}$.

9. Set up the "isotopic" formalism defined in § 14 for a system of $\mathcal{N}$ fermions of r similar but different types, namely n_1 fermions of type 1, n_2 fermions of type 2, ..., n_r fermions of type r $(n_1 + n_2 + \ldots + n_r = \mathcal{N})$. How must this formalism be modified if the $\mathcal{N}$ particles are bosons rather than fermions?

CHAPTER XV*)

INVARIANCE AND CONSERVATION THEOREMS. TIME REVERSAL

1. Introduction

This chapter is devoted to a systematic examination of the invariance properties that the equations of motion of a physical system may have with respect to certain transformations and the consequences of these in terms of the behavior of the system.

Certain mathematical complements are given in section I.

In section II we examine the general properties of transformations and of groups of transformations. With every transformation $\mathscr{T}$ of the variables and dynamical states there is associated a transformation operator T for the ket vectors representing the states. T is of necessity *either unitary linear or unitary antilinear*, and is defined up to an arbitrary phase factor by the law for the transformation of the fundamental observables of the system. In practice, T is always linear except for time reversal. With the various transformations of physics one can form a certain number of groups of transformations and with each such group $\mathscr{G}$ associate a certain group of transformation operators G [1]). After reviewing the most important of these groups we show with some simple examples how one constructs G in the case when $\mathscr{G}$ is a finite group, and in the case when $\mathscr{G}$ is a continuous group whose finite transformations can all be defined as a succession of infinitesimal transformations.

Questions relating specifically to invariance are taken up in section III. The transformations of this section are time-independent and linear, and the results obtained are simple generalizations of those previously obtained for rotations (Ch. XIII). To postulate that the equations of motion of the dynamical states are invariant under the transformations of a certain group $\mathscr{G}$, is equivalent to supposing that the Hamiltonian H commutes with the operators of the group G.

*) The fourth part (Ch. XVI–XIX), apart from a few specifically marked passages which may be omitted in a first reading, is independent of the present chapter and may be taken up first if desired.

[1]) The few notions regarding groups and their representations used in this chapter are given in sections I and II of Appendix D.

Thus any observable formed with the operators of G is a constant of the motion, and hence to any G-invariance there correspond a certain number of *conservation laws*. By taking into account these symmetries of H, we can simplify its diagonalization and make certain predictions regarding the degeneracies of its eigenvalues.

The invariance under time reversal differs from the rest by its physical significance and by the fact that the corresponding operator is antilinear. It is discussed in section IV.

Throughout this chapter, the striking analogy between Classical Mechanics and Quantum Mechanics will be apparent, especially in the definitions of the transformations, in the connection between the invariance properties of the equations of motion and the symmetries of the Hamiltonian, and in the existence of conservation laws.

I. MATHEMATICAL COMPLEMENTS. ANTILINEAR OPERATORS

2. Three Useful Theorems

THEOREM I. A necessary and sufficient condition that two linear operators A and B be equal, is that

$$\langle u|A|u\rangle = \langle u|B|u\rangle \text{ for any } |u\rangle.$$

THEOREM II. A necessary and sufficient condition that two linear operators A and B be equal to within a phase,

$$A = B\,e^{i\alpha} \tag{XV.1}$$

is that

$$|\langle u|A|v\rangle| = |\langle u|B|v\rangle| \text{ for any } |u\rangle \text{ and } |v\rangle. \tag{XV.2}$$

THEOREM III. If between the vectors of the space $\mathscr{E}$ there exists a one-to-one correspondence $\mathscr{T}$ which is defined up to an arbitrary constant and which conserves the scalar product, then the phases can always be chosen so as to make $\mathscr{T}$ either unitary linear or unitary antilinear.

Theorem I was given in Chapter VII (§ 5). It is included here simply for completeness.

With regard to theorem II, it is obvious that (XV.2) is a consequence of (XV.1). To show that the converse is also true, we choose a particular representation in which we denote the matrix

elements of A and B by A_{ij} and B_{ij} respectively. Since the basis vectors of the representation satisfy (XV.2), we have

$$|A_{ij}| = |B_{ij}| \text{ for all } i \text{ and } j. \tag{XV.3}$$

Letting $|u\rangle$ be the ith basis vector and $|v\rangle$ a linear combination of the jth and the kth basis vectors, we similarly find

$$|A_{ij}\, x_j + A_{ik}\, x_k| = |B_{ij}\, x_j + B_{ik}\, x_k|$$

for all values of the complex coefficients x_j and x_k. Taking into account (XV.3), this last equation may be written

$$\text{Re}\,[x_j\, x_k^*\, (A_{ij}\, A_{ik}^* - B_{ij}\, B_{ik}^*)] = 0.$$

In order that this be true for any $x_j\, x_k{}^*$, it is necessary that

$$A_{ij}\, A_{ik}^* - B_{ij}\, B_{ik}^* = 0.$$

This, together with (XV.3), gives

$$\frac{A_{ij}}{B_{ij}} = \frac{A_{ik}}{B_{ik}}. \tag{XV.4}$$

For a given i, the same argument may be repeated with different column indices j and k, thus showing that the ratio A_{ij}/B_{ij} is independent of j. Also, we can repeat the demonstration interchanging the role of the rows and the columns, and thereby show that this ratio is also independent of i. Since according to (XV.3) the moduli of the matrix elements of A and B are equal, this ratio must be a number of modulus unity:

$$\frac{A_{ij}}{B_{ij}} = e^{i\alpha} \qquad \text{for all } i \text{ and } j.$$

In other words, the operators A and B are equal to within a phase factor $e^{i\alpha}$. Q.E.D.

Consider next theorem III. By hypothesis, to each vector $|u\rangle$ of $\mathscr{E}$ there corresponds a vector $|u'\rangle$ given by the law $\mathscr{T}$. This vector is determined up to a phase factor. Let us make a particular choice for the phase of each of the vectors $|u'\rangle$. Then the law $\mathscr{T}$ sets up a one-to-one correspondence between the vectors of $\mathscr{E}$:

$$|u'\rangle = \mathscr{T}\,[|u\rangle] \qquad |u\rangle = \mathscr{T}^{-1}\,[|u'\rangle] \tag{I}$$

and this correspondence conserves the modulus of the scalar product

$$|\langle u'|v'\rangle| = |\langle u|v\rangle|. \tag{II}$$

Let

$$|1\rangle, |2\rangle, \ldots, |n\rangle, \ldots \tag{XV.5}$$

be a complete set of orthonormal vectors in $\mathscr{E}$. The corresponding set

$$|1'\rangle, |2'\rangle, \ldots, |n'\rangle, \ldots \tag{XV.5'}$$

is also a complete orthonormal set; it is orthonormal because, by (II), $\mathscr{T}$ conserves the norm and the orthogonality; complete because if there existed a vector $|a'\rangle$ orthogonal to all the vectors of (XV.5'), the vector $|a\rangle \equiv \mathscr{T}^{-1}[|a'\rangle]$ would be orthogonal to all the vectors of (XV.5), which is impossible by hypothesis.

Put

$$u_n \equiv \langle n|u\rangle \qquad u_n' \equiv \langle n'|u'\rangle \tag{XV.6}$$

We want to prove that, with a suitable choice of phases for the "primed" ket vectors, either one or the other of the following relations holds:

$$u_n' = u_n, \quad \text{whatever } |u\rangle \text{ and } n \tag{XV.7a}$$

$$u_n' = u_n^*, \quad \text{whatever } |u\rangle \text{ and } n. \tag{XV.7b}$$

Note that, from condition (II), we obviously have:

$$|u_n'| = |u_n|, \quad \text{whatever } |u\rangle \text{ and } n.$$

Thus, we only have to investigate the phase relationship between u_n' and u_n.

To this effect, we fix the phase of each basis ket $|n'\rangle$ by requiring that $|1'\rangle + |n'\rangle$ correspond to $|1\rangle + |n\rangle$.

We first prove relations (XV.7) in the case when $|u\rangle$ is a "real" ket, i.e. when all the u_n are real. Applying (II) to the scalar product of $|u\rangle$ by $|1\rangle + |n\rangle$ we obtain:

$$|u_1 + u_n| = |u_1' + u_n'|$$

If we choose the phase of $|u'\rangle$ in such a way that $u_1' = u_1$, we obtain the desired result: $u_n' = u_n$. The above argument does not apply when $u_1 = 0$. Its extension to include this case is straightforward, and will not be given here.

Consider now an arbitrary ket $|u\rangle$. Applying (II) to the scalar product of $|u\rangle$ by the "real" ket $|j\rangle + |j+1\rangle + \ldots + |j+k\rangle$ we obtain:

$$\left|\sum_{s=0}^{k} u_{j+s}\right| = \left|\sum_{s=0}^{k} u_{j+s}'\right| \tag{XV.8}$$

This holds for any j and s. In order to visualize this result, it is convenient to use the following geometrical representations for $|u\rangle$

and $|u'\rangle$. We associate with $|u\rangle$ the polygonal line (Γ) obtained by putting end to end the representative vectors in the complex plane of its successive components $u_1, u_2, \ldots, u_n, \ldots$. Similarly $|u'\rangle$ is represented by the polygonal line (Γ') built up from the representative vectors of $u_1', u_2', \ldots, u_n', \ldots$. (XV.8) means that the distance between any two virtices of (Γ) is equal to the distance between the corresponding virtices of (Γ'). As a consequence either (a) (Γ) can be brought into coincidence with (Γ') by a rotation, or (b) (Γ) can be brought into coincidence with (Γ') by a rotation and a reflection in the real axis.

In case (a), we choose the phase of $|u'\rangle$ by requiring that $u_1' = u_1$. With this choice, (Γ) and (Γ') coincide, i.e.

$$u_n' = u_n, \text{ whatever } n.$$

In case (b), our choice of phase is such that $u_1' = u_1^*$. Then, (Γ') is the mirror image of (Γ) with respect to the real axis, i.e.

$$u_n' = u_n^*, \text{ whatever } n.$$

Finally, we have to show that there are in fact only two possibilities: either all the ket vectors fall into case (a), or they all fall into case (b).

Assume, to be specific, that the particular vector $|j\rangle + e^{i\alpha}|k\rangle$ $(\alpha \neq n\pi)$ falls into case (a). Then, any vector $|u\rangle$ whose components u_j, u_k do not vanish and have a relative phase different from $n\pi$ necessarily also falls into case (a) as is readily seen by applying property (II) to the scalar product of $|u\rangle$ by $|j\rangle + e^{i\alpha}|k\rangle$. The argument is easily extended to include vectors whose components u_j, u_k either vanish or have a relative phase equal to $n\pi$. The same argument applies to case (b).

In conclusion, we have two possibilities:

Case (a). With a suitable choice of phases, relation (XV.7*a*) applies. The correspondence $\mathscr{T}$ is obviously *linear*. More precisely, we have

$$\langle u'|\, v'\rangle = \langle u\,|v\rangle \qquad \text{(XV.9}a\text{)}$$

showing that it is a unitary linear correspondence.

Case (b). With a suitable choice of phases, relation (XV.7*b*) applies. The correspondence $\mathscr{T}$ is obviously *antilinear*. More precisely, we have:

$$\langle u'\,|v'\rangle = \langle u\,|v\rangle^*. \qquad \text{(XV.9}b\text{)}$$

According to the definition of unitarity for antilinear operators, which is given below, it is a unitary antilinear correspondence. Q.E.D.

3. Antilinear Operators in Hilbert Space

The properties of antilinear operators in Hilbert space are analogous to those of linear operators. We shall briefly list them in the same order as we listed the properties of linear operators in Chapter VII.

DEFINITION. ACTION ON THE KETS

If to each ket $|u\rangle$ of Hilbert space there corresponds a certain ket $|v\rangle$, and if this correspondence is antilinear, then we say that $|v\rangle$ is the result of the action of a certain antilinear operator A on $|u\rangle$, and we write

$$|v\rangle = A\,|u\rangle. \tag{XV.10}$$

The antilinearity property is written

$$A\,(\lambda_1\,|1\rangle + \lambda_2\,|2\rangle) = \lambda_1^*\,(A\,|1\rangle) + \lambda_2^*\,(A\,|2\rangle). \tag{XV.11}$$

An antilinear operator is completely defined by its action on each of the vectors of a complete set of linearly independent vectors of $\mathscr{E}$, and notably by its action on the vectors of a basis in $\mathscr{E}$.

ALGEBRAIC OPERATIONS

Same definition as for linear operators:

(i) *Multiplication by a constant* c. If $c \neq c^*$, note that $c\,A \neq A\,c$;

$$\boxed{c\,A = A\,c^*} \tag{XV.12}$$

(ii) *Sum* of two antilinear operators (mentioned for completeness);

(iii) *Products*: if A_1, A_2 are antilinear operators, the product $A_1\,A_2$ defined by

$$(A_1\,A_2)\,|u\rangle = A_1\,(A_2\,|u\rangle)$$

is a linear operator. If A is antilinear and B linear, the product $A\,B$ is antilinear. More generally, if $A, B, \ldots, L$ contains in all $p+q$ operators of which p are linear and q antilinear, the product $(A\,B \ldots L)$ is linear or antilinear according as q is even or odd.

These products are all associative, but in general not commutative. The definition of commutators is the same as for linear operators and the corresponding algebraic rules (V.63–66) are still valid.

INVERSE

If the correspondence (XV.10) between $|u\rangle$ and $|v\rangle$ is of a one-to-one character, then it also defines the inverse A^{-1} of A.

$$|u\rangle = A^{-1}\,|v\rangle.$$

The antilinear operators A, B are by definition each the inverse of the other if we simultaneously have

$$A\,B = 1 \qquad B\,A = 1. \tag{XV.13}$$

If each of the operators $A, B, C, \ldots, L$ is either linear or antilinear, and if each of them has an inverse, then the inverse of their product exists and is given by

$$(A\,B\,C\,\ldots\,L)^{-1} = L^{-1}\,\ldots\,C^{-1}\,B^{-1}\,A^{-1}. \tag{XV.14}$$

ACTION ON THE BRAS

Let A be an antilinear operator and $\langle\chi|$ a bra vector; the complex conjugate of the scalar product $\langle\chi|(A\,|u\rangle)$, being a linear function of $|u\rangle$, defines a certain bra (cf. § VII.3), which we denote by $\langle\eta|$. By definition

$$\langle\eta| = \langle\chi|\,A. \tag{XV.15}$$

The correspondence between $\langle\chi|$ and $\langle\eta|$ is antilinear:

$$(\lambda_1\langle 1| + \lambda_2\langle 2|)\,A = \lambda_1^*\,(\langle 1|\,A) + \lambda_2^*\,(\langle 2|\,A). \tag{XV.16}$$

According to this definition, we have the following identity between scalar products

$$\boxed{(\langle\chi|\,A)\,|u\rangle = [\langle\chi|\,(A\,|u\rangle)]^*} \tag{XV.17}$$

This is to be compared with the identity (VII.17) for linear operators. In the present case we cannot omit the parentheses.

In order to define the three algebraic operations and the inverse for operators acting on bras, we proceed in the same way as for linear operators. The multiplication by a constant c gives [cf. eq. (XV.12)]:

$$\langle\chi|\,(c\,A) = c^*\,(\langle\chi|\,A) = \langle\chi|\,(A\,c^*). \tag{XV.18}$$

The rest is without change.

IMPORTANT NOTE

Relations (XV.12) and (XV.17) summarize the differences between ordinary manipulations and those involving antilinear operators:

(i) considered as an operator in ket-space or bra-space, a given constant c does not commute with antilinear operators unless it is real;

(ii) in the scalar product one must explicitly show whether A acts on the ket to its right or on the bra to its left.

In practice, parentheses are used as often as is necessary in order to avoid any possible confusion in the meaning of the symbols. Consider, for example, the product $A_1 A_2$ of two antilinear operators; the symbol $\langle u | A_1 A_2 | v \rangle$ is ambiguous, while on the other hand we can write without any possible confusion

$$\langle u | (A_1 A_2) | v \rangle = (\langle u | A_1 A_2) | v \rangle = [(\langle u | A_1) (A_2 | v \rangle)]^* = \langle u | (A_1 A_2 | v \rangle). \tag{XV.19}$$

Similarly, consider the product $(A | u \rangle \langle v |)$ of the linear operator $| u \rangle \langle v |$ by the antilinear operator A; the symbols $A | u \rangle \langle v | w \rangle$ and $\langle w | A | u \rangle \langle v |$ lead to confusion, while we can write without any possible confusion:

$$(A | u \rangle \langle v |) | w \rangle = A (| u \rangle \langle v | w \rangle) = (A | u \rangle) \langle v | w \rangle^* \tag{XV.20}$$

and

$$\langle w | (A | u \rangle \langle v |) = (\langle w | A) | u \rangle \langle v | = [\langle w | (A | u \rangle)]^* \langle v |. \tag{XV.21}$$

4. Antilinear Transformations

ADJOINT (OR HERMITEAN CONJUGATE) ANTILINEAR OPERATORS

$A^\dagger$ is by definition the adjoint of the antilinear operator A if $A^\dagger | u \rangle$ is the ket conjugate to $\langle u | A$ for any $| u \rangle$. It is an antilinear operator.

It follows that for any $| u \rangle$ and $| t \rangle$:

$$\langle t | (A^\dagger | u \rangle) \equiv \langle u | (A | t \rangle). \tag{XV.22}$$

This identity is to be compared with (VII.20). Apart from this, all of the properties set forth in § VII.7 can be taken over without change. In particular:

If each of the operators $A, B, C, \ldots, L$ is either linear or antilinear, then [cf. eq. (XV.14)]

$$(A\,B\,C\,\ldots\,L)^\dagger = L^\dagger \ldots C^\dagger\,B^\dagger\,A^\dagger. \tag{XV.23}$$

ANTIUNITARY OPERATOR

A is antiunitary if it is antilinear and if $A^{-1} = A^\dagger$:

$$A\,A^\dagger = A^\dagger\,A = 1.$$

If $A, B, C, \ldots, L$ contains in all $p+q$ operators of which p are unitary and q antiunitary, the product $A\,B\,C\,\ldots\,L$ is unitary or antiunitary according as q is even or odd.

ANTIUNITARY TRANSFORMATION OF LINEAR OPERATORS AND OF VECTORS

An antiunitary operator K defines an antiunitary transformation of the vectors and linear operators of $\mathscr{E}$, in which:

$$\text{any ket } |u\rangle \text{ transforms into } |\hat{u}\rangle \equiv K\,|u\rangle;$$
$$\text{any linear operator } B \text{ into } \hat{B} \equiv K\,B\,K^\dagger;$$
$$\text{any bra } \langle v| \text{ into } \langle \hat{v}| \equiv \langle v|K^\dagger.$$

In such a transformation:

(i) conjugate relations between bras and kets and Hermitean-conjugate relations between operators are conserved. If B is an observable, $\hat{B}$ is an observable with the same eigenvalue spectrum, the subspace of each eigenvalue of B transforming into the subspace of the same eigenvalue of $\hat{B}$;

(ii) scalar products transform into their complex conjugates

$$\langle \hat{u}\,|\hat{B}\,|\hat{v}\rangle = \langle u\,|B\,|v\rangle^*; \tag{XV.24}$$

(iii) considered as an operator, any constant c transforms into its complex conjugate

$$K\,c\,K^\dagger = c^*; \tag{XV.25}$$

(iv) any relation between vectors and/or operators is also satisfied by the transforms of these quantities if we replace all coefficients by their complex conjugates. In other words, the transformation K

conserves equations between vectors and/or operators if we agree to treat the coefficients involved as operators. For example the commutation relations

$$[q, p] = \mathrm{i}\hbar \qquad [J_x, J_y] = \mathrm{i}\hbar J_z \tag{XV.26a}$$

transform respectively into

$$[\hat{q}, \hat{p}] = -\mathrm{i}\hbar \qquad [\hat{J}_x, \hat{J}_y] = -\mathrm{i}\hbar \hat{J}_z. \tag{XV.26b}$$

5. Antilinear Operators and Representations

COMPLEX-CONJUGATION OPERATOR K_Q ASSOCIATED WITH A REPRESENTATION $\{Q\}$

By definition, K_Q is the operator that transforms the wave functions of the representation $\{Q\}$ into their complex conjugates. *It depends on the representation in question and notably on the phases of the basis vectors.*

Let $|1\rangle, |2\rangle, \ldots, |n\rangle, \ldots$ denote the basis vectors of $\{Q\}$. K_Q is the antilinear operator which leaves these vectors invariant

$$K_Q |n\rangle = |n\rangle. \tag{XV.27}$$

K_Q is thereby unambiguously defined. It is obvious that

$$K_Q{}^\dagger = K_Q \qquad K_Q{}^2 = 1 \tag{XV.28}$$

and therefore that K_Q is antiunitary. It clearly has the aforementioned property, namely that:

In the antiunitary transformation K_Q the matrices of the $\{Q\}$ representation are transformed into their complex conjugates, for we have

$$\langle n | (K_Q |u\rangle) = \langle n | u\rangle^* \qquad (\langle v | K_Q) | n\rangle = \langle v | n\rangle^*$$
$$\langle m | (K_Q B K_Q) | n\rangle = \langle m | B | n\rangle^* \quad (B, \text{ linear operator}).$$

Thus in the $\{Q\}$ representation the action of K_Q merely consists in taking the complex conjugate. The action of any other antilinear operator A may easily be deduced if we note that A is the product of K_Q and a linear operator, i.e. A can be written

$$A = (A K_Q) K_Q = K_Q (K_Q A) \tag{XV.29}$$

and $(A K_Q)$ and $(K_Q A)$ are linear operators, obtained from one

another by the transformation K_Q:

$$K_Q A = K_Q (A K_Q) K_Q$$

[if A is antiunitary, $(K_Q A)$ and $(A K_Q)$ are unitary].

CHANGE OF REPRESENTATION

Consider another representation $\{\Xi\}$. We denote by K_Ξ the complex-conjugation operator associated with this representation, and for the rest follow the notation of § VII.21; in particular, the matrix for the transformation of vectors and linear operators is the unitary matrix $S(\xi; n) \equiv \langle \xi | n \rangle$.

If this matrix is real, the vectors of the new basis are invariant under K_Q; in other words,

$$\text{if } S = S^*, \qquad K_\Xi = K_Q.$$

In this case, the linear operators $A K_\Xi$ and $K_\Xi A$ associated, in the $\{\Xi\}$ representation, with a given antilinear operator A, are the same as the ones that are associated with A in the $\{Q\}$ representation.

When this is not the case, AK_Ξ differs from $K_\Xi A$. It is useful to be able to deduce the matrix $(A K_\Xi)_\Xi$ representing $A K_\Xi$ in the $\{\Xi\}$ representation from the matrix $(A K_Q)_Q$ representing $A K_Q$ in the $\{Q\}$ representation. We have

$$\begin{aligned}\langle \xi | (A K_\Xi) | \xi' \rangle &= \langle \xi | (A K_Q)(K_Q K_\Xi) | \xi' \rangle \\ &= \sum_{mn} \langle \xi | m \rangle \langle m | (A K_Q) | n \rangle \langle n | (K_Q K_\Xi) | \xi' \rangle \\ &= \sum_{mn} \langle \xi | m \rangle \langle m | (A K_Q) | n \rangle \langle n | \xi' \rangle^*\end{aligned}$$

i.e.

$$(A K_\Xi)_\Xi = S (A K_Q)_Q \tilde{S}. \tag{XV.30}$$

Similarly

$$(K_\Xi A)_\Xi = (A K_\Xi)_\Xi^* = S^* (K_Q A)_Q S^\dagger. \tag{XV.31}$$

II. TRANSFORMATIONS AND GROUPS OF TRANSFORMATIONS

6. Transformations of the Dynamical Variables and Dynamical States of a System

We have already defined what is meant by "rotation of a physical system" and by "permutation of the particles of a physical system".

More generally, to effect a transformation $\mathscr{T}$ of a system, is to replace each of its variables by a new variable and each of its states by a new state, *while conserving the physical properties of the system.*

Thus the transformation $\mathscr{T}$ sets up a one-to-one correspondence between the dynamical variables: a given variable B transforms into another variable

$$B' \equiv \mathscr{T}[B]$$

of the system. By hypothesis, the transform B' has the same spectrum as B, and the eigenstates of each eigenvalue of B' are the transforms of the eigenstates of the same eigenvalue of B; these two conditions express the conservation of physical properties in the transformation of dynamical variables. Although we have been able to formulate this conservation condition without making reference to methods of measurement, the transformation $\mathscr{T}$ is particularly easy to visualize as a transformation applied to the measuring apparatus attached to the variable B: the transform of that apparatus is the apparatus attached to the variable B'. One so defines the displacements of the dynamical variables (rotations, translations), the reflections of the dynamical variables (reflection in a point, in a plane), etc.

From the concept of the transformation of variables we easily pass over to that of the transformation of states. Let $|u\rangle$ be a vector representing a possible dynamical state of the system. It can always be considered as the common eigenvector of a complete set of commuting observables, and is thereby defined up to a phase. Its transform

$$|u'\rangle \equiv \mathscr{T}[|u\rangle]$$

is the common eigenvector of the transformed observables. $\mathscr{T}$ therefore sets up a one-to-one correspondence between state vectors defined up to a phase.

By hypothesis the transformation conserves the physical properties

of dynamical states: the system being in the state $|u'\rangle$, the probability that a measurement will give a result corresponding to the state $|v'\rangle$, the transform of $|v\rangle$, is equal to the probability that the same measurement will give a result corresponding to the state $|v\rangle$ when the system is in the state $|u\rangle$. In other words

$$|\langle u'|v'\rangle|^2=|\langle u|v\rangle|^2 \text{ for all } |u\rangle \text{ and } |v\rangle.$$

The one-to-one correspondence in question therefore conserves the modulus of the scalar product. According to theorem III, the phases of the transformed vectors can always be fixed so as to make the transformation either unitary or antiunitary. We can then write

$$|u'\rangle=T|u\rangle, \tag{XV.32}$$

T being a unitary or antiunitary operator according to the transformation; in both cases we have

$$T T^\dagger=T^\dagger T=1. \tag{XV.33}$$

From the law (XV.32) for the transformation of vectors, we can easily deduce the law for the transformation of the density operator:

$$\varrho'=T\varrho T^\dagger. \tag{XV.34}$$

Consider once more the transformation of the observables. Since physical properties are conserved by our transformation, so are average values. Thus for an observable B we have

$$\langle u'|B'|u'\rangle=\langle u|B|u\rangle \text{ for any } |u\rangle$$

or again [1])

$$\langle u|(T^\dagger B' T)|u\rangle=\langle u|B|u\rangle \text{ for any } |u\rangle.$$

From theorem I and property (XV.33) we therefore have

$$B'=T B T^\dagger \qquad B=T^\dagger B' T. \tag{XV.35}$$

From relations (XV.35) we have the important property that any algebraic relation between the observables of the system is conserved in the transformation if T is linear, and is replaced by the complex conjugate relation if T is antilinear.

---

[1]) This result, obvious for T linear, is also true for T antilinear, due to $\langle u|B|u\rangle$ being real.

This places very restricting conditions on the transformation laws for observables. Any observable B is a certain *real* function $F(\xi)$ of the fundamental observables $\xi_1, \xi_2, \ldots, \xi_n, \ldots$, of the system. Its transform is necessarily $B' \equiv F(\xi')$. *The transformation $\mathscr{T}$ is therefore entirely determined once we know the transformation laws for the fundamental observables*, i.e. once we know the functions $f_1(\xi), \ldots, f_n(\xi), \ldots$, such that

$$\mathscr{T}[\xi_n] \equiv \xi_n' = f_n(\xi).$$

These will also determine the commutation relations for the ξ'. The latter must satisfy the property of conservation of algebraic relations given above. There are then only two possibilities: *either the transformation conserves the fundamental commutation relations, or it changes the sign* [1]). In the first case, the operator T associated with the transformation is linear; in the second, it is antilinear [cf. eq. (XV.26*a–b*)].

The operator T must satisfy the relations:

$$\xi_n' = T\,\xi_n\,T^\dagger. \tag{XV.36}$$

All physical properties of T are contained in these relations. They do not, however, suffice to completely determine T. Let T_1 be another unitary (or antiunitary) operator satisfying the same relations. We have (dropping the index n to simplify the writing)

$$T_1^\dagger\,\xi'\,T_1 = \xi$$

and therefore

$$T_1^\dagger\,T\,\xi\,T^\dagger\,T_1 = \xi$$

or

$$[T_1^\dagger\,T,\,\xi] = 0 \text{ for any } \xi. \tag{XV.37}$$

If we make the hypothesis that state space $\mathscr{E}$ is *irreducible* with respect to the observables ξ, i.e., that $\mathscr{E}$ contains no subspace invariant with respect to the ξ, then (XV.37) will be satisfied if, and only if, $T_1^\dagger\,T$ is a constant.

---

[1]) The transformations of Classical Mechanics are those which preserve the Poisson brackets

$$\{A_{\mathrm{cl}},\,B_{\mathrm{cl}}\}$$

of each pair $(A_{\mathrm{cl}},\,B_{\mathrm{cl}})$ of dynamical variables of the system (canonical transformations). Similarly the transformations that conserve physical properties in Quantum Mechanics are those which conserve the corresponding expressions:

$$1/i\hbar\,[A,\,B]$$

This result is contained in Schur's Lemma (§ D.8). It can be directly proven as follows. Let $|u\rangle$ be a simultaneous eigenvector of a *complete* set of commuting observables. Since $C \equiv T_1^\dagger T$ commutes with each of these, $|u\rangle$ is necessarily an eigenvector of C: $C|u\rangle = c|u\rangle$. Since C commutes with any function $F(\xi)$ of the observables of the system, we also have

$$C\ F(\xi)\ |u\rangle = c\ F(\xi)\ |u\rangle \text{ for any } F(\xi).$$

But the space spanned by the vectors $F(\xi)\ |u\rangle$ is an invariant subspace of $\mathscr{E}$ with respect to the ξ, and since by hypothesis $\mathscr{E}$ is irreducible, this can only be $\mathscr{E}$ itself. Therefore $C = c$.

If we suppose, as we have always done up to the present, that every vector of state space may be considered as an eigenvector of a complete set of commuting observables, then this property of irreducibility is automatically assured (Problem XV.1). Clearly the above discussion is meaningful only if the observables mentioned are physical observables. We shall always suppose $\mathscr{E}$ to be irreducible with respect to physical observables [1]). If such an assumption is made, the constant c is necessarily of modulus 1 since T and T_1 are unitary. Thus we can write

$$T_1 = e^{i\alpha}\, T.$$

In summary, *with each transformation $\mathscr{T}$ there is associated a unitary or antiunitary operator T, defined up to a phase by the transformation laws for the fundamental variables of the system* [eq. (XV.36)]. *T is unitary if the transformation conserves the commutation relations, antiunitary if it changes their sign.*

The phase of T may be arbitrarily chosen and has no effect on the physical properties of the transformation, not affecting the transformation law for observables, the transformation law for density operators, nor the different algebraic manipulations between operators.

7. Groups of Transformations

With the various transformations at our disposal, we can form a certain number of *groups* of transformations, the word group being taken in its mathematical sense (cf. § D.2).

The *product* $\mathscr{T}_{21} \equiv \mathscr{T}_2 \mathscr{T}_1$ of transformations $\mathscr{T}_1$ and $\mathscr{T}_2$ is the

[1]) Cf. reference in note 2, p. 536, in which the irreducibility hypothesis is abandoned.

transformation realized by applying first $\mathscr{T}_1$ and then $\mathscr{T}_2$. The transformation operator T_{21} is to within a phase factor equal to the product $T_2 T_1$. This type of product is therefore *associative*; it is not necessarily commutative.

The *identity* $\mathscr{I}$ is the transformation in which each observable transforms into itself; the corresponding operator is the multiplication by an arbitrary phase factor.

The *inverse* $\mathscr{T}^{-1}$ of a transformation $\mathscr{T}$ is defined by the relations $\mathscr{T}^{-1}\mathscr{T} = \mathscr{T}\mathscr{T}^{-1} = \mathscr{I}$; since $\mathscr{T}$ defines a one-to-one correspondence, the inverse will always exist.

Thus, the transformations defined in § 6 may each be regarded as an element of a certain group $\mathscr{G}$.

Of all the groups, the one whose physical significance is the most immediately apparent is *the group of spatial transformations* (translations, rotations, reflections) *and its various subgroups*. Of the latter we may cite the translation group, the group of rotations (about a point), the group of displacements (translation × rotation), the group of reflections in a point, in a plane, the rotation-reflection group (rotation about a point × reflection in the same point), the groups of symmetries in a crystal.

In the preceding chapter we encountered groups of a different category, the *groups of permutations* of similar particles. We also considered permutations of certain variables to the exclusion of others, and notably the group of permutations in charge space, with which is associated, in the case of nucleon systems, the *group of isotopic rotations* or group of rotations in charge space.

It is convenient to treat separately the transformations that explicitly bring in the time. In the first place there are the Galilean coordinate transformations, mentioned here for completeness (Problem XV.7). These are the non-relativistic equivalents of the pure Lorentz transformations which form with the spatial rotations, the proper Lorentz group, to be discussed in Part 5. There is also the group of time translations, and finally time reversal. Translation and reversal of time are studied in section IV. In the rest of the present section we deal only with transformations not explicitly involving the time.

8. Groups of Transformation Operators

Let $[\mathscr{T}]$ be a set of transformations $\mathscr{T}_1, \mathscr{T}_2, \ldots, \mathscr{T}_i, \ldots$. With each element $\mathscr{T}_i$ of the set we can associate an operator T_i defining the

transformation of the vectors and operators of state space. T is defined by its physical properties only to within a phase factor, which we choose arbitrarily for the moment. We thereby obtain a set of transformation operators $[T]$ whose elements are in one-to-one correspondence with those of $[\mathscr{T}]$.

Suppose now that $[\mathscr{T}]$ is a certain group $\mathscr{G}$. It does not necessarily follow that $[T]$ has the group property. In the correspondence between $[\mathscr{T}]$ and $[T]$, products are conserved only to within a phase factor; for each product

$$\mathscr{T}_k = \mathscr{T}_j \mathscr{T}_i$$

we have

$$T_k = e^{i\alpha_{ji}^k} T_j T_i,$$

where α_{ji}^k is a certain phase depending on the choice of phase for T_i, T_j and T_k. In order that $[T]$ be a group all of the α_{ji}^k must vanish, and in this case $[T]$ is isomorphic to the group $\mathscr{G}$.

When the phases of the T_i can be chosen in such a way that $[T]$ has the group property, then this is obviously the most convenient choice. This possibility exists for some groups but not for all. It exists for the permutation group (Ch. XIV) but not for the rotation group (§ XIII.15) when the system has an odd number of particles with half-integral spins. In this last case the rotation operators $R(\alpha\,\beta\,\gamma)$ defined by equation (XIII.60) do form a group, but there are two R operators, differing by a sign, for each rotation $\mathscr{R}$; if we take one of them as an element of $[R]$ we thereby establish a one-to-one correspondence between the rotations and the rotation operators, but the product is only conserved up to a sign, and the set $[R]$ does not have the group property.

To obtain a set of transformation operators with the group property, it may be necessary to associate with each $\mathscr{T}_i$ not a single operator T_i but a set (T_i) of operators differing one from another by a phase. If the (T_i) are suitably chosen, the set of transformation operators $\{T\}$ thereby obtained forms a group G and this group is homomorphic to the group $\mathscr{G}$. Let (1) be the set of operators associated with the transformation $\mathscr{I}$; the elements of (1) are the operator 1 and perhaps other elements obtained by multiplying 1 by a phase factor. It is an invariant subgroup of G and the quotient group $G/(1)$ is isomorphic to $\mathscr{G}$ (cf. § D.5).

We can thus form many sets $\{T\}$ having the group property [1]). In practice, we choose one of them once and for all; it is obviously chosen to be as simple as possible. We thereby obtain a group G of transformation operators homomorphic to the group $\mathcal{G}$.

In all the cases encountered up to the present in Quantum Theory, one can always choose G so as to have corresponding to each element of $\mathcal{G}$ *either a single operator of* G (isomorphism) *or, failing this, two operators of* G *of which one is the negative of the other*. The first of these possibilities can always be realized when the system has an even number of half-integral spins. We have already had an example of the second in the rotation of half-integral spins; we shall have another in connection with time reversal; it always appears when the system has an odd number of half-integral spins.

9. Continuous Groups and Infinitesimal Transformations. Translations. Rotations

As an illustration of the theory, we shall construct the group G for some of the groups $\mathcal{G}$. We consider first the continuous groups – those with an infinite number of elements depending on one or several continuous parameters – and more particularly those for which the finite transformations can all be generated by a series of infinitesimal transformations. Such is the case for the rotation group, and for the group of spatial translations. We need then only to know the infinitesimal transformations of the observables to have their transformation under any operation of the group. With each of the infinitesimal transformations of the group can be associated an infinitesimal transformation operator, i.e. a unitary operator differing by an infinitesimal from 1 [2]).

Suppose for simplicity that the group elements depend only on a

[1]) One of them, the most complicated of all, is obtained by associating with each $\mathcal{T}_i$ all of the operators obeying (XV.36): in this case, the elements of (T_i) are all obtained by multiplying one of them by an arbitrary phase factor. In practice, it is always possible to impose a reality condition on the T_i that fixes the phase factor up to a sign without violating the group property. The set (T_i) then has just two elements, differing from each other by a sign (cf. note, p. 669).

[2]) It could not be antiunitary. Consider two non-commuting observables: in an infinitesimal transformation they are subject to infinitesimal modifications; their commutator therefore cannot undergo a finite modification and in particular cannot change sign.

continuous parameter α and that the latter be chosen so as to have

$$\mathscr{T}(\alpha) \underset{\alpha \to 0}{\longrightarrow} \mathscr{I}.$$

To the first order in $\delta\alpha$, the infinitesimal transformation operator associated with $\mathscr{T}(\delta\alpha)$ is given by

$$T(\delta\alpha) = 1 - \mathrm{i}\,\Theta\,\delta\alpha$$

where Θ is a Hermitean operator (since T is unitary).

If the observable ξ is transformed into $\xi + \delta\xi$ in the transformation $\mathscr{T}(\delta\alpha)$, we have [eq. (VII.96)]

$$\delta\xi = -\mathrm{i}\,\delta\alpha[\Theta, \xi],$$

i.e.

$$[\Theta, \xi] = \mathrm{i}\frac{\delta\xi}{\delta\alpha}. \tag{XV.38}$$

For a given transformation $\mathscr{T}(\delta\alpha)$, the $\delta\xi/\delta\alpha$ are known and relations (XV.38) define Θ up to a constant [1]).

Consider, for example, the translations of a particle along the axis Ox. Let $\boldsymbol{r} \equiv (x, y, z)$ be the position, $\boldsymbol{p} \equiv (p_x, p_y, p_z)$ the momentum and $\boldsymbol{s} \equiv (s_x, s_y, s_z)$ the spin of the particle. In the translation $\mathscr{T}_x(a)$ of distance a along the axis Ox, the nine fundamental variables are invariant with the exception of x which goes over into $x - a$ [2]):

$$T(a)\, x\, T^{\dagger}(a) = x - a. \tag{XV.39}$$

---

[1]) In Classical Mechanics, the variation $\delta\alpha\,\partial\xi/\partial\alpha$ of each variable ξ in the infinitesimal displacement $\mathscr{T}(\delta\alpha)$ is defined by the Poisson bracket

$$\frac{\partial\xi}{\partial\alpha} = \{\tau, \xi\},$$

where τ is the conjugate momentum associated with the variable α. The observable $\hbar\Theta$ is the quantum analogue of τ. [In an infinitesimal rotation about $\boldsymbol{u}$, τ is the component of the angular momentum along $\boldsymbol{u}$; in a translation along $\boldsymbol{u}$, τ is the component of the momentum along $\boldsymbol{u}$.]

[2]) The transform x' of x is indeed $x - a$ and not $x + a$. Let $|b\rangle$ be an eigenvector of x corresponding to the eigenvalue b. Its transform $|b'\rangle$ is an eigenvector of x corresponding to the eigenvalue $b + a$:

$$x\,|b\rangle = b\,|b\rangle \qquad x\,|b'\rangle = (b + a)\,|b'\rangle.$$

But, from the definition of the transformation of observables

$$x'\,|b'\rangle = b\,|b'\rangle$$

whence

$$x' = x - a.$$

This argument is to be compared with that of § XIII.12.

In particular, in the infinitesimal transformation $\mathscr{T}_x(\delta a)$, $\delta x = -\delta a$ and all of the other variations $\delta y, \ldots, \delta s_z$ vanish. The corresponding Hermitean operator Θ_x therefore obeys the commutation rules

$$[\Theta_x, x] = -\mathrm{i}$$
$$[\Theta_x, y] = \ldots = [\Theta_x, s_z] = 0$$

which gives

$$\Theta_x = \frac{p_x}{\hbar} + k_0,$$

where k_0 is a real arbitrary constant which we put equal to zero. This treatment is easily extended to translations of a set of N particles and gives

$$\boxed{\Theta_x = \frac{P_x}{\hbar}} \qquad \text{(XV.40)}$$

where $P_x = \sum_{i=1}^{N} p_x^{(i)}$ *is the component along* Ox *of the total momentum* of the N particles.

With the infinitesimal translation $\mathscr{T}_x(\delta a)$ there therefore is associated the infinitesimal unitary operator

$$T_x(\delta a) = 1 - \frac{\mathrm{i}}{\hbar} P_x \delta a.$$

The finite-transformation operator $T_x(a)$ can be chosen to be

$$T_x(a) = \exp(-\mathrm{i} P_x a/\hbar)$$

as may easily be verified by substituting into relation (XV.39) which defines it. The operators thus defined form a group isomorphic to the group of translations along Ox. We have in particular

$$T_x(a)\ T_x(b) = T_x(b)\ T_x(a) = T_x(a+b).$$

The group of translations along Ox is a subgroup of the translation group proper. A particular translation $\mathscr{T}(\boldsymbol{a})$ is defined by the vector $\boldsymbol{a}$ giving the displacement of the dynamical states of the system. The translation group therefore depends on three continuous parameters – the components of $\boldsymbol{a}$. Its composition law is

$$\mathscr{T}(\boldsymbol{a})\ \mathscr{T}(\boldsymbol{b}) = \mathscr{T}(\boldsymbol{b})\ \mathscr{T}(\boldsymbol{a}) = \mathscr{T}(\boldsymbol{a}+\boldsymbol{b}).$$

Generalizing the preceding results to any translation whatever, we associate with the infinitesimal translation $\mathscr{T}(\boldsymbol{\epsilon})$ the operator

$$\boxed{T(\boldsymbol{\epsilon}) \simeq 1 - \frac{\mathrm{i}}{\hbar}(\boldsymbol{P}\cdot\boldsymbol{\epsilon})} \qquad \text{(XV.41)}$$

where $\boldsymbol{P}$ is the total momentum of the N particles of the system. It follows that the operator associated with the translation $\mathscr{T}(\boldsymbol{a})$ is

$$\boxed{T(\boldsymbol{a}) = \mathrm{e}^{-\mathrm{i}\boldsymbol{P}\cdot\boldsymbol{a}/\hbar}} \qquad \text{(XV.42)}$$

The T operators thus defined form a group isomorphic to the translation group since

$$T(\boldsymbol{a})\ T(\boldsymbol{b}) = T(\boldsymbol{b})\ T(\boldsymbol{a}) = T(\boldsymbol{a}+\boldsymbol{b}).$$

The rotation group provides another example of a continuous group with three parameters. The rotation operators have already been defined in Chapter XIII. The defining formula (XIII.55) may be compared with (XV.41) above. The total angular momentum $\boldsymbol{J}$ plays the same role for the rotation group as the total momentum $\boldsymbol{P}$ for the translation group. The component $(\boldsymbol{u}\cdot\boldsymbol{J})$ of the total angular momentum along $\boldsymbol{u}$ is defined up to a constant by the commutation relations (XIII.56) and (XIII.57), which give the infinitesimal rotations about $\boldsymbol{u}$ of scalar and vector observables respectively; the arbitrary constant is fixed by the condition that $\boldsymbol{J}$ itself be a vector operator [1]).

The rotation operators R are defined as products of infinitesimal rotations. This leads to formula (XIII.60) which may be compared with (XV.42) above. The operators in question form a group. This group is isomorphic to the rotation group if the system contains an even number of half-integral spins, but only homomorphic to the latter if the system has an odd number of half-integral spins. This point has already been discussed and will not be taken up again here.

The set of T and R operators defined by (XV.42) and (XIII.60), taken together with all of their products, also forms a group. If the system contains an even number of half-integral spins this group is isomorphic to the displacement group; if not, it is merely homomorphic

[1]) This condition is equivalent to the condition that the $R_{\boldsymbol{u}}(\varepsilon)$ be the infinitesimal operators of a group.

to the displacement group, two operators differing by a sign corresponding to each element of the latter.

10. Finite groups. Reflections

Of all the groups, the group of reflections in a point is without doubt the simplest. It has in all two elements, the identity $\mathscr{I}$ and the reflection $\mathscr{S}_0$: $\mathscr{S}_0^2 = \mathscr{I}$. In the transformation $\mathscr{S}_0$ the polar vectors $\boldsymbol{r}$, $\boldsymbol{p}$ change sign and the axial vectors $\boldsymbol{r} \times \boldsymbol{p}$, $\boldsymbol{s}$ are invariant. Since $\boldsymbol{r}$ and $\boldsymbol{p}$ simultaneously change sign, the transformation conserves the commutation relations of the orbital variables; it also conserves those of the components of spin. The operator S_0 defining the reflection of a particle is therefore linear; it is the unitary operator obeying the relations

$$\begin{aligned} S_0\, \boldsymbol{r}\, S_0^\dagger &= -\boldsymbol{r} \\ S_0\, \boldsymbol{p}\, S_0^\dagger &= -\boldsymbol{p} \\ S_0\, \boldsymbol{s}\, S_0^\dagger &= \boldsymbol{s}. \end{aligned} \tag{XV.43}$$

It is defined by these up to a phase. In order that in addition the operators S_0 and 1 form a group isomorphic to the reflection group, we must have

$$S_0^2 = 1 \qquad (S_0 = S_0^\dagger) \tag{XV.44}$$

which fixes the phase of S_0 to within a sign.

S_0 is completely determined by specifying its action on the basis vectors of a representation, the $\{\boldsymbol{r}, s_z\}$ representation for example. We shall adopt the following definition

$$S_0\, |\boldsymbol{r}\, \mu\rangle = |(-\boldsymbol{r})\, \mu\rangle. \tag{XV.45}$$

which is consistent with (XV.43) and (XV.44) (Problem XV.2). According to this definition, the reflection of a wave function is given by the rule

$$S_0\, \psi(\boldsymbol{r}, \mu) = \psi(-\boldsymbol{r}, \mu).$$

The operator S_0 is none other than the *parity* operator introduced in § XIII.23. It is an observable with the two eigenvalues ± 1. These considerations can be extended without difficulty to systems of several particles.

The reflection S_0 commutes with all of the rotations $\mathscr{R}$. The product of the operations of the reflection group with those of the rotation

group forms the rotation-reflection group. Note too that S_0 commutes with any rotation operator since the latter are functions of the total angular momentum $\boldsymbol{J}$, and S_0 commutes with $\boldsymbol{J}$, since by (XV.43)

$$S_0 \boldsymbol{J} S_0^\dagger = \boldsymbol{J}.$$

Thus the set formed of S_0, the R operators, and their products also forms a group. When the group $[R]$ is isomorphic to the rotation group (integral total spin) this group is isomorphic to the rotation-reflection group; if not (half-integral total spin), it is only homomorphic to the rotation-reflection group; with each element of the rotation-reflection group we have associated two operators, differing by a sign; in particular, the two operators associated with the pure reflection are $+S_0$ and $-S_0$ and the two operators associated with the identity are $+1$ and -1.

We next consider a different type of reflection, reflection in a plane. Let $\mathscr{S}_u$ be the reflection in the plane perpendicular to the unit vector $\boldsymbol{u}$. $\mathscr{S}_u$ is a particular transformation of the rotation-reflection group, namely the product of $\mathscr{S}_0$ with a rotation through an angle π about $\boldsymbol{u}$ (or $-\boldsymbol{u}$):

$$\mathscr{S}_u = \mathscr{S}_0 \mathscr{R}_u(\pi). \tag{XV.46}$$

We note that

$$\mathscr{S}_u^2 = \mathscr{I}. \tag{XV.47}$$

Consequently $\mathscr{S}_u$ and $\mathscr{I}$ form a group, and the study of this group can be patterned after the study of reflections in a point. In addition, with the aid of (XV.43) we can deduce the properties of one of these groups from those of the other.

We shall consider only the case of the single particle. The operator S_u is the unitary linear operator obeying the relations

$$\begin{aligned} S_u \boldsymbol{r} S_u^\dagger &= \boldsymbol{r} - 2\boldsymbol{u}(\boldsymbol{u}\cdot\boldsymbol{r}) \\ S_u \boldsymbol{p} S_u^\dagger &= \boldsymbol{p} - 2\boldsymbol{u}(\boldsymbol{u}\cdot\boldsymbol{p}) \\ S_u \boldsymbol{s} S_u^\dagger &= -\boldsymbol{s} + 2\boldsymbol{u}(\boldsymbol{u}\cdot\boldsymbol{s}). \end{aligned} \tag{XV.48}$$

We can take for S_u the expression

$$S_u = S_0 R_u(\pi) = S_0 \, e^{-i\pi(\boldsymbol{J}\cdot\boldsymbol{u})} \tag{XV.49}$$

which gives

$$S_u^2 = S_0^2 \, e^{-2\pi i(\boldsymbol{J}\cdot\boldsymbol{u})} = (-)^{2J}. \tag{XV.50}$$

With this choice of phase we therefore have

$$S_u^2 = -1$$

in the case of a half-integral spin. To have a set of transformation operators that form a group, we must associate with $\mathscr{S}_u$ the two operators [1]) S_u and $-S_u$.

Another type of finite group is the group of permutations of n similar particles. This group was studied in Chapter XIV. With each permutation we associated a linear, unitary permutation operator; the set of operators thereby obtained forms a group isomorphic to the permutation group [2]). We shall not return to these questions here. We add only the important remark that the permutations commute with the spatial transformations and that, just from the way in which they have been constructed, the permutation operators and the spatial transformation operators have the same property.

III. INVARIANCE OF THE EQUATIONS OF MOTION AND CONSERVATION LAWS

11. Invariant Observables

We shall now consider the question of invariance itself. Let $\mathscr{G}$ be a certain group of transformations. We shall denote by G the group of corresponding transformation operators and by T_t a particular element of G. We suppose that all of the T_t are linear (and unitary). The only transformation leading to the introduction of antilinear operators is time reversal; it will be studied in section IV.

The fact that an observable Q is invariant with respect to the transformations of the group is expressed by the condition $T_t\, Q T_t^\dagger = Q$, i.e.,

$$[Q, T_t] = 0 \text{ for any } T_t. \qquad \text{(XV.51)}$$

The consequences of these commutation rules have already been

[1]) If we consider only the group $\{S_u, \mathscr{S}\}$, it is preferable to put $S_u = \mathrm{i}^{2J} S_0 R_u(\pi)$; this gives $S_u^2 = 1$. With this new choice of phase, the multiplication rule (XV.46) is conserved only up to a phase.

[2]) The same result could have been obtained by replacing all the operators associated with odd permutations by their negatives. The work of Chapter XIV may all be carried through with this new phase convention, with the sole difference that the formulas are complicated with $(-)$ signs without any of the results being changed.

analysed for the case of the rotation group (§ XIII.17). They can be formulated in a very general way using the group concept and the properties of the linear representations of the group G (cf. Appendix D and in particular § D.9). We shall denote by $|\tau\, j\, \mu\rangle$ the basis vectors of a standard representation suited to the group G. These vectors are labelled with three quantum numbers (or 3 sets of quantum numbers). j indicates the irreducible representation to which $|\tau\, j\, \mu\rangle$ belongs. μ distinguishes between the basis vectors of a given irreducible representation. τ is an additional quantum number permitting, if need be, to distinguish between orthogonal, equivalent, irreducible subspaces. We shall intentionally use the same symbols as in § XIII.6, of which the present discussion constitutes an obvious generalization. The essential property of Q is the generalization of (XIII.120), namely

$$\langle \tau\, j\, \mu\, |Q\, |\tau'\, j'\, \mu'\rangle = \delta_{jj'}\, \delta_{\mu\mu'}\, Q^{(j)}_{\tau\tau'}. \qquad \text{(XV.52)}$$

A complete proof is given in Appendix D [cf. eq. (D.20)].

In many cases, it is possible to derive this relation without referring explicitly to the general theory of group representation. To this effect, one merely needs to find among the functions of T_t:

(i) a set J of observables invariant under *all* operations of the group and whose eigenvalues can be labelled with the quantum number (or set of quantum numbers) j;

(ii) a set M of observables that commute with each other but not with all of the operations of the group, and whose eigenvalues are labelled by μ.

Q, J, and M clearly form a set of commuting observables and therefore Q has a specially simple representative matrix in any representation in which J and M are diagonal, namely the one given by relation (XV.52).

This method was successfully applied to the rotation group [cf. eq. (XIII.120)]. In this particular case, we found one observable of category (i) namely $\boldsymbol{J}^2$, and one observable of category (ii), namely J_z. It also applies to the rotation-reflection group with $\boldsymbol{J}^2$ and S_0 — that is, the total angular momentum and the parity — for set (i), and the observable J_z for set (ii). Note that there is a certain arbitrary in the choice of the M; in the case of rotations, the usual practice is to take J_z, but J_x or J_y or any other component of $\boldsymbol{J}$ would do as well.

The eigenstates of Q are obtained by separately diagonalizing the matrices $Q^{(j)}_{\tau\tau'}$, which each correspond to a well-defined eigenvalue of J. Thus to each j there corresponds a set $q_s^{(j)}$ of eigenvalues of Q which we distinguish one from another, if need be, by the quantum number (or numbers) s. Let d_j be the number of possible values for μ ($2j+1$ in the case of rotations); to each of these values there corresponds the same matrix $Q^{(j)}_{\tau\tau'}$; thus each non-degenerate eigenvalue of this matrix is a d_j-fold degenerate eigenvalue of Q, each p-fold degenerate eigenvalue of this matrix is a pd_j-fold degenerate eigenvalue of Q. If $d_j \neq 1$, all the eigenvalues of Q corresponding to the quantum number j are degenerate, and the order of their degeneracy is a multiple of d_j. This type of degeneracy is a direct consequence of the invariance of Q with respect to G and will be called a *G-degeneracy*.

12. Symmetry of the Hamiltonian and Conservation Laws

Suppose that the Hamiltonian H is invariant in the transformations of a group $\mathcal{G}$. All that we have said about invariant observables will then apply to it. Starting from the commutation rules

$$[H, T_i] = 0 \text{ for any operator } T_i \text{ of } G, \qquad \text{(XV.53)}$$

we form observables of type J and of type M. These observables form with H a set of observables that can be simultaneously diagonalized. In addition, the spectrum of H has a *G-degeneracy*.

As in Classical Mechanics, the symmetries of the Hamiltonian lead to conservation laws, i.e. since any observable (not explicitly dependent on time) that commutes with H is a constant of the motion, we have the obvious property:

If H is invariant in the transformations of a group, any observable that is a function of the operators of the group is a constant of the motion.

In particular, the observables of the set (J, M) defined above have this property, and since they commute, they can be given precise values simultaneously and keep this set of precise values in the course of time.

Thus with each group is associated a certain number of conservation laws. Whenever the conserved observable has a classical analogue,

these laws are identical with the corresponding classical conservation laws.[1]) The most common of these are listed below.

(i) *Translational invariance and conservation of the total momentum.* A necessary and sufficient condition for the Hamiltonian to be invariant under translation is that it be invariant with respect to infinitesimal translations, i.e. that [eq. (XV.41)]

$$[H, \boldsymbol{P}] = 0, \tag{XV.54}$$

where $\boldsymbol{P}$ is the total momentum of the system. Thus the three components of total momentum are constants of the motion, and the total momentum is conserved. In addition, since P_x, P_y and P_z commute, they can simultaneously be precisely defined, so that they conserve their respective values in the course of time.

(ii) *Rotational invariance and conservation of the angular momentum* (already studied in Chapter XIII; rotational invariance is expressed by the condition: $[H, \boldsymbol{J}] = 0$).

(iii) *Reflection invariance and conservation of the parity.* It results from the invariance of the Hamiltonian under reflection in a point,

$$[H, S_0] = 0, \tag{XV.55}$$

that the *parity* S_0 is a constant of the motion.

(iv) *Invariance under permutation and conservation of symmetry.* In a system containing identical particles, the Hamiltonian H is invariant in any permutation P of these particles:

$$[H, P] = 0 \qquad \text{for any } P.$$

It follows that any observable built up from the P is a constant of the motion; this is notably the case with the projectors S and A onto the symmetrical and antisymmetrical states respectively:

$$[H, S] = 0 \qquad [H, A] = 0.$$

S and A are therefore constants of the motion. It was seen in Chapter XIV that this property of the operators S and A is a necessary condition for the internal coherence of the symmetrization postulate.

---

[1]) This is the case with the observables associated with infinitesimal displacements (cf. note on p. 650).

(v) *Charge independence and conservation of isotopic spin.* Let $\boldsymbol{T}$ be the total isotopic spin of a system of nucleons. If the nucleon-nucleon forces are charge-independent, the Hamiltonian of the system will be invariant with respect to rotations in isotopic-spin space, i.e.

$$[H, \boldsymbol{T}] = 0. \tag{XV.56}$$

Therefore, the components of $\boldsymbol{T}$ are constants of motion, and so is any observable that is a function of these components. In particular $\boldsymbol{T}^2$ and T_z are constants of the motion (N.B. The conservation of T_z is simply the conservation of charge).

13. Invariance Properties and the Evolution of Dynamical States

We shall now show that if H is invariant with respect to a group $\mathcal{G}$, then so also is the evolution operator $U(t, t_0)$. The latter is by definition the solution of the integral equation

$$U(t, t_0) = 1 + \frac{1}{i\hbar} \int_0^t H\, U(t', t_0)\, dt'.$$

Multiplying both sides on the left by T_i and on the right by $T_i^\dagger$, and taking into account the unitarity of T_i and property (XV.53), we find

$$T_i\, U(t, t_0)\, T_i^\dagger = 1 + \frac{1}{i\hbar} \int_{t_0}^t H\, T_i\, U(t', t_0)\, T_i^\dagger\, dt'.$$

Since U and $T_i\, U\, T_i^\dagger$ satisfy the same integral equation, they are equal, which gives

$$[U(t, t_0), T_i] = 0 \qquad \text{for any operator } T_i \text{ of } G. \tag{XV.57}$$

Among the physical consequences of the invariance of U in the transformations of the group are the conservation laws of the previous paragraph. Moreover, if $|\psi(t)\rangle$ is a solution of the equations of motion, then since U and T_i commute, the transformed vector $T_i\, |\psi(t)\rangle$ is also a solution. The dynamical states represented by these two vectors therefore correspond at each instant by the transformation $\mathcal{T}_i$. Consequently the law of motion of the dynamical *states* is invariant with respect to the transformations of the group $\mathcal{G}$:

Two dynamical states initially the transforms one of the other by a certain transformation $\mathcal{T}_i$ of the group $\mathcal{G}$, conserve this property in the course of time.

This invariance property may be equivalently expressed in terms of measuring operations:

Suppose that after having prepared the system in a certain way at time t_0, we perform a specific measuring operation upon it at a later time t. The result of this measurement is not modified if, without changing anything else, we effect a given transformation $\mathcal{T}_i$ of the group both on the initial state (i.e. on the apparatus used to prepare the system) *and on the quantity (quantities) measured* (i.e. on the apparatus used to observe it).

Suppose, for example, that the system is prepared in the pure state represented by the vector $|\varphi\rangle$ and that we measure the probability for it to be found in the pure state represented by $|\chi\rangle$. The transforms of these states are respectively represented by the vectors $T_i|\varphi\rangle$ and $T_i|\chi\rangle$. From the commutation rule (XV.57), we deduce the equality of the probabilities

$$|\langle\chi|T_i^\dagger U(t, t_0) T_i|\varphi\rangle|^2 = |\langle\chi|U(t, t_0)|\psi\rangle|^2. \qquad \text{(XV.58)}$$

More generally, let ϱ_0 be the density operator representing the state of the system at the time of preparation t_0, and ϱ that representing it at the time the measurement is made, t. A typical measurement would consist in determining the probability that the value(s) of the quantity (quantities) measured will be found in a certain domain D. If we denote by P_D the projector onto the subspace of the eigenstates corresponding to this domain, that probability is

$$w = \mathrm{Tr}\,\varrho\, P_D = \mathrm{Tr}\, U(t, t_0)\, \varrho_0\, U^\dagger(t, t_0)\, P_D. \qquad \text{(XV.59)}$$

Suppose now that we start from the initial state $\varrho_0' \equiv T_i \varrho_0 T_i^\dagger$ and that the measurement is made on the transform(s) of the quantity (quantities) considered in the first experiment. The projector corresponding to the eigenvalues of the domain D is $P_D' \equiv T_i P_D T_i^\dagger$. The result of this new measurement will therefore be

$$w' = \mathrm{Tr}\, U(T_i \varrho_0 T_i^\dagger)\, U^\dagger(T_i P_D T_i^\dagger). \qquad \text{(XV.60)}$$

Taking into account the properties of the trace and the unitarity of T_i, relation (XV.57) gives

$$w' = w. \qquad \text{(XV.61)}$$

In fact, if (XV.58) is valid for any $|\varphi\rangle$ and $|\chi\rangle$, (XV.61) necessarily follows.

What has been done above was to assume that H is invariant with respect to the transformations of a given group $\mathcal{G}$, and to deduce, among other things, that the "law of motion" of the dynamical states is invariant with respect to $\mathcal{G}$. Conversely, we may *postulate that the law of motion is invariant with respect to* $\mathcal{G}$, *and explore the consequences of this postulate.* Such a postulate is equivalent to assuming that for every transformation $\mathcal{T}_i$ of the group, equation (XV.58) is satisfied whatever $|\varphi\rangle$ and $|\chi\rangle$, or (theorem II), that U and $T_i^\dagger U T_i$ are equal to within a phase:

$$T_i^\dagger U(t, t_0) T_i = e^{i\alpha_i} U(t, t_0). \tag{XV.62}$$

The phase factors $e^{i\alpha_i}$ are subject to severe limitations. For most of the groups encountered in physics, they are necessarily all equal to 1 [1]).

We shall always suppose this phase condition fulfilled, even when considerations of internal coherence do not require it. The invariance postulate may then be written

$$[T_i, U(t, t_0)] = 0.$$

In the case of the infinitesimal U operator, $U(t + dt, t) = 1 - (i/\hbar) H\, dt$, we recover the symmetry of the Hamiltonian:

$$[T_i, H] = 0.$$

Thus any postulate of invariance for the law of motion leads to a symmetry property of the Hamiltonian. For an isolated quantum system (no external field) one usually postulates invariance with respect to the displacement group; this amounts to supposing space

---

[1]) This is due to the fact that the $e^{i\alpha_i}$ constitute a one-dimensional representation of the group $\mathcal{G}$, and that they are continuous functions of t tending to 1 when $t \to t_0$. If $\mathcal{G}$ is a finite group, or at least if $\mathcal{T}_i$ is an element of a finite group, there exists an integer p such that $\mathcal{T}_i^p = \mathcal{I}$; it follows that $e^{i\alpha_i}$ is one of the pth roots of unity. The above-mentioned continuity condition requires that $e^{i\alpha_i} = 1$. Reflections in a point and reflections in a plane are examples of such transformations. If the only one-dimensional representation of the group $\mathcal{G}$ is the identical representation we obviously have $e^{i\alpha_i} = 1$ for every operation of the group; the rotation group and the displacement group fall into this category. On the other hand, if the invariance group of the law of motion is limited to the group of translations, this alone does not necessarily lead to all of the phases α_i being null.

to be homogeneous (translational invariance) and isotropic (rotational invariance). Up to the present this postulate has never come into conflict with experiment.

For a great many years it was also thought that the motion of physical systems was invariant under reflection. This is verified experimentally for all phenomena where the only interactions involved are the electromagnetic interactions and the so-called nuclear interactions, i.e. the interactions responsible for the cohesion of atomic nuclei. However, experiment shows it to be violated by certain interactions, and in particular by those responsible for the β decay of atomic nuclei. These interactions are much weaker than the preceding ones; whenever they can be neglected the motion of a physical system is invariant under reflection and the parity is conserved. This is notably the case in atomic physics where only electromagnetic interactions are involved.

When an external field is present, the invariance properties of the law of motion depend on the symmetries of the field. By way of illustration we shall consider two examples taken from atomic physics – the Stark effect and the Zeeman effect.

14. Symmetries of the Stark and Zeeman Effects

STARK EFFECT

Consider an atomic system in an external electric field $\mathscr{E}$ directed, say, along the z axis. This field is invariant under translation, under rotation about the z axis and under reflection in planes parallel to the z axis; the invariance group involved here is the product of the translation group by the group of reflections in planes containing Oz.

We suppose the motion of the center of mass already separated out, and consider only the possible states of the Hamiltonian H of the relative variables; the latter is invariant with respect to the group of reflections in planes containing Oz. The considerations to follow are based uniquely on this invariance property; they involve neither the detail nor the strength of the coupling of the system with the electric field.

Let $\mathscr{S}_u$ be the reflection in an arbitrarily chosen plane containing Oz, and S_u the corresponding operator. The phase of S_u will be taken so as to have $S_u^2 = 1$. Since the group in question is generated by the reflection $\mathscr{S}_u$ and the infinitesimal rotations about Oz, any transformation operator of the group is a function of S_u and of J_z. Thus

we have in all two independent constants of the motion S_u and J_z. Since they do not commute,

$$S_u J_z S_u = -J_z. \tag{XV.63}$$

Certain of the eigenvalues of H are necessarily degenerate.

Continuing, we note that J_z^2 commutes with both J_z and S_u and therefore with all the operators of the group. J_z^2 is the observable of type J of the group (definition of § 11). As the observable of type M we may take either J_z or S_u. In the first case we classify the stationary states according to the eigenvalues of J_z, in the second according to those of the set (J_z^2, S_u).

Suppose that the system contains an even number of half-integral spins. J_z can then take all integral values; we denote a particular one of these by m. If $|\rangle$ is a stationary state corresponding to this value, then $S_u |\rangle$ is a stationary state of the same energy corresponding to the eigenvalue $-m$ of J_z. If $m \neq 0$, these two states are necessarily orthogonal. Thus, if we classify the stationary states according to the eigenvalues of J_z, two opposite eigenvalues m and $-m$ give the same energy spectrum, each level of the spectrum having the same order of degeneracy. In other words, the energy levels depend only on $|m|$ and all the levels corresponding to $|m| \neq 0$ present a degeneracy of even order.

We can equally well obtain this result by classifying the stationary states according to the eigenvalues of the pair (J_z^2, S_u). To denote these eigenvalues we use the symbols m^+ and m^-; m is a non-negative integer whose square is equal to the eigenvalue of J_z^2 and the superscript is positive or negative according as the eigenvalue of S_u is $+1$ or -1. As an example consider a state m^+:

$$J_z^2 |m^+\rangle = m^2 |m^+\rangle \qquad S_u |m^+\rangle = |m^+\rangle.$$

If $m \neq 0$, the vector $J_z |m^+\rangle$ does not vanish and we have, from (XV.63),

$$S_u (J_z |m^+\rangle) = -J_z S_u |m^+\rangle = -(J_z |m^+\rangle).$$

Thus J_z operating on any m^+ state gives a m^- state; if the first is a stationary state, the second is also and corresponds to the same energy level: the energy levels depend only on the positive integer m and all present a degeneracy of even order. The case $m = 0$ is an exception: the energy spectra of the states 0^+ and 0^- may be different.

ZEEMAN EFFECT

Consider now an atomic system in a constant magnetic field $\mathscr{H}$ directed along the z axis. In a reflection in a plane parallel to this axis, $\mathscr{H}$ changes sign. On the other hand, $\mathscr{H}$ is invariant in a reflection $\mathscr{S}_0$ in the origin. The invariance group of the external field is thus the group generated by the translations, the rotations about Oz and the reflection $\mathscr{S}_0$. As in the case of the Stark effect, we examine only the symmetries of the Hamiltonian H of the relative variables. H is invariant in rotation about Oz and in reflection in the origin. Let S_0 be the parity operator and let its phase be fixed so as to have $S_0^2 = 1$. All of the transformations of the group are expressible as functions of the observables J_z and S_0 and these two observables commute:

$$S_0 J_z S_0 = J_z.$$

Therefore, the invariance of H with respect to the transformations of this group does not lead to any systematic degeneracy [1]).

H, S_0 and J_z can be simultaneously diagonalized and each common eigenvector of these three observables represents a stationary state invariant with respect to the transformations of the group. This result is based uniquely on the symmetry properties of H; it is independent of the detail and of the strength of the coupling of the system with the magnetic field.

IV. TIME REVERSAL AND THE PRINCIPLE OF MICROREVERSIBILITY

15. Time-translation and Conservation of Energy

Of all the transformations involving the time the simplest are the time-translations. In Classical Mechanics, the invariance of the equations of motion under time-translation leads to the well-known conservation of energy law (it requires the Hamiltonian function to be time-independent). We obtain an analogous property in Quantum Mechanics.

Let $|\psi(t)\rangle$ be a possible solution of the equations of motion. To state that the law of motion of the system is invariant under a translation

[1]) In other words, the irreducible representations of the group are necessarily all of degree 1.

τ of the time is equivalent to stating that there is another solution $|\psi'(t)\rangle$ which represents at time t the dynamical state that the previous solution represents at time $t+\tau$, i.e.,

$$|\psi'(t)\rangle = e^{i\alpha(t,\tau)}\,|\psi(t+\tau)\rangle.$$

In order for every solution of the equations of motion to have this property, it is necessary that the operator U have the property

$$U(t, 0) = e^{i\alpha(t,\tau)}\,U(t+\tau, \tau), \tag{XV.64}$$

where $\alpha(t, \tau)$ is a phase that may depend on t and τ. For infinitely small t, putting $f(\tau) = \partial\alpha/\partial t|_{t=0}$, we have

$$1 - \frac{i}{\hbar}H(0)\,dt = (1 + i\,f(\tau)\,dt)\left(1 - \frac{i}{\hbar}H(\tau)\,dt\right),$$

i.e.

$$H(\tau) = H(0) + \hbar\,f(\tau). \tag{XV.65}$$

If the law of motion is invariant in any time-translation then (XV.64) must hold for all τ. In other words, the Hamiltonian is constant to within the addition of a (real) function of the time. In actual fact, this function can be supposed null without modifying any of the physical properties of the system. To replace the Hamiltonian $H(t)$ by its value at time $t=0$ only has the effect of multiplying $U(t, t_0)$ by the phase factor $[i\int_0^t f(t')\,dt']$ (Problem XV.6).

Thus, we may suppose the Hamiltonian to be time-independent if the law of motion is invariant under time-translations. This will always be done in what follows. Eq. (XV.64) then reduces to the invariance of $U(t, t_0)$ in time-translations

$$U(t+\tau, \tau) = U(t, 0). \tag{XV.66}$$

16. Time Reversal in Classical Mechanics and in Quantum Mechanics

The systems to be considered in the remainder of this section are all conservative. It often occurs that the laws of motion for such systems are invariant not only with respect to time-translation but also with respect to time reversal. We also encounter this type of invariance in Classical Mechanics.

The Lagrangian function $L_{\text{cl}}(\dot{q}, q)$ of Classical Mechanics is a second-degree polynomial with respect to the velocities. For many systems, terms of the first degree are absent and we have

$$L_{\text{cl}}(\dot{q}, q) = L_{\text{cl}}(-\dot{q}, q).$$

Systems of isolated particles always have this symmetry property. It is not necessarily destroyed by the introduction of a static external field; it is preserved, for example, in a pure electric field. On the other hand a magnetic field introduces a linear coupling with respect to the velocities and therefore destroys it. When it holds, the momenta p are homogeneous linear functions of the velocities and the Hamiltonian function has an analogous symmetry property.

To make the discussion less formal we shall examine the consequences of such a symmetry for the simple particular case of a particle in a static potential. In this case we have

$$H(\boldsymbol{p}, \boldsymbol{r}) \equiv \frac{\boldsymbol{p}^2}{2m} + V(\boldsymbol{r}) = H(-\boldsymbol{p}, \boldsymbol{r}). \tag{XV.67}$$

It follows that all solutions $\boldsymbol{r}(t)$ of the equations of motion are reversible with respect to time: the function $\boldsymbol{r}_{\text{rev}}(t)$ defined by

$$\boldsymbol{r}_{\text{rev}}(t) = \boldsymbol{r}(-t) \tag{XV.68}$$

is also a solution of the equations of motion. The correspondence between these two solutions is represented in Figure XV.1. The

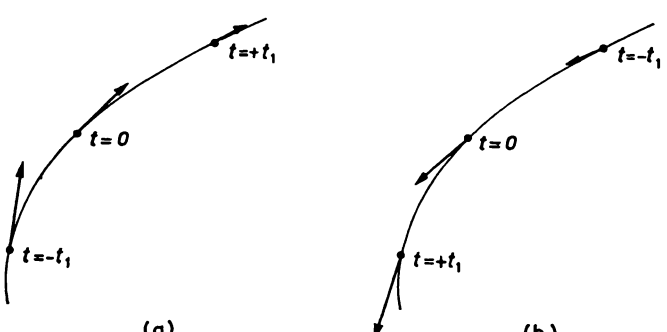


Fig. XV.1. Configuration space representation of two classical trajectories (a) and (b) that correspond by time reversal:

$$\boldsymbol{r}_a(t) = \boldsymbol{r}_b(-t), \qquad \dot{\boldsymbol{r}}_a(t) = -\dot{\boldsymbol{r}}_b(-t).$$

position of the particle at time t in one of the solutions is equal to its position at time $-t$ in the other; its velocity at time t in one solution is the opposite of its velocity at time $-t$ in the other. The correspondence between momenta is the same as the correspondence between velocities:

$$\boldsymbol{p}_{\text{rev}}(t) = -\boldsymbol{p}(-t). \tag{XV.69}$$

Consider now the analogous quantum system. Its Schrödinger equation is

$$i\hbar \frac{\partial}{\partial t} \psi(\boldsymbol{r}, t) = \left[-\frac{\hbar^2}{2m} \triangle + V(\boldsymbol{r}) \right] \psi(\boldsymbol{r}, t). \tag{XV.70}$$

The Hamiltonian being a real operator, if we change t into $-t$ and take the complex conjugate of both sides we have

$$i\hbar \frac{\partial}{\partial t} \psi^*(\boldsymbol{r}, -t) = \left[-\frac{\hbar^2}{2m} \triangle + V(\boldsymbol{r}) \right] \psi^*(\boldsymbol{r}, -t). \tag{XV.71}$$

In other words, if $\psi(\boldsymbol{r}, t)$ is a solution of the Schrödinger equation, then so is the function

$$\psi_{\text{rev}}(\boldsymbol{r}, t) \equiv \psi^*(\boldsymbol{r}, -t). \tag{XV.72}$$

The correspondence between ψ and ψ_{rev} is strikingly analogous to the correspondence between the two classical solutions discussed above [eqs. (XV.68–69)]; $P(\boldsymbol{r}, t)$ and $\Pi(\boldsymbol{p}, t)$ denoting respectively the probability densities for position and momentum at time t, one can easily show that

$$P_{\text{rev}}(\boldsymbol{r}, t) = P(\boldsymbol{r}, -t) \tag{XV.68'}$$

$$\Pi_{\text{rev}}(\boldsymbol{p}, t) = \Pi(-\boldsymbol{p}, -t). \tag{XV.69'}$$

17. The Time-reversal Operation. Spinless Particle

As seen in the preceding example, the reversibility of the solutions of the Schrödinger equation with respect to time is due to the invariance of the Hamiltonian $H(\boldsymbol{p}, \boldsymbol{r})$ when $\boldsymbol{p}$ is changed into $-\boldsymbol{p}$. Indeed, this is why it is represented by a real differential operator in Wave Mechanics. We are thus led to define a transformation of the dynamical variables and dynamical states, which we shall call *time reversal*, in which $\boldsymbol{r}$ and $\boldsymbol{p}$ transform respectively into $\boldsymbol{r}$ and $-\boldsymbol{p}$. We denote the operator effecting this transformation by K and the transformation itself by $\mathscr{K}$. By definition

$$K\, \boldsymbol{r}\, K^{\dagger} = \boldsymbol{r} \qquad K\, \boldsymbol{p}\, K^{\dagger} = -\boldsymbol{p}. \tag{XV.73}$$

The transformation changes the sign of the commutation relations, and K is therefore an *antiunitary operator* (cf. § 6). It is defined by relations (XV.73) up to a phase factor. Let K_0 be the complex-conjugation operator associated with the representation of Wave Mechanics (this type of antiunitary operator was defined in § 5). Since the matrices representing $\boldsymbol{r}$ and $\boldsymbol{p}$ in this representation are respectively real and pure imaginary, K_0 obviously satisfies relations (XV.73). Therefore we may take K_0 as our time reversal operator:

$$K = K_0.$$

With this choice of phase, the action of K on a wave function gives its complex conjugate:

$$K\Phi(\boldsymbol{r}) = \Phi^*(\boldsymbol{r}).$$

To suppose H invariant in the substitution of $-\boldsymbol{p}$ for $\boldsymbol{p}$ amounts to supposing that

$$[K, H] = 0. \tag{XV.74}$$

The application of the (antiunitary) operator K to both sides of the Schrödinger equation then gives

$$-i\hbar \frac{\partial}{\partial t} K |\psi(t)\rangle = HK |\psi(t)\rangle,$$

i.e.

$$i\hbar \frac{\partial}{\partial t} (K |\psi(-t)\rangle) = H(K |\psi(-t)\rangle).$$

Thus, if $|\psi(t)\rangle$ satisfies the Schrödinger equation, then so does the vector

$$|\psi(t)\rangle_{\text{rev}} \equiv K |\psi(-t)\rangle. \tag{XV.75}$$

The dynamical state represented by the vector $|\psi\rangle_{\text{rev}}$ at a given time t is the time-reversal transform of the state represented by the vector $|\psi\rangle$ at time $-t$. This is just the reversibility property of the solutions of the Schrödinger equation, as found in the preceding paragraph.

From the relations (XV.73) that define it, there results that the transformation $\mathscr{K}$ commutes with all of the spatial transformations (translations, rotations and reflections). We also note that K commutes

with the spatial-transformation operators, as defined in section II [1]). In particular K anticommutes with the three components of momentum and therefore commutes with the infinitesimal-translation operators; similarly it anticommutes with the three components of angular momentum,

$$K\,(\boldsymbol{r}\times\boldsymbol{p})\,K^{\dagger}=-(\boldsymbol{r}\times\boldsymbol{p}), \tag{XV.76}$$

and therefore commutes with the infinitesimal-rotation operators.

18. General Definition of Time Reversal

In order to extend the notion of time reversal to the most general systems of particles we must define the time reversal of the spin variables. Since the spin is a particular angular momentum, it must transform like the orbital angular momentum [eq. (XV.76)], i.e.

$$K\,\boldsymbol{s}\,K^{\dagger}=-\boldsymbol{s}. \tag{XV.77}$$

The time-reversal operation inverts the spin. This definition preserves the property of commutation of $\mathscr{K}$ with the spatial transformations and notably with the rotations. Moreover, according to (XV.76) and (XV.77), K anticommutes with the components of the total angular momentum $\boldsymbol{J}$:

$$K\,\boldsymbol{J}\,K^{\dagger}=-\boldsymbol{J} \tag{XV.78}$$

and consequently commutes with the rotation operators (cf. Problem XV.8), i.e. since K is antilinear, (XV.78) leads to:

$$K\mathrm{e}^{-\mathrm{i}(\boldsymbol{J}\cdot\boldsymbol{u})\varphi/\hbar}\,K^{\dagger}=\exp\left[+\frac{\mathrm{i}}{\hbar}\,\varphi(K(\boldsymbol{J}\cdot\boldsymbol{u})K^{\dagger})\right]=\mathrm{e}^{-\mathrm{i}(\boldsymbol{J}\cdot\boldsymbol{u})\varphi/\hbar}.$$

We next construct the time-reversal operator for a particle of spin s.

---

[1]) If the transformation $\mathscr{K}$ commutes with another transformation $\mathscr{T}$, we necessarily have $KT=\mathrm{e}^{\mathrm{i}\alpha}TK$, the phase factor $\mathrm{e}^{\mathrm{i}\alpha}$ depending on the choice of the operator T. As may easily be seen, the latter is defined up to a sign if we further require that $KT=TK$.

This permits us to draw certain conclusions as to the possible complexity of the group of operators G associated with a group of transformations $\mathscr{G}$, when the latter commutes with $\mathscr{K}$. We may then take for (T_i) [notation of § 8] the two transformation operators that commute with K. In particular (1) denotes the pair $(+1,-1)$. The set $\{T\}$ formed of these pairs (the one the negative of the other) of operators is obviously a group, and may be chosen as the group G.

This operator is defined by (XV.73) and (XV.77). We denote by K_0 the complex-conjugation operator associated with the $\{\boldsymbol{r}, s_z\}$ representation, where the relative phases of the basis vectors are fixed by the usual conventions and in particular the basis vectors of spin space are those of the "standard" choice defined in Chapter XIII. We therefore have

$$K_0\,\boldsymbol{r}\,K_0=\boldsymbol{r} \qquad K_0\,\boldsymbol{p}\,K_0=-\boldsymbol{p} \tag{XV.79}$$

$$K_0\,s_x\,K_0=s_x \qquad K_0\,s_y\,K_0=-s_y \qquad K_0\,s_z\,K_0=s_z. \tag{XV.80}$$

Put

$$K=TK_0. \tag{XV.81}$$

Since $T=KK_0$ and $T^\dagger=K_0K^\dagger$, the (*linear*) unitary transformation T gives:

$$T\,\boldsymbol{r}\,T^\dagger=\boldsymbol{r} \qquad T\,\boldsymbol{p}\,T^\dagger=\boldsymbol{p} \tag{XV.82}$$

$$T\,s_x\,T^\dagger=-s_x \qquad T\,s_y\,T^\dagger=s_y \qquad T\,s_z\,T^\dagger=-s_z. \tag{XV.83}$$

Eqs. (XV.82) and (XV.83) give the transformation of the variables $\boldsymbol{r}$, $\boldsymbol{p}$, $\boldsymbol{s}$, *in a rotation of the spin alone through an angle* π *about the* y *axis.* Let $Y^{(s)}$ be the operator effecting this rotation

$$Y^{(s)}=\mathrm{e}^{-\mathrm{i}\pi s_y/\hbar}.$$

T and $Y^{(s)}$ differ only by a phase factor, and since this factor has no physical significance we shall put it equal to 1. This gives

$$K=Y^{(s)}K_0=\mathrm{e}^{-\mathrm{i}\pi s_y/\hbar}\,K_0. \tag{XV.84}$$

For the special case of a particle of spin $\frac{1}{2}$

$$K=-\mathrm{i}\sigma_yK_0. \tag{XV.85}$$

All of this can be extended without difficulty to systems of N particles. The operator K becomes the tensor product of the time-reversal operators for each individual particle. If K_0 is the complex-conjugation operator associated with the standard representation $\{\boldsymbol{r}^{(1)}\,s_z^{(1)}\ldots\boldsymbol{r}^{(N)}\,s_z^{(N)}\}$, and $Y^{(S)}$ the operator for the rotation of the intrinsic spins through an angle π about the $0y$ axis, we find, with the choice of phase adopted above,

$$\boxed{K=Y^{(S)}K_0=\mathrm{e}^{-\mathrm{i}\pi S_y/\hbar}\,K_0} \tag{XV.86}$$

19. Time Reversal and Complex Conjugation

The time-reversal operation $\mathscr{K}$ presents many analogies with complex conjugation. By extension, two linear operators that are time-reversal transforms one of the other will henceforth be called *complex conjugates*. In particular, an operator Q is

(i) *real* if $K\,Q\,K^{\dagger}=Q$

(ii) *pure imaginary* if $K\,Q\,K^{\dagger}=-Q$.

Any real constant is a real operator. The constant i is a pure imaginary operator. The product of i with a real operator is a pure imaginary; the sum and product of two real operators are real operators.

The notion of reality is not to be confused with that of Hermiticity. In this respect we make the following two remarks:

(a) Contrary to what happens in Hermitean conjugation, two complex-conjugate operators are not necessarily represented in a given representation by complex-conjugate matrices.

(b) The above definition of complex conjugation is not the only one possible; any antiunitary transformation whose square is equal to $\mathscr{I}$ can be treated as a complex conjugation.

In *our* definition of complex conjugation, the position observables are all real, the momenta and spins are all imaginary, and the spatial-transformation operators, as defined in section II, are real (Problem XV.8).

The notion of complex conjugation can be extended to include vectors. We define the *complex conjugate of the vector* $|\rangle$ *to be the vector* $K\,|\rangle$. This is not a reciprocal correspondence unless $K^2=1$. It does not necessarily follow from $\mathscr{K}^2=\mathscr{I}$ that $K^2=1$, but only that K^2 commutes with all the dynamical variables, and is therefore a constant. It may easily be shown (Problem XV.9) that the value of this constant is independent of the choice of phase adopted for the definition of K and that the only possible values are ± 1. Moreover this value may be calculated from expression (XV.86) for K; since K_0 commutes with $Y^{(S)}$ and $K_0^2=1$, we find

$$K^2=(Y^{(S)})^2=\mathrm{e}^{-2\pi i S_y/\hbar} \tag{XV.87}$$

or, denoting by n the number of particles in the system with half-integral spins,

$$\boxed{K^2=(-)^n} \tag{XV.88}$$

If $K^2=1$ (n even), the complex conjugation of vectors is a reciprocal correspondence and we can define real vectors. A vector $|r\rangle$ is real if

$$K\,|r\rangle=|r\rangle.$$

By definition a *real representation* is a representation whose basis vectors are all real.

To set up a real basis we may proceed in the following way. We start with an arbitrarily chosen vector $|a\rangle$ and by linear combination of $|a\rangle$ and $K\,|a\rangle$ form the real vector of norm 1

$$|a^r\rangle=c\,|a\rangle+c^*(K\,|a\rangle).$$

The real character of $|a^r\rangle$ is obvious. The constant c is to be adjusted so as to make the norm equal to 1; one easily verifies that it is always possible to do so. We then take a vector $|b\rangle$ orthogonal to $|a^r\rangle$ and form in the same way the real vector $|b^r\rangle$ of norm 1. Since $K\,|a^r\rangle=|a^r\rangle$ and since by hypothesis $\langle a^r\,|b\rangle=0$, we have

$$\langle a^r\,|(K\,|b\rangle)=(\langle a^r\,|K^\dagger)\,(K\,|b\rangle)=\langle a^r\,|b\rangle^*=0$$

and consequently $|b^r\rangle$ is orthogonal to $|a^r\rangle$. We then take another vector $|c\rangle$ orthogonal to $|a^r\rangle$ and $|b^r\rangle$ and form the vector $|c^r\rangle$, and so on, until we have formed a complete set.

Real representations have interesting properties. The associated complex-conjugation operator is the operator K itself. Any real operator is represented by a real matrix; two complex-conjugate operators are represented by complex-conjugate matrices. The unitary matrices relating two real representations are real matrices.

If $K^2=-1$ (n odd), there are no real vectors. However, since $K=-K^\dagger$, we have for any vector $|u\rangle$

$$\langle u\,|(K\,|u\rangle)\equiv\langle u\,|(K^\dagger\,|u\rangle)=-\langle u\,|(K\,|u\rangle)=0.$$

We therefore have the interesting property that:

Two complex-conjugate vectors $|\rangle$ *and* $K\,|\rangle$ *are orthogonal.*

In addition, if a vector $|b\rangle$ is orthogonal to two complex-conjugate vectors $|a\rangle$ and $K\,|a\rangle$, then so is its complex conjugate $K\,|b\rangle$. From this property, easy to verify, we deduce, by an argument patterned on the one for the case $K^2=1$, that there exist bases entirely made up of pairs of complex-conjugate vectors.

20. Principle of Microreversibility

Just as we postulate that the law of motion of a given system is invariant under certain spatial transformations (§ 13), we can also postulate that it is reversible with respect to time, i.e. that if $|\psi(t)\rangle = U(t, 0)\,|\psi\rangle$ represents a possible state of motion of the system, then the vector

$$|\psi(t)\rangle_{\text{rev}} = KU(-t, 0)\,|\psi\rangle,$$

also represents a possible state of motion of the system, and is therefore equal, to within a phase factor, to $U(t, 0)\,K\,|\psi\rangle$. This property being assumed to hold for all $|\psi\rangle$, the postulate of reversibility with respect to the time is written

$$U(t, 0)\,K = e^{i\alpha(t)}\,KU(-t, 0).$$

This postulate applies only to conservative systems. If we denote the Hamiltonian by H we have

$$U(t, 0) = e^{-iHt/\hbar} \qquad U(-t, 0) = e^{+iHt/\hbar} = U^{\dagger}(t, 0)$$

from which

$$U(t, 0) = e^{i\alpha}(KU^{\dagger}(t, 0)\,K^{\dagger}). \tag{XV.89}$$

Multiplying both sides on the left by K and on the right by $K^{\dagger}$, and using the fact that $K^2 = K^{\dagger 2} = (-)^n$ we find that

$$U^{\dagger}(t, 0) = e^{i\alpha}(KU(t, 0)\,K^{\dagger}).$$

Comparing this relation with the Hermitean conjugate of (XV.89)

$$U^{\dagger}(t, 0) = e^{-i\alpha}(KU(t, 0)\,K^{\dagger}),$$

we see that $e^{i\alpha}$ can take only the values ± 1; but since $e^{i\alpha}$ is a continuous function of t and is equal to 1 when $t = 0$, we necessarily have $e^{i\alpha} = 1$. Applying (XV.89) to the infinitesimal operator

$$U(\mathrm{d}t, 0) = 1 - \frac{i}{\hbar} H\,\mathrm{d}t$$

gives

$$1 - \frac{i}{\hbar} H\,\mathrm{d}t = K\left[1 + \frac{i}{\hbar} H\,\mathrm{d}t\right] K^{\dagger},$$

i.e.

$$KHK^{\dagger} = H. \tag{XV.90}$$

Thus, *if the law of motion of a conservative system is reversible with*

respect to the time, the Hamiltonian is real, and conversely (the converse was proved in § 17).

The reversibility postulate is usually formulated under the name of the *principle of microreversibility*, in a way that differs somewhat from the above. This other formulation is as follows.

Let w be the probability of finding the system (supposed conservative) at time t in a certain state $|\chi\rangle$ when it was prepared at time t_0 in the state $|\varphi\rangle$; let w_{rev} be the probability of finding it at time t in the state $K|\varphi\rangle$ when it was prepared at time t_0 in the state $K|\chi\rangle$. The microreversibility principle states that

$$w_{\text{rev}} = w \tag{XV.91}$$

for all $|\varphi\rangle$, $|\chi\rangle$, t_0 and t.

Condition (XV.91) can be written

$$|(\langle\varphi|K^\dagger)(U(t, t_0)\,K|\chi\rangle)|^2 = |\langle\chi|U(t, t_0)|\varphi\rangle|^2.$$

This relation may be compared with relation (XV.58). Since we have

$$(\langle\varphi|K^\dagger)(UK|\chi\rangle) = \langle\varphi|(K^\dagger UK)|\chi\rangle^* = \langle\chi|(K^\dagger U^\dagger K)|\varphi\rangle$$

it can be written

$$|\langle\chi|(K^\dagger U^\dagger K)|\varphi\rangle| = |\langle\chi|U|\varphi\rangle|.$$

Since this relation is satisfied whatever $|\varphi\rangle$ and $|\chi\rangle$, U and $K^\dagger U^\dagger K$ are equal to within a phase factor (Theorem II). The argument used above in connection with $e^{i\alpha}$ can be repeated to show that this phase factor is necessarily equal to 1. The microreversibility principle therefore requires that

$$\boxed{U(t, t_0) = K^\dagger U^\dagger(t, t_0)K} \tag{XV.92}$$

which, in the case of the infinitesimal U operator, is just the condition for the invariance of the Hamiltonian, (XV.90) [1].

It is currently assumed that any quantum system evolving in the absence of external fields satisfies the microreversibility principle. Up

---

[1] Since $U^\dagger(t_2, t_1) = U(t_1, t_2)$, equation (XV.92) may also be written

$$U(t_2, t_1) = K^\dagger U(t_1, t_2)K.$$

to now this hypothesis has not come into conflict with experiment. In the presence of an external field the principle is, or is not, satisfied according to whether the field is, or is not, invariant with respect to time reversal. An electrostatic field has this invariance property. This is easily understood since the sources of such a field are fixed electric charges, and a static charge distribution is not modified by time reversal. The sources of a static magnetic field, on the other hand, are fixed electric currents: time reversal reverses the currents and therefore also the fields. Hence the microreversibility principle cannot be satisfied in the presence of a magnetic field even a time-indepenent magnetic field.

21. Consequence: Kramers Degeneracy

As with any other symmetry, the realness of H is reflected in certain special properties of the corresponding eigenvalue problem.

If $|u\rangle$ is an eigenvector of H,

$$H|u\rangle = E|u\rangle,$$

its complex conjugate $K|u\rangle$ is also an eigenvector of H and corresponds to the same eigenvalue E, i.e. by hypothesis

$$HK = KH$$

and, since E is real,

$$H(K|u\rangle) = K(H|u\rangle) = (KE|u\rangle) = E(K|u\rangle).$$

Consequently the subspace $\mathscr{E}_E$ of the eigenvalue E is invariant in the antiunitary transformation K and we may apply the results of § 19. There are two cases to be considered according as $K^2 = \pm 1$.

1st CASE. $K^2 = +1$ (even number of spins $\frac{1}{2}$).

In each subspace $\mathscr{E}_E$ we can choose an orthonormal basis whose vectors are all real. Thus H has (at least) one basis whose vectors are all *real.*

2nd CASE. $K^2 = -1$ (odd number of spins $\frac{1}{2}$).

In each subspace $\mathscr{E}_E$ we can choose an orthonormal basis made up of pairs of complex-conjugate vectors. Thus each subspace $\mathscr{E}_E$ has an even number of dimensions: *each eigenvalue of H is at least twofold degenerate and its degeneracy is necessarily of even order.* This type of degeneracy is called a *Kramers degeneracy.*

Certain systems have no symmetries other than the invariance under time reversal. Such is the case for the atoms in an asymetrical crystal network. Such an atom may be treated as a system in a purely electrostatic external field. Its Hamiltonian is real. If it has an odd number of electrons, all of its levels are doubly degenerate. The degeneracy is removed by the introduction of a magnetic field.

22. Real Rotation-invariant Hamiltonian

If the Hamiltonian H has symmetry properties other than time reversal, the results of the preceding paragraph remain valid but lose much of their interest. Wigner [1]) has systematically studied the properties of H in the case when

(i) H is real: $[K, H]=0$

(ii) H is invariant in the transformations T_i of a group of *linear* transformations: $[T_i, H]=0$

(iii) the transformations T_i commute with time reversal:

$$[K, T_i]=0.$$

We shall discuss here only the case when the invariance group is the rotation group [2]).

Let Y be the operator for rotation through an angle π about the y

---

[1]) E. P. Wigner, Göttinger Nachrichten, 31 (1932) 546.

[2]) For a given group G, the fact that H is real has the effect of either doubling the "G degeneracy" or making possible the diagonalization of H in a representation whose basis vectors obey a certain "reality" condition. The detailed result is as follows (cf. Wigner, *loc. cit.*):

Let $\mathsf{G}^{(j)}$ be an irreducible representation of G, $\mathsf{G}^{(j)*}$ the conjugate representation. K acting on a vector of one of these representations gives a vector of the other. Either $\mathsf{G}^{(j)}$ and $\mathsf{G}^{(j)*}$ are equivalent or they are inequivalent; in the first case, the matrix S permitting to pass from the one to the other ($\mathsf{G}^{(j)*}= S\mathsf{G}^{(j)}S^{\dagger}$) has the property $SS^* = \pm 1$. Call G the representation of the group G defined in the subspace of a given eigenvalue of H. Three cases have to be considered:

(*a*) If $\mathsf{G}^{(j)*}$ is inequivalent to $\mathsf{G}^{(j)}$, these two representations appear the same number of times in the decomposition of G in irreducible components (degeneracy doubled);

(*b*) if $\mathsf{G}^{(j)*} \simeq \mathsf{G}^{(j)}$ and $SS^*K^2 = -1$, $\mathsf{G}^{(j)}$ appears an even number of times in the decomposition of G (degeneracy doubled);

(*c*) if $\mathsf{G}^{(j)*} \simeq \mathsf{G}^{(j)}$ and $SS^*K^2 = +1$ the basis vectors of each component $\mathsf{G}^{(j)}$ of G can be chosen to satisfy the "reality" condition: $KS\,|u\rangle = |u\langle$.

The representations of the rotation group all correspond to case (c).

axis (not to be confused with the operator for the rotation of the spins alone, $Y^{(S)}$, introduced above). Put

$$K_y \equiv Y^\dagger K \equiv e^{i\pi J_y/\hbar} K. \tag{XV.93}$$

$Y^\dagger$ and K commute with $\boldsymbol{J}^2$ and anticommute with J_z. Therefore K_y commutes with $\boldsymbol{J}^2$ and J_z. It also commutes with J_+ and J_-:

$$Y^\dagger K(J_x \pm i J_y) = Y^\dagger(-J_x \pm i J_y)K = (J_x \pm i J_y)Y^\dagger K.$$

Moreover, since $[Y, K] = 0$ and $Y^2 = K^2 = (-)^n$ [cf. relation (XV.88)],

$$K_y^2 = 1. \tag{XV.94}$$

The antiunitary transformation K_y can be treated as a complex conjugation, as was K in § 19. To distinguish it from the latter, we shall use quotation marks to indicate this new type of conjugation. Thus, a "real" (linear) operator is one that commutes with K_y. The "complex conjugate" of a vector $|\rangle$ is by definition the vector $K_y|\rangle$. Since $K_y^2 = 1$, there exist "real" vectors, "real" representations. The action of a "real" operator on a "real" vector gives a "real" vector; in a "real" representation, the "real" operators are represented by real matrices.

Since $\boldsymbol{J}^2$ and J_z are "real", we can construct a basis in the space of angular momentum $(j\,j)$ whose vectors are all "real". J_+ and J_- also being "real", the vectors of the standard basis that can be constructed from these vectors by the method of § XIII.6, are all "real".

Let $|\tau j\mu\rangle$ be the vectors of a "real" standard basis. If H is invariant under rotation and under time reversal, it will commute with both $\boldsymbol{J}$ and K_y, and will be represented in the $\{\tau j\mu\}$ representation by a real matrix of the form (XV.52), i.e.

$$\langle \tau j\mu | H | \tau' j' \mu' \rangle = \delta_{jj'}\, \delta_{\mu\mu'}\, H^{(j)}_{\tau\tau'}$$

and the $H^{(j)}_{\tau\tau'}$ are all real. It follows that *H has a rotational degeneracy and has in addition at least one orthonormal system of eigenvectors forming a "real" standard basis.*

We can set up a "real" basis in $\mathscr{E}$ by taking tensor products of the basis vectors of simpler spaces. Suppose that we have

$$\mathscr{E} = \mathscr{E}_1 \otimes \mathscr{E}_2 \otimes \dots$$

and that we have defined rotation and time-reversal operators in each component space $\mathscr{E}_i$. With each $\mathscr{E}_i$ there is associated a "complex-conjugation" operator $K_y^{(i)} = Y^{(i)\dagger} K^{(i)}$, and the operator K_y for the whole space is the tensor product of these

$$K_y = K_y^{(1)} K_y^{(2)} \ldots .$$

The tensor product of "real" vectors is a "real" vector. In particular, we can form a set of "real" basis vectors in $\mathscr{E}$ by taking tensor products of vectors forming a "real" standard basis of angular momentum in each component space. Starting from this basis in $\mathscr{E}$ we can form a standard basis for the total angular momentum in this space by the addition of angular momenta. Since the Clebsch–Gordon coefficients are real, the vectors of the standard basis thus formed are all "real".

We end this chapter with a remark concerning the "reality" of the spherical harmonics.

Let $\omega = (\theta, \varphi)$ be the angular coordinates of the position vector $\boldsymbol{r}$ of a particle, and $Y_l^m(\omega)$ the spherical harmonic representing the state of angular momentum (lm). In this particular case ($\hbar = 1$),

$$K_y = \mathrm{e}^{-i\pi l_y} K_0$$

so that [eqs. (B.92) and (C.62)]

$$\begin{aligned} K_y Y_l^m(\omega) &= \mathrm{e}^{-i\pi l_y} Y_l^{m*}(\omega) = \mathrm{e}^{-i\pi l_y} (-)^m Y_l^{-m}(\omega) \\ &= (-)^l Y_l^m(\omega). \end{aligned}$$

The spherical harmonic Y_l^m is therefore "real" if l is even, "pure imaginary" if l is odd; on the other hand, the functions

$$\mathscr{Y}_l^m(\omega) \equiv \mathrm{i}^l Y_l^m(\omega)$$

form a "real" standard basis for the orbital angular momentum.

Let $\omega_p = (\theta_p, \varphi_p)$ be the angular coordinates of the momentum $\boldsymbol{p}$ of the same particle and $Y_l^m(\omega_p)$ the spherical harmonic representing the state of angular momentum (lm) in the $\{\boldsymbol{p}\}$ representation. One can show (Problem XV.10) that

$$K_0 = S_0 K_p,$$

where S_0 is the parity operator [eq. (XV.45)] and K_p the complex-conjugation operator associated with the $\{\boldsymbol{p}\}$ representation. It follows that

$$K_p\, Y_l^m(\omega_p) = Y_l^m(\omega_p).$$

The $Y_l^m(\omega_p)$ form a "real" standard basis for the orbital angular momentum. These considerations of "reality" can be extented without difficulty to irreducible tensor operators in general (cf. Problem XV.12).

EXERCISES AND PROBLEMS

1. We suppose that the state vector space $\mathscr{E}$ of a system is the direct sum $\mathscr{E}_1 + \mathscr{E}_2 + \ldots$ of a certain number of subspaces invariant with respect to the ensemble of the *physical* observables of the system. Show that the projectors $P_1, P_2, \ldots$ onto these subspaces commute with each physical observable, and that the vectors of $\mathscr{E}$ do not all have the property of being simultaneous eigenvectors of a complete set of commuting physical observables.

2. Show that the operator S_0 defined by (XV.45) verifies relations (XV.43) and (XV.44) and notably the relation $S_0\, \boldsymbol{p}\, S_0^{\dagger} = -\boldsymbol{p}$.

3. The operator of reflection in a plane perpendicular to $\boldsymbol{u}$ for a particle of spin $\frac{1}{2}$ is the product of an operator acting on the orbital variables alone with an operator acting on the spin variables alone. Show that for a particular choice of phase factor, the last mentioned operator is equal to $(\boldsymbol{\sigma}\cdot\boldsymbol{u})$. Show that the product $(\boldsymbol{\sigma}\cdot\boldsymbol{v})\,(\boldsymbol{\sigma}\cdot\boldsymbol{u})$ representing two successive reflections in planes perpendicular to the vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ respectively is equal to the operator associated with a rotation of the spin through an angle $2(\boldsymbol{u}, \boldsymbol{v})$ about the axis $(\boldsymbol{u} \times \boldsymbol{v})/|\boldsymbol{u} \times \boldsymbol{v}|$.

4. A beam of particles of spin s is scattered by a spin-dependent potential. Let $\boldsymbol{p}_i$ be the initial momentum. The particles scattered into a certain direction are selected with the aid of a diaphragm. Let $\boldsymbol{p}_f$ be the momentum of these particles and D the scattering plane, i.e. the plane containing $\boldsymbol{p}_i$ and $\boldsymbol{p}_f$. We suppose that the incident particles are non-polarized so that their spin state is represented by the density operator $\varrho_i = 1/(2s+1)$. Show that, if the interaction potential is *invariant under rotation and under reflection in the origin*, then the density operator ϱ_f representing the spin state of the scattered particles is symmetrical with respect to the plane D.

5. Show that the density operator representing the spin state of a particle of spin $\frac{1}{2}$ can be expressed in the form

$$\varrho = \tfrac{1}{2}\,(1 + \boldsymbol{P}\cdot\boldsymbol{\sigma});$$

$\boldsymbol{\sigma} \equiv (\sigma_x, \sigma_y, \sigma_z)$ is twice the spin vector of the particle and $\boldsymbol{P}$ is a vector of length between 0 and 1 that completely defines the state of polarization of the particle.

Returning to Problem XV.4 and supposing $s = \frac{1}{2}$ show that the vector $\boldsymbol{P}_f$ defining the polarization of the scattered particles is *perpendicular to the scattering plane.*

6. Show that if any real function of time is added to the Hamiltonian of a system the evolution operator $U(t, t_0)$ is simply multiplied by a phase factor.

7. In a Galilean transformation of the coordinate system, the variables $\boldsymbol{r}$, $\boldsymbol{p}$, $\boldsymbol{s}$ of a particle are transformed respectively into $\boldsymbol{r} - \boldsymbol{v}t$, $\boldsymbol{p} - m\boldsymbol{v}$, $\boldsymbol{s}$. Show that such a transformation is realized by the operator

$$G(\boldsymbol{v}, t) = \exp\left[\mathrm{i}\boldsymbol{v}\cdot(m\boldsymbol{r} - \boldsymbol{p}t)/\hbar\right].$$

The same transformation applied to a system of N particles is realized by the operator $G(\boldsymbol{v}, t) = \exp\left[\mathrm{i}\boldsymbol{v}\cdot(M\boldsymbol{R} - \boldsymbol{P}t)/\hbar\right]$, where M, $\boldsymbol{R}$ and $\boldsymbol{P}$ are respectively the mass, the position and the momentum of the center of mass.

[N.B. For a given value of $\boldsymbol{v}$, the Galilean transformations form a group; with the phase adopted here, the transformation operators also form a group.] The law of motion is invariant under this transformation if

$$G^{\dagger}(\boldsymbol{v}, t)\, U(t, t_0)\, G(\boldsymbol{v}, t_0) = U(t, t_0) \times \text{phase factor}.$$

Show that this condition is equivalent to the "condition of Galilean invariance of the Schrödinger equation":

$$G^{\dagger}(\boldsymbol{v}, t)\left[\mathrm{i}\hbar\frac{\partial}{\partial t} - H\right]G(\boldsymbol{v}, t) = \left[\mathrm{i}\hbar\frac{\partial}{\partial t} - H\right] + f,$$

where f is an arbitrary real function of the time, and that this condition is indeed fulfilled by the Schrödinger Hamiltonian (one finds: $f = 0$).

8. Show that the translation and rotation operators and the reflection operator S_0 [definition (XV.45)] commute with the time-reversal operator K and that this property does not depend on the choice of the arbitrary phase involved in the definition of K.

9. Show that if the square of an antiunitary operator is equal to a constant, that constant is either $+1$ or -1.

10. Let $K_{\boldsymbol{p}}$ be the complex-conjugation operator associated with the $\{\boldsymbol{p}\}$ representation, K_0 the complex-conjugation operator associated with the $\{\boldsymbol{r}\}$ representation, S_0 the reflection operator [definition (XV.45)]. Show that $K_0 = S_0 K_{\boldsymbol{p}}$. Deduce that in time-reversal the wave function $\Phi(\boldsymbol{p})$ transforms into

$$K\Phi(\boldsymbol{p}) = \Phi^*(-\boldsymbol{p}).$$

11. Suppose that we have a system that is invariant under time-reversal. Let B be a pure imaginary (cf. § 19) observable of this system (momentum, spin, orbital angular momentum, etc.). Show that

(i) if a stationary state is non-degenerate, the average value of B in that state is null;

(ii) the trace of the projection of B onto the subspace of any energy eigenvalue is null.

12. Denote by $T_q^{(k)}$ $(q = -k, -k+1, \ldots, +k)$ the standard components of an irreducible tensor operator $\mathbf{T}^{(k)}$. By definition the complex conjugate of $\mathbf{T}^{(k)}$ is the tensor operator whose $q = 0$ component is the time-reversal transform of $T_0^{(k)}$ (with this definition the Cartesian components of real vector operators are invariant under time reversal, and those of imaginary vector operators change sign). Show that the standard components of the complex conjugate of $\mathbf{T}^{(k)}$ are the operators that are the "complex conjugates" (*in the sense of* § 22) of the corresponding components of $(-)^k \mathbf{T}^{(k)}$; and that in particular in a "real" standard representation the matrix element $\langle \tau j \mu \mid T_q^{(k)} \mid \tau' j' \mu' \rangle$ is real if $i^k \mathbf{T}^{(k)}$ is real, and pure imaginary if $i^k \mathbf{T}^{(k)}$ is pure imaginary.

[N.B. This property has a direct application in angular correlation problems, for it gives, to within a sign, the relative phases of the contributions from the different multipoles.]

« Si les plats que je vous offre sont mal préparés, c'est moins la faute de mon cuisinier que celle de la chimie, qui est encore dans l'enfance. »
(*La Rôtisserie de la reine Pédauque*).

ANATOLE FRANCE

PART FOUR

METHODS OF APPROXIMATION

CHAPTER XVI

STATIONARY PERTURBATIONS

1. General Introduction to Part Four

The few problems in Quantum Mechanics that can be exactly solved relate to very special and very simple systems. In practice no physical system can be studied without approximation; the physicist's art consists precisely in determining the relative importance of the different effects and adopting in each case the appropriate approximation method. There are, properly speaking, as many methods of approximation as there are problems and there is no question of treating them all here. In this book we give only those methods sufficiently general to merit a systematic development. Some have already been treated, notably the classical approximation and the WKB method of Chapter VI, and the phase shift method for scattering problems of Chapter X. The others will be the subject of this fourth part.

Mathematically, the study of a quantum system consists in determining its evolution operator $U(t, t_0)$, or at least determining certain characteristic properties of $U(t, t_0)$. When the Hamiltonian H is independent of time, this reduces to solving the eigenvalue problem for H, for in this case,

$$U(t, t_0) = e^{-\frac{i}{\hbar}H(t-t_0)},$$

and the properties of U are thus directly related to those of H. We have a more complicated situation when H depends on the time; clearly, however, the eigenvalues and eigenfunctions of H (which are then time-dependent) will be of importance for determining the properties of U.

On the whole, collision problems, which bring in the continuous spectrum of eigenvalues, are more difficult than problems relating to bound states. Collision problems are considered in the last chapter of this part (Chapter XIX). The three other chapters (Ch. XVI–XVIII) are concerned essentially with the study of bound states, although the methods set forth can eventually be extended to the study of non-bound states. The bound state problems can themselves be separated into two categories. We have, on the one hand, the

determination of stationary states, or more precisely the solutions of the eigenvalue problem for H; on the other, the problem of transitions from one state to another. Three types of methods are commonly employed for the approximate solution of the eigenvalue problem for H. The first, the WKB method (Ch. VI), is based on the classical approximation: the large quantum number and small wave-length approximation. The second, the stationary perturbation method, starts from the exact solution of the eigenvalue problem for an operator H_0 differing little from H, and the eigenfunctions and eigenvalues of H are then expressed as power series in the difference $H - H_0$. The third, the variational method, can be used, failing the other two, when one has an intuitive idea of the general form of the sought-for eigenfunctions. Stationary perturbations are treated in the present chapter and the variational method in Chapter XVIII. Chapter XVII is devoted to the calculation of transitions and, more generally, to the principle methods of investigating the evolution of bound states when the Hamiltonian is modified in the course of time.

I. PERTURBATION OF A NON-DEGENERATE LEVEL

2. Expansion in Powers of the Perturbation

It is assumed that the Hamiltonian H can be written as the sum of an "unperturbed Hamiltonian" H_0 and a perturbation term which it is convenient to write in the form λV, where λ is a real parameter and V, like H_0, a time-independent Hermitean operator:

$$H = H_0 + \lambda V. \tag{XVI.1}$$

We suppose the eigenvalue problem for H_0 solved. Thus we know its eigenvalues

$$E_1^0, E_2^0, \ldots, E_i^0, \ldots$$

and the corresponding set of eigenvectors $|E_i^0\, \alpha\rangle$; the quantum number α distinguishes between the eigenvectors belonging to a degenerate eigenvalue:

$$H_0|E_i^0\, \alpha\rangle = E_i^0|E_i^0\, \alpha\rangle. \tag{XVI.2}$$

The spectrum of H varies continuously with λ, coinciding with the spectrum of H_0 when $\lambda = 0$. Consider a given eigenvalue E_a^0 of

H_0. We wish to calculate the eigenvalue (or eigenvalues) of H that tend to E_a^0 when $\lambda \to 0$, and to determine the corresponding eigenstates of H. To simplify the writing, we suppose that the spectrum of H_0 is entirely discrete; all what follows remains valid if a part of the spectrum of H_0 is continuous so long as E_a^0 belongs to the discrete spectrum.

The method is particularly simple *when the eigenvalue E_a^0 is non-degenerate*. Throughout the present section we shall assume this to be so.

Let E be the eigenvalue of H that tends to E_a^0 when $\lambda \to 0$. It will also be a non-degenerate eigenvalue. The corresponding eigenvector, $|\psi\rangle$, is defined to within a constant which may be arbitrarily fixed. We shall adopt the following definition for $|\psi\rangle$:

$$H|\psi\rangle = E|\psi\rangle \tag{XVI.3}$$

$$\langle 0|\psi\rangle = \langle 0|0\rangle = 1 \tag{XVI.4}$$

where we have put: $|E_a^0\rangle = |0\rangle$. With this definition, $|\psi\rangle$ tends to $|0\rangle$ when $\lambda \to 0$.

If the perturbation λV is sufficiently small it is reasonable to assume that E and $|\psi\rangle$ can be expanded into rapidly converging power series in λ i.e. we may write:

$$E = E_a^0 + \lambda\varepsilon_1 + \lambda^2\varepsilon_2 + \ldots + \lambda^n\varepsilon_n + \ldots \tag{XVI.5}$$

$$|\psi\rangle = |0\rangle + \lambda|1\rangle + \lambda^2|2\rangle + \ldots + \lambda^n|n\rangle + \ldots. \tag{XVI.6}$$

By retaining only the first terms in these series one obtains approximate expressions for E and $|\psi\rangle$.

The perturbation method consists in determining the successive expansion coefficients in (XVI.5) and (XVI.6). To this end, we substitute expression (XVI.1), (XVI.5) and (XVI.6) into both sides of (XVI.3), which then becomes an equality between two power series in λ. In order that this equality be satisfied, the coefficients of each power of λ must separately be equal, giving, successively:

$$(H_0 - E_a^0)|0\rangle = 0 \tag{XVI.7^0}$$

$$(H_0 - E_a^0)|1\rangle + (V - \varepsilon_1)|0\rangle = 0 \tag{XVI.7^1}$$

$$(H_0 - E_a^0)|2\rangle + (V - \varepsilon_1)|1\rangle - \varepsilon_2|0\rangle = 0 \tag{XVI.7^2}$$

$$\ldots\ldots \qquad \ldots\ldots$$

$$(H_0 - E_a^0)|n\rangle + (V - \varepsilon_1)|n-1\rangle \ldots - \varepsilon_n|0\rangle = 0. \tag{XVI.7^n}$$

$$\ldots\ldots \qquad \ldots\ldots$$

Condition (XVI.4) becomes

$$\langle 0|1\rangle = \langle 0|2\rangle = \ldots = \langle 0|n\rangle = \ldots = 0. \tag{XVI.8}$$

Equation (XVI.7^0) determines the eigenvalue and the eigenvector to the zeroth order. With conditions (XVI.8), equation (XVI.7^1) determines the first-order corrections to these two quantities, equation (XVI.7^2) the second-order corrections, ..., equation (XVI.7^n) the nth-order corrections.

Let us show that equation (XVI.7^n) effectively determines ε_n and $|n\rangle$ in terms of the lower order corrections. To do this we project onto the basis vectors of H_0. Projecting onto $|0\rangle$ we obtain, with the aid of (XVI.8)

$$\varepsilon_n = \langle 0|V|n-1\rangle. \tag{XVI.9}$$

Projecting onto the other basis vectors of H_0 we obtain the corresponding components of $|n\rangle$ along each of them:

$$\langle E^0\alpha|n\rangle = \frac{1}{E_a{}^0 - E^0}\,[\langle E^0\alpha|(V-\varepsilon_1)|n-1\rangle - \varepsilon_2\langle E^0\alpha|n-2\rangle - \ldots \\ \ldots - \varepsilon_{n-1}\langle E^0\alpha|1\rangle] \\ (E_a{}^0 \neq E^0).$$

Since $\langle 0|n\rangle = 0$, $|n\rangle$ is thereby completely determined. It is convenient to put

$$Q_0 \equiv 1 - |0\rangle\langle 0| = \sum_{E^0 \neq E_a{}^0} \sum_\alpha |E^0\alpha\rangle\langle E^0\alpha|$$

and

$$\frac{Q_0}{a} \equiv Q_0 \frac{1}{E_a{}^0 - H^0} Q_0 = \sum_{E^0 \neq E_a{}^0} \frac{\sum_\alpha |E^0\alpha\rangle\langle E^0\alpha|}{E_a{}^0 - E_0}. \tag{XVI.10}$$

With these notations one may write

$$|n\rangle = \frac{Q_0}{a}\,[(V-\varepsilon_1)|n-1\rangle - \varepsilon_2|n-2\rangle - \ldots - \varepsilon_{n-1}|1\rangle]. \tag{XVI.11}$$

Equations (XVI.9) and (XVI.11) are equivalent to (XVI.7^n), which completes the demonstration.

3. First-order Perturbations

The first-order corrections, from equation (XVI.7^1), are obtained by writing equations (XVI.9) and (XVI.11) for the case $n=1$.

Equation (XVI.9) gives the first-order correction to the energy level

$$\boxed{\varepsilon_1 = \langle 0|V|0\rangle} \tag{XVI.12}$$

whence the expression for E to the first order:

$$E = \langle 0|H|0\rangle + O(\lambda^2). \tag{XVI.13}$$

This is just the average value of the Hamiltonian H calculated with the eigenvector of the unperturbed Hamiltonian H_0.

Equation (XVI.11) gives the first-order correction to the eigenvector:

$$|1\rangle = \frac{Q_0}{a}(V - \varepsilon_1)|0\rangle$$

but since $Q_0|0\rangle = 0$ this expression reduces to

$$|1\rangle = \frac{Q_0}{a} V|0\rangle. \tag{XVI.14}$$

The vector $|\psi\rangle$ is therefore given to the first order by the expression

$$|\psi\rangle = \left(1 + \lambda \frac{Q_0}{a} V\right)|0\rangle + O(\lambda^2) \tag{XVI.15}$$

and its norm by

$$\begin{aligned}\langle \psi|\psi\rangle &= \langle 0|0\rangle + \langle \psi|Q_0|\psi\rangle \\ &\simeq \langle 0|0\rangle + \lambda^2 \langle 1|1\rangle = 1 + \lambda^2 \langle 1|V \frac{Q_0}{a^2} V|0\rangle \\ &= 1 + O(\lambda^2).\end{aligned}$$

According to (XVI.14), the components of the first-order correction to $|0\rangle$ along the other basis vectors of H_0 are given by the equation:

$$\lambda \langle E^0\alpha|1\rangle = \frac{\langle E^0\alpha|(\lambda V)|0\rangle}{E_a{}^0 - E^0}. \quad (E^0 \neq E_a{}^0) \tag{XVI.16}$$

The component along $|E^0\alpha\rangle$ is therefore equal to the perturbation matrix element linking $|0\rangle$ to $|E^0\alpha\rangle$ divided by the *energy difference* between these two unperturbed eigenstates. The smallness of this quantity is a measure of how rapidly the perturbation series converges.

4. Ground State of the Helium Atom

As a first example [1]) we use the perturbation method to evaluate the ground-state energy of the helium atom and, more generally, the ground-state energy of any $(Z-2)$ times ionized atom. Such an atom is made up of a nucleus of charge Ze and two electrons. The nucleus is supposed infinitely heavy so that the Hamiltonian H is the Hamiltonian of two electrons in the potential

$$-\frac{Ze^2}{r_1}-\frac{Ze^2}{r_2}+\frac{e^2}{r_{12}}$$

where $\boldsymbol{r}_1$, and $\boldsymbol{r}_2$ are the position vectors of the first and second electrons respectively and $r_{12}=|\boldsymbol{r}_1-\boldsymbol{r}_2|$ their mutual distance.

If one neglects the mutual repulsion term e^2/r_{12}, the Hamiltonian reduces to that of two independent particles in the Coulomb field $-Ze^2/r$, and the corresponding eigenvalue problem can be solved exactly. We shall take this as our "unperturbed Hamiltonian" and treat the e^2/r_{12} as a perturbation. The ground-state of H_0 corresponds to both electrons in the 1s state.

Let E_{H} be the binding energy of the ground-state of the hydrogen atom. The energy of the ground-state of H_0, being the sum of the energies of the two electrons, will then be

$$E_a^0 = -\,2Z^2E_{\mathrm{H}}.$$

The corresponding eigenfunction is the product of the eigenfunctions for each particle, namely (cf. § B3)

$$\Phi_a(\boldsymbol{r}_1, \boldsymbol{r}_2) = \mathrm{e}^{-(r_1+r_2)/a}/\pi a^3,$$

where

$$a=a_0/Z=\hbar^2/Zme^2(\simeq Z^{-1}\times 0.53\times 10^{-8}\ \mathrm{cm}).$$

The effect of the perturbation $V=e^2/r_{12}$ is given, in the first order, by formula (XVI.12) $(\lambda=1)$:

$$\begin{aligned}\varepsilon_1 &= \int \Phi_a V\, \Phi_a \,\mathrm{d}\boldsymbol{r}_1\,\mathrm{d}\boldsymbol{r}_2\\ &= \frac{e^2}{\pi^2a^6}\int \frac{\mathrm{e}^{-2(r_1+r_2)/a}}{|\boldsymbol{r}_1-\boldsymbol{r}_2|}\,\mathrm{d}\boldsymbol{r}_1\,\mathrm{d}\boldsymbol{r}_2.\end{aligned} \tag{XVI.17}$$

---

[1]) This example is taken from L. Pauling and E. B. Wilson, *loc. cit.*, note, p. 614/615.

This is the electrostatic energy of two spherical distributions of electricity with densities $-e\varrho_1(r_1)$ and $-e\varrho_2(r_2)$ respectively, where

$$\varrho_1(r) = \varrho_2(r) = \mathrm{e}^{-2r/a}/\pi a^3. \tag{XVI.18}$$

To calculate the integral

$$I = \int \mathrm{d}\boldsymbol{r}_1 \int \mathrm{d}\boldsymbol{r}_2 \frac{\varrho_1(r_1)\,\varrho_2(r_2)}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|}$$

we employ the expansion (B.99). After integration over the angles we obtain

$$\begin{aligned} I &= 16\pi^2 \int_0^\infty \mathrm{d}r_1 \int_0^\infty \mathrm{d}r_2\, \varrho_1(r_1)\,\varrho_2(r_2)\, r_1^2\, r_2^2/r_> \\ &= 16\pi^2 \int_0^\infty \varrho_1(r_1)\, r_1\, \mathrm{d}r_1 \left\{ \int_0^{r_1} \varrho_2(r_2)\, r_2^2\, \mathrm{d}r_2 + r_1 \int_{r_1}^\infty \varrho_2(r_2)\, r_2\, \mathrm{d}r_2 \right\}. \end{aligned} \tag{XVI.19}$$

After calculating I, with the aid of (XVI.18), we find that $\varepsilon_1 = e^2 I$ is equal to

$$\varepsilon_1 = \tfrac{5}{4} Z E_{\mathrm{H}}. \tag{XVI.20}$$

One expects the error of this approximation to be smaller when the energy of mutual repulsion of the electrons is smaller compared with the energy of attraction of the nucleus, and thus for higher Z. This is seen to be the case in table XVI.1, where the observed binding

TABLE XVI.1

Binding energies for He, Li^+, Be^{++} *in the ground state*

The energies are all given in electronvolts. Columns 2, 3, 4 give respectively the unperturbed energy, the first-order perturbation and the sum of the two i.e., the ground state energy to the first order in perturbation theory. Column 5 gives the result of the variational calculation of § XVIII.6, Column 6 the experimental value for the energy of the ground state (from L. Pauling and E. B. Wilson, *loc. cit.*).

	1	2	3	4	5	6
	Z	E_a^0	ε_1	$E_{\mathrm{pert}} = E_a^0 + \varepsilon_1$	E_{var}	E_{exp}
He	2	−108	34	− 74	− 76.6	− 78.6
Li^+	3	−243.5	50.5	−193	−195.6	−197.1
Be^{++}	4	−433	67.5	−365.5	−368.1	−370.0

energies of the ground states of He, Li^+ and Be^{++} are compared with those given by the perturbation theory. The perturbation expansion converges more rapidly than *a priori* considerations might lead one to think and the calculation of the binding energy to the first order gives a value in reasonable accord with observation, even in the case of helium ($Z=2$).

5. Coulomb Energy of Atomic Nuclei

In an atomic nucleus the distance between nucleons is of the order of 10^{-13} cm. At such distances the so-called nuclear forces are much stronger than the Coulomb repulsion of the protons and the latter may therefore be treated as a perturbation. H_0 will then be the sum of the kinetic energy of the nucleons and the potential energy of specifically nuclear origin; V will be made up of the coulomb repulsion terms

$$V = \sum_{i<j\leqslant Z} \frac{e^2}{r_{ij}} \tag{XVI.21}$$

($r_{ij}=|\boldsymbol{r}_i-\boldsymbol{r}_j|$ = distance between protons i and j).

We wish to apply perturbation theory to calculate the shift E_c due to these Coulomb terms.

Let j denote the spin of the unperturbed state. This state has a degeneracy of order $(2j+1)$, i.e. there exist $(2j+1)$ orthonormal unperturbed states Φ_μ^j, corresponding each to a different value μ of the J_z component of the total angular momentum.

In spite of this degeneracy, we can apply the results of the perturbation theory developed above. This is because the perturbing potential V being, like H_0, invariant under rotation, we can treat the eigenvalue problem in the subspace $\mathscr{E}(j\mu)$ of the states of given angular momentum $(j\mu)$; in this subspace the ground state is not degenerate and the first-order correction to its energy is correctly given by formula (XVI.12):

$$E_c \simeq \langle \Phi_\mu^j | V | \Phi_\mu^j \rangle. \tag{XVI.22}$$

This correction is seen to be independent of μ: the Coulomb perturbation does not remove the degeneracy. This is true, in fact, in all orders: since H, like H_0, is invariant under rotation, it necessarily presents, like H_0, a rotational degeneracy.

The function Φ_μ^j depends on the coordinates and intrinsic spins

of the Z protons and $N=\mathcal{N}-Z$ neutrons. Since it is antisymmetrical with respect to the protons, the $\frac{1}{2}Z(Z-1)$ terms of V make equal contributions; we therefore have $E_c=\frac{1}{2}Z(Z-1)\,\varepsilon$, where

$$\begin{aligned}\varepsilon &= \langle \Phi_\mu{}^j | (e^2/r_{12}) | \Phi_\mu{}^j \rangle \\ &= \sum_{\text{spins}} \int |\Phi_\mu{}^j|^2 \, (e^2/r_{12}) \, \mathrm{d}\boldsymbol{r}_1 \, \mathrm{d}\boldsymbol{r}_2 \ldots \mathrm{d}\boldsymbol{r}_Z \, \mathrm{d}\boldsymbol{r}_{Z+1} \ldots \mathrm{d}\boldsymbol{r}_{\mathcal{N}}\end{aligned}$$

$\sum_{\text{spins}}$ extends over all the spin variables of the nucleons. Let us put

$$\varrho(\boldsymbol{r}_1, \boldsymbol{r}_2) = \sum_{\text{spins}} \int |\Phi_\mu{}^j|^2 \, \mathrm{d}\boldsymbol{r}_3 \ldots \mathrm{d}\boldsymbol{r}_{\mathcal{N}}.$$

$\varrho(\boldsymbol{r}_1, \boldsymbol{r}_2)$ is the probability density for simultaneously finding proton 1 at point $\boldsymbol{r}_1$, and proton 2 at point $\boldsymbol{r}_2$. We then have

$$\varepsilon = \int (e^2/r_{12}) \, \varrho(\boldsymbol{r}_1, \boldsymbol{r}_2) \, \mathrm{d}\boldsymbol{r}_1 \, \mathrm{d}\boldsymbol{r}_2.$$

To evaluate E_c we employ the very crude model where correlations between the two protons are neglected [1]),

$$\varrho(\boldsymbol{r}_1, \boldsymbol{r}_2) \simeq \varrho(\boldsymbol{r}_1) \, \varrho(\boldsymbol{r}_2) \tag{XVI.23}$$

and where the density $\varrho(\boldsymbol{r})$ is taken to be that of a uniform distribution of charge in a sphere of radius R:

$$\varrho(\boldsymbol{r}) = \begin{cases} \dfrac{3}{4\pi R^3} & \text{if} \quad r < R \\ 0 & \text{if} \quad r > R \end{cases} \tag{XVI.24}$$

The calculation is analogous to the one of § 4. Substituting the above densities into the integral I, as given by (XVI.19), we finally obtain

$$E_c \simeq \frac{3}{5} Z(Z-1) \frac{e^2}{R}. \tag{XVI.25}$$

Although formula (XVI.25) constitutes a relatively rough evaluation of E_c, it can be used to verify the charge-independence of nuclear forces (cf. § XIV.15). The isotopic spin formalism lends itself well to this purpose. According to the independence hypothesis, the unperturbed Hamiltonian is invariant under rotations in charge space, and states of the same isotopic spin multiplet have the same binding energy. The observed differences between the binding energies

---

[1]) Thus we certainly over-estimate E_c for, all things being equal, the Pauli principle requires $\varrho(\boldsymbol{r}_1, \boldsymbol{r}_2)$ to be smaller when $\boldsymbol{r}_1 = \boldsymbol{r}_2$.

of two members of a given isotopic-spin multiplet will then be equal to the difference of their Coulomb energies. If we take

$$R = 1.45 \times \mathcal{N}^{\frac{1}{3}} \times 10^{-13} \text{ cm},$$

it is possible with formula (XVI.25) to reasonably account for the difference in the ground state energies of mirror nuclei.

6. Higher-order Corrections

The second-order corrections, from equation (XVI.7[2]), are obtained by writing equations (XVI.9) and (XVI.11) for the special case $n=2$ and using the first-order corrections (XVI.12) and (XVI.14). This gives

$$\varepsilon_2 = \langle 0|V|1\rangle = \langle 0|V\frac{Q_0}{a}V|0\rangle \tag{XVI.26}$$

$$|2\rangle = \frac{Q_0}{a}(V-\varepsilon_1)|1\rangle = \left[\frac{Q_0}{a}V\frac{Q_0}{a} - \frac{Q_0}{a^2}V|0\rangle\langle 0|\right]V|0\rangle. \tag{XVI.27}$$

Corrections of higher order are similarly obtained. The formulas are lengthy to write out but simplify in the important special case when all the lower-order corrections to the unperturbed energy vanish. Thus, if

$$\varepsilon_1 = \varepsilon_2 = \ldots = \varepsilon_{n-1} = 0, \tag{XVI.28}$$

the recurrence formulas (XVI.9) and (XVI.11) give

$$\varepsilon_n = \langle 0|V\left(\frac{Q_0}{a}V\right)^{n-1}|0\rangle \tag{XVI.29}$$

$$|n\rangle = \frac{Q_0}{a}V|n-1\rangle = \left(\frac{Q_0}{a}V\right)^2|n-2\rangle = \ldots = \left(\frac{Q_0}{a}V\right)^n|0\rangle. \tag{XVI.30}$$

and condition (XVI.28) reads

$$\langle 0|V|0\rangle = \langle 0|V\frac{Q_0}{a}V|0\rangle = \ldots = \langle 0|V\left(\frac{Q_0}{a}V\right)^{n-2}|0\rangle = 0. \tag{XVI.28'}$$

It rarely happens that the calculations are continued to the nth-order when this condition is not fulfilled. The perturbation method is useful only if convergence is sufficiently rapid for the lowest-order corrections to suffice. The complexity of the calculations increases rapidly with order and soon becomes prohibitive.

Nonetheless, the general behavior of the higher-order corrections is of interest for investigating the convergence of the perturbation expansion. Writing (XVI.26) in terms of the matrix elements of V in the $\{|E^0\alpha\rangle\}$ representation we find for ε_2 the expression

$$\varepsilon_2 = \sum_{E^0 \neq E_a^0} \frac{\sum_\alpha |\langle 0|V|E^0\alpha\rangle|^2}{E_a^0 - E^0}. \tag{XVI.31}$$

The magnitude of the effect caused by the perturbation varies as the ratio of the matrix elements $\lambda\langle 0|V|E^0\alpha\rangle$ to the energy difference $|E_a^0 - E^0|$. More generally, one may say, roughly speaking, that the convergence of the perturbation series depends on the smallness of the ratio of the matrix elements $\lambda\langle E^0\alpha|V|E^{0\prime}\alpha'\rangle$ linking two eigenstates of H_0 of energy E^0 and $E^{0\prime}$ to the differences $|E_a^0 - E^0|$ and $|E_a^0 - E^{0\prime}|$ between these energies and the unperturbed energy E_a^0.

One can obtain an upper limit for ε_2 by over-estimating in absolute value each term on the right-hand side of (XVI.31). Replacing each denominator by the separation $\delta E_{\min}$ of E_a^0 from its nearest neighbor, we get

$$|\varepsilon_2| \leqslant \frac{1}{\delta E_{\min}} \sum_{E^0 \neq E_a^0} \sum_\alpha \langle 0|V|E^0\alpha\rangle \langle E^0\alpha|V|0\rangle$$

or

$$|\varepsilon_2| \leqslant \frac{1}{\delta E_{\min}} \langle 0|V\, Q_0\, V|0\rangle.$$

Replacing Q_0 by its definition we have

$$\langle 0|VQ_0V|0\rangle = \langle 0|V(1 - |0\rangle\langle 0|)V|0\rangle = \langle 0|V^2|0\rangle - \langle 0|V|0\rangle^2 = (\Delta V)^2$$

where $(\Delta V)^2$ is the root-mean-square deviation of the perturbing potential in the state $|0\rangle$. Thus we can write

$$|\varepsilon_2| \leqslant \frac{(\Delta V)^2}{\delta E_{\min}}. \tag{XVI.32}$$

7. Stark Effect for a Rigid Rotator

In general, the calculation of ε_2, which involves the infinite series of matrix elements in (XVI.31), is much more difficult than the calculation of ε_1, which involves a single matrix element of V. However,

it sometimes happens that most of the matrix elements of (XVI.31) vanish, so that the sum on the right-hand side contains only a limited number of terms.

As an example we consider the Stark level-shift for a rigid rotator. This type of problem occurs in the study of the polarizability of diatomic molecules in an electric field. The rigid rotator represents the motion of the nuclei in a diatomic molecule in the limit when the vibrational quanta are treated as infinitely large. The only degrees of freedom of the rotator are the angular variables (θ, φ) fixing its spatial orientation. Let $\boldsymbol{L}$ denote the rotator's angular momentum and $I = mr_0^2$ its moment of inertia ($m \equiv$ reduced mass, $r_0 \equiv$ mutual distance of the nuclei); its Hamiltonian is:

$$H_0 = \frac{\boldsymbol{L}^2}{2I}$$

and its eigenfunctions are the spherical harmonics $Y_l^m(\theta, \varphi)$. We denote the vector represented by Y_l^m by $|lm\rangle$. The corresponding energy eigenvalue depends only on l:

$$H_0|lm\rangle = E_l^0|lm\rangle \qquad E_l^0 = \frac{\hbar^2}{2I}\, l(l+1).$$

In the presence of a uniform electric field $\mathscr{E}$, directed along the z axis, the Hamiltonian contains the additional term

$$V = -d\mathscr{E}\cos\theta,$$

where d is the electric dipole moment of the rotator. We shall calculate the effect of this term using the perturbation theory.

In the $\{|lm\rangle\}$ representation, nearly all of the matrix elements of V vanish [1]). In order to have

$$\langle l_1 m_1|V|l_2 m_2\rangle \neq 0$$

it is necessary [cf. (C.16) and (C.17)] that

$$m_1 = m_2 \quad \text{and} \quad l_1 = l_2 \pm 1.$$

[1]) These noteworthy properties are due to the fact that V is the $q = 0$ component of an irreducible tensor operator of order 1 and that its parity is (-1).

When these conditions are fulfilled, we can deduce the matrix element from the following formula [cf. (B.90)]:

$$\langle lm|\cos\theta\,|l-1\;m\rangle = \langle l-1\;m|\cos\theta\,|lm\rangle = \left(\frac{l^2-m^2}{4l^2-1}\right)^{\frac{1}{2}}. \quad \text{(XVI.33)}$$

Except for the level $l=0$, all of the unperturbed levels are degenerate. However, H_0 and $H \equiv H_0+V$ both commute with L_z and therefore one can separately solve the eigenvalue problem for H in each of the subspaces $\mathscr{E}_m$ of a given eigenvalue m of L_z. In each of these subspaces the spectrum of H_0 is non-degenerate and the perturbation method described above is applicable.

Thus we consider the unperturbed state $|lm\rangle$ in $\mathscr{E}_m$. By the selection rules given above,

$$\langle lm|V|lm\rangle = 0.$$

Thus the first-order correction to the energy level vanishes. By these same selection rules the second-order correction, given by (XVI.31), reduces to the two terms corresponding to $l \pm 1$:

$$\varepsilon_2{}^{lm} = \frac{2I(d\mathscr{E})^2}{\hbar^2}\sum_{l'\neq l}\frac{|\langle lm|\cos\theta\,|l'm\rangle|^2}{l(l+1)-l'(l'+1)}$$

$$= \frac{2I(d\mathscr{E})^2}{\hbar^2}\left[\frac{|\langle lm|\cos\theta\,|l+1\;m\rangle|^2}{l(l+1)-(l+1)\,(l+2)} + \frac{|\langle lm|\cos\theta\,|l-1\;m\rangle|^2}{l(l+1)-l(l-1)}\right].$$

The bracket on the right-hand side is easily calculated with the aid of (XVI.33) and gives

$$\varepsilon_2{}^{lm} = \frac{(d\mathscr{E})^2}{E_l{}^0}\,\frac{l(l+1)-3m^2}{2(2l-1)\,(2l+3)}.$$

Thus, the energy spectrum is given to the second order by the formula

$$E_{lm} \simeq E_l{}^0\left[1+\left(\frac{d\mathscr{E}}{E_l{}^0}\right)^2\frac{l(l+1)-3m^2}{2(2l-1)\,(2l+3)}\right]. \quad \text{(XVI.34)}$$

The degeneracy is only partly removed since the energy E_{lm} depends only on l and m^2. States with m differing by a sign will still coincide. This residual degeneracy remains in all orders; it results from H being invariant with respect to the symmetries of the Stark effect, that is, invariant with respect to the group of reflections in planes containing Oz. We shall return to this question in § 13.

II. PERTURBATION OF A DEGENERATE LEVEL

8. Elementary Theory

Let us suppose that the eigenvalue E_a^0 of the unperturbed Hamiltonian is g_a-fold degenerate. We retain the notations of § 2 and denote by $\mathscr{E}_a^0$ the subspace corresponding to E_a^0 and by P_0 the projector onto $\mathscr{E}_a^0$.

We can now have more than one eigenvalue of H tending to E_a^0 when $\lambda \to 0$. We shall denote these eigenvalues by $E_1, E_2, \ldots, E_n$, their orders of degeneracy by $g_1, g_2, \ldots, g_n$ and their respective subspaces by $\mathscr{E}_1, \mathscr{E}_2, \ldots, \mathscr{E}_n$. We have

$$g_1 + g_2 + \ldots + g_n = g_a$$

and the space $\mathscr{E}_1 + \mathscr{E}_2 + \ldots + \mathscr{E}_n$ tends toward $\mathscr{E}_a^0$ when $\lambda \to 0$. If P is the projector onto $\mathscr{E}_1 + \mathscr{E}_2 + \ldots + \mathscr{E}_n$, it is a continuous function of λ and

$$P \xrightarrow[\lambda \to 0]{} P_0.$$

The perturbation theory determination of the eigenvalues and eigenfunctions of H is more complicated than when the unperturbed eigenvalue is non-degenerate. In section III we shall give a rigorous solution to the problem allowing its resolution to any order. Here we leave rigour aside and limiting ourselves to the lowest orders investigate what the method of § 2 gives when applied to the present problem.

Let E be one of the eigenvalues $E_1, E_2, \ldots, E_n$ and $|\psi\rangle$ one of the eigenvectors corresponding to E. In the limit when $\lambda \to 0$, $|\psi\rangle$ tends toward a certain vector $|0\rangle$, which for the moment can only be specified to the extent of saying it belongs to the space $\mathscr{E}_a^0$. We assume that E and $|\psi\rangle$ can be represented by expansions (XVI.5) and (XVI.6) with the normalization condition (XVI.4). The coefficients are related to each other by equations (XVI.7) and (XVI.8) and can be determined by recurrence.

Equation (XVI.7^0) requires that $|0\rangle$ belong to $\mathscr{E}_a^0$:

$$P_0|0\rangle = |0\rangle. \tag{XVI.35}$$

Projection of (XVI.7^1) onto $\mathscr{E}_a^0$ gives

$$P_0(V - \varepsilon_1)|0\rangle = 0 \tag{XVI.36}$$

and projection onto the complementary space gives

$$Q_0|1\rangle = \frac{Q_0}{a} V|0\rangle, \tag{XVI.37}$$

where we have put

$$Q_0 = 1 - P_0 \tag{XVI.38}$$

and defined Q_0/a in accordance with (XVI.10).

Equation (XVI.36) is an eigenvalue equation in the subspace $\mathscr{E}_a{}^0$: ε_1 is the eigenvalue of the operator $P_0 V P_0$ in $\mathscr{E}_a{}^0$ and $|0\rangle$ the corresponding eigenvector. In the $|E^0\alpha\rangle$ representation it becomes

$$\sum_{\alpha'} \langle E_a{}^0\alpha|V|E_a{}^0\alpha'\rangle \langle E_a{}^0\alpha'|0\rangle = \varepsilon_1 \langle E_a{}^0\alpha|0\rangle.$$

Thus the first-order correction ε_1 is obtained by diagonalizing the $g_a \times g_a$ matrix whose elements are given by

$$V_{\alpha\alpha'} \equiv \langle E_a{}^0\alpha|V|E_a{}^0\alpha'\rangle.$$

The possible values for ε_1 are the eigenvalues of this matrix. If there are g_a different eigenvalues they are all non-degenerate and the perturbation has completely removed the degeneracy. If there are less than g_a different eigenvalues, some of them will be degenerate and the degeneracy will only have been partly removed.

If the first-order correction ε_1 is a non-degenerate eigenvalue, the corresponding eigenstate $|0\rangle$ is completely determined to the zeroth order; it is defined to within a constant by (XVI.7^0) and (XVI.7^1). The projection $Q_0|1\rangle$ of the first-order correction to $|\psi\rangle$ onto the complement of $\mathscr{E}_a{}^0$ is given by (XVI.37). Its projection onto $\mathscr{E}_a{}^0$ remains undetermined except for condition (XVI.4). If $\mathscr{E}_1$ is g_1-fold degenerate, equations (XVI.7^0) and (XVI.7^1) show only that $|0\rangle$ belongs to the corresponding g_1-dimensional subspace; for a more precise determination of $|0\rangle$ one must proceed to higher orders.

Having chosen one of the values ε_1, we obtain the second-order correction ε_2 by projecting (XVI.7^2) onto the subspace of ε_1. This subspace, which we denote $\mathscr{E}_a{}^{(1)}$, is contained in $\mathscr{E}_a{}^0$; we denote the corresponding projector by $P^{(1)}$ and the projector onto its complement in $\mathscr{E}_a{}^0$ by P':

$$P_0 = P^{(1)} + P' \qquad P^{(1)} + P' + Q_0 = 1.$$

We then have

$$\begin{aligned} P^{(1)}H_0 &= P^{(1)}P_0H_0 \\ &= E_a{}^0P^{(1)} \\ P^{(1)}V &= P^{(1)}V(P^{(1)}+P'+Q_0) \\ &= \varepsilon_1P^{(1)}+P^{(1)}VQ_0. \end{aligned}$$

The projection of equation (XVI.7[2]), gives

$$P^{(1)}VQ_0|1\rangle-\varepsilon_2P^{(1)}|0\rangle = 0$$

whence, with the aid of (XVI.37),

$$P^{(1)}\left[\left(V\frac{Q_0}{a}V\right)-\varepsilon_2\right]|0\rangle = 0. \qquad \text{(XVI.39)}$$

Equation (XVI.39) is for the second-order corrections the analogue of (XVI.36) for the first. In the same way as ε_1 was the eigenvalue of P_0VP_0 in $\mathscr{E}_a{}^0$, ε_2 is the eigenvalue of $P^{(1)}V(Q_0/a)P^{(1)}$ in $\mathscr{E}_a{}^{(1)}$. The corresponding eigenvector is $|0\rangle$.

If ε_1 is a non-degenerate eigenvalue ($g_1=1$), then $|0\rangle$ is determined by the equations of lower order and we have

$$\varepsilon_2 = \langle 0|V\frac{Q_0}{a}V|0\rangle$$

as in the non-degenerate case. If on the contrary $g_1>1$, the calculation of ε_2 involves finding the eigenvalues of a $g_1\times g_1$ matrix. If these are found to be all different, the degeneracy is completely removed in the second order; if not, we must, if need be, proceed to higher orders.

It sometimes happens that a degeneracy remains in all orders. We have already had examples in the Coulomb energy of nuclei (§ 5) and in the Stark effect for a rigid rotator (§ 7). Examination of the symmetries of H_0 and H usually enables one to predict to what extent the degeneracy of an unperturbed level is likely to be removed by the perturbation λV. A systematic discussion of this point is given in § 13; we first illustrate the method with some examples of first-order calculations taken from atomic physics.

9. Atomic Levels in the Absence of Spin-Orbit Forces

In Chapter XIV, we studied the energy levels of a Z-electron atom

in the central field approximation (§ XIV.12). In this approximation the Hamiltonian H is replaced by

$$H_C = \sum_{i=1}^{z} \left[\frac{\boldsymbol{p}_i^2}{2m} + V_C(r_i) \right].$$

This takes into account only the average value of the electrostatic repulsion of the electrons. Let us denote by V_1 the difference between the exact Coulomb interaction and the potential of H_C

$$V_1 = \sum_{i<j} \frac{e^2}{|\boldsymbol{r}_i - \boldsymbol{r}_j|} - \sum_i \left[V_C(r_i) + \frac{Ze^2}{r_i} \right]. \tag{XVI.40}$$

If we neglect the terms in H depending on electron spin, we can write

$$H \simeq H_C + V_1 \tag{XVI.41}$$

and the spectrum of H can be deduced from that of H_C by treating V_1 as a perturbation. The modifications required by the inclusion of spin-dependent force will be considered in § 10.

As a general rule, the eigenvalues of H_C are all strongly degenerate. The perturbation V_1 at least partly removes this degeneracy. The levels of H in the neighbourhood of a given eigenvalue of H_C are obtained by diagonalizing V_1 in the subspace of that eigenvalue. In particular, the ground and first excited states of H are obtained by diagonalizing V_1 in the subspace $\mathscr{E}_0$ of the lowest eigenvalue E_0 of H_C.

The symmetries of H considerably simplify the problem of diagonalization. Since H (like H_C) is independent of the spins, it is invariant not only under overall rotations, but also under rotations of the orbital variables alone and the spin variables alone; H commutes not only with the total angular momentum $\boldsymbol{J}$ but also with the total orbital angular momentum $\boldsymbol{L}$ and the total spin $\boldsymbol{S}$ [1]). Since H_C has the same symmetries as H, the operator $P_0 V_1 P_0$ will also commute with $\boldsymbol{L}$ and $\boldsymbol{S}$; its eigenstates in $\mathscr{E}_0$ can be labelled by the eigenvalues of $\boldsymbol{L}^2$, $\boldsymbol{S}^2$, L_z and S_z and its eigenvalues depend only on L and S and each have a degeneracy of order $(2L+1)(2S+1)$.

Denote the vectors of a standard basis $\{\boldsymbol{L}^2 L_z, \boldsymbol{S}^2 S_z\}$ in $\mathscr{E}_0$ by $|\gamma L S M_L M_S\rangle$; the quantum number γ distinguishes between vectors having the same orbital angular momentum and spin. In the corre-

[1]) Moreover, H is invariant under permutation of the orbital variables alone, but by the theorem of § D.18 this is equivalent, because of the Pauli principle, to the rotational invariance of the spins.

sponding representation, the matrix of $P_0 V_1 P_0$ takes the very simple form:

$$\langle \gamma L S M_L M_S | V_1 | \gamma' L' S' M_L' M_S' \rangle = \delta_{LL'} \, \delta_{SS'} \, \delta_{M_L M_L'} \, \delta_{M_S M_S'} \, \mathscr{V}_{\gamma\gamma'}^{(LS)}.$$

To effect the diagonalization we need only to diagonalize the matrices $\mathscr{V}^{(LS)}$ corresponding to each pair of quantum numbers (LS).

As an example we consider the carbon atom. Its ground state configuration is $1s^2\ 2s^2\ 2p^2$, i.e., two complete shells, 1s and 2s, and an incomplete 2p shell with 2 electrons. It is $\binom{6}{2} = 15$ fold degenerate. For the purpose of finding the possible values for the pair (LS) and their degeneracy, i.e., the number of corresponding series of $(2L+1)$ $(2S+1)$ vectors, one may ignore the closed shells and consider only the two 2p electrons (Problem XVI.3).

With neglect of the Pauli principle, the different spectral terms that can be formed with two 2p electrons are

$$^3S \quad ^3P \quad ^3D \quad ^1S \quad ^1P \quad ^1D.$$

As the triplet and singlet spin states are respectively symmetrical and antisymmetrical in the exchange of the spins, and as the states S and D on the one hand, and P on the other, are respectively symmetrical and antisymmetrical in the interchange of the orbital variables, the terms of this list that satisfy the Pauli principle are

$$^3P \quad ^1S \quad ^1D,$$

in all $9+1+5=15$ linearly independent antisymmetrical states, as predicted. In the $\{LSM_L M_S\}$ basis (in this particular case the quantum number γ is superfluous), the perturbation V_1 is necessarily diagonal. It has, in all, three different eigenvalues, $\mathscr{V}(^3P)$, $\mathscr{V}(^1S)$ and $\mathscr{V}(^1D)$, respectively 9-, 1- and 5-fold degenerate; these are not difficult to calculate numerically when the individual-state wave functions are known. The calculation shows that the level ^3P is distinctly lower than the other two (Fig. XVI.1) [1].

---

[1] In general, the levels of a given configuration are in order of decreasing total spin (*Hund's rule*). This is because the second term in (XVI.40) gives the same contribution to all levels so that the order of succession of the levels depends only on the magnitude of the repulsive term $\sum e^2/|r_i - r_j|$. This term decreases with the mutual distance of the electrons and therefore as the orbital part of the wave function becomes more antisymmetrical. The spin part of

10. Spin-Orbit Forces. LS and jj Coupling

H is only approximately given by (XVI.41). It also contains spin-dependent terms which we shall denote by V_2:

$$H = H_C + V_1 + V_2. \tag{XVI.42}$$

In a first approximation, each electron moves independently of the others in the potential $V_C(r)$ and its spin $\boldsymbol{s}$ is coupled to its orbital angular momentum in accordance with (XIII.95); thus to a very good approximation

$$V_2 \simeq \sum_{i=1}^{Z} (\boldsymbol{l}_i \cdot \boldsymbol{s}_i)\, g(r_i) \tag{XVI.43}$$

where

$$g(r) = \frac{1}{2m^2 c^2} \frac{1}{r} \frac{\mathrm{d}V_C}{\mathrm{d}r}. \tag{XVI.44}$$

To correctly take this effect into account we must replace V in the treatment of § 9 by $V_1 + V_2$. However, $V_1 + V_2$ is less symmetrical than V_1: $V_1 + V_2$ commutes with $\boldsymbol{J}$, but neither with $\boldsymbol{L}$ nor with $\boldsymbol{S}$. The problem of diagonalizing the perturbation in the subspace of the unperturbed eigenvalue is therefore greatly complicated by adding the spin-orbit term. It becomes relatively simple only when one of V_1 or V_2 is much smaller than the other.

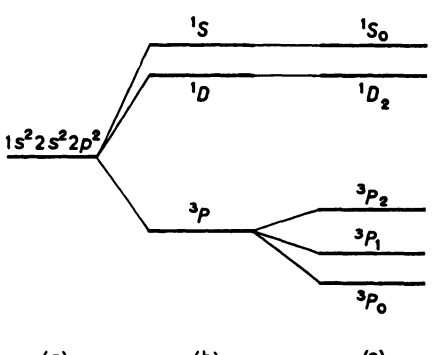


Fig. XVI.1. The levels of the ground state of the carbon atom; (a) in the central field approximation ($V_1 = V_2 = 0$); (b) neglecting spin-orbit coupling ($V_2 = 0$); (c) including spin-orbit coupling.

---

the wave function is more symmetrical for larger S (§ D.18) and therefore the orbital part is more antisymmetrical for larger S.

If $V_1 \gg V_2$, one may neglect V_2 in a first approximation. Each configuration will then give a series of levels each corresponding to a well-defined pair (LS) and each having (cf. § 9) a degeneracy of order $(2L+1)(2S+)1$; the corresponding eigenvectors are particular linear combinations of Slater determinants formed of the individual-states of the configuration. They are eigenvectors of $\mathbf{L}^2$, $\mathbf{S}^2$, L_z and S_z. V_2 is then treated as a small perturbation in the subspace corresponding to each of these levels. To each possible value of $J(J=L+S, L+S-1, \ldots, |L-S|)$ there corresponds an eigenvalue of V_2 with a degeneracy of order $(2J+1)$. The corresponding eigenvectors $|\gamma LSJM\rangle$ are eigenvectors of $\mathbf{L}^2$, $\mathbf{S}^2$, $\mathbf{J}^2$ and J_z. This method of coupling the Slater determinants of the configuration to form the eigenvectors of the total angular momentum is called *Russel–Saunders coupling* or *LS coupling*.

If $V_2 \gg V_1$, one may neglect V_1 in a first approximation. The Hamiltonian then becomes $H_C + V_2$ which is the Hamiltonian for independent particles moving in the potential $V_C + (\boldsymbol{l}\cdot\boldsymbol{s})g(r)$. Let the total angular momentum of each individual particle be $\boldsymbol{j}=\boldsymbol{l}+\boldsymbol{s}$. The spin-orbit coupling term $(\boldsymbol{l}\cdot\boldsymbol{s})g(r)$ partly removes the degeneracy of the individual states of non-zero orbital angular momentum by splitting each of them into the two levels $j = l \pm \frac{1}{2}$. The corresponding eigenvectors are specified by the quantum numbers $(nljm)$. The treatment of $H_C + V_2$ is otherwise in all ways analogous to the treatment of H_C itself (§ XIV.12); each configuration of H_C gives rise to a certain number of configurations of $H_C + V_2$. Once these are determined, V_1 is treated in each of them as a perturbation. To each possible eigenvalue of J there corresponds one (or several) eigenvalue(s) with a degeneracy of order $(2J+1)$. The corresponding eigenvectors are eigenvectors of the $\boldsymbol{j}_i^2$, and of $\mathbf{J}^2$ and J_z. This method of coupling the Slater determinants of the ground-state configuration to form the eigenvectors of the total angular momentum is called *jj coupling*.

The following table summarizes these two methods.

LS Coupling		*jj* Coupling	
(V_1)	$\mathbf{L} = \sum_i \boldsymbol{l}_i \qquad \mathbf{S} = \sum_i \boldsymbol{s}_i$	(V_2)	$\boldsymbol{j}_i = \boldsymbol{l}_i + \boldsymbol{s}_i$
(V_2)	$\mathbf{J} = \mathbf{L} + \mathbf{S}$	(V_1)	$\mathbf{J} = \sum_i \boldsymbol{j}_i$

The relative importance of V_2 increases rapidly with Z [1]). In the light and medium atoms $V_1 \gg V_2$ and LS coupling is an excellent approximation; in heavier atoms (starting from Pb say), V_1 and V_2 are of the same order of magnitude and the level structure of the ground state configuration is intermediate between those given by LS and jj coupling.

11. The Atom in LS Coupling. Splitting Due to Spin-Orbital Coupling

Let us consider the effect of the spin-orbit coupling in the LS coupling case ($V_1 \gg V_2$). We treat V_2 as a perturbation of the levels of $H_C + V_1$.

The spectrum of the unperturbed Hamiltonian has been described in § 9. Each level corresponds to a specified value of L and S and has a degeneracy of order $(2L+1)(2S+1)$; the subspace of the corresponding eigenvectors is spanned by a set of vectors of well-defined orbital angular momentum and spin which we shall denote by $|\alpha LSM_L M_S\rangle$. The quantum number α indicates the configuration of H_C to which the level belongs, and also distinguishes between levels of the same configuration which have the same values for L and S.

The perturbation energy is obtained by diagonalizing V_2 in the subspace $\mathscr{E}(\alpha LS)$ of each of the unperturbed levels. It will now be shown that the matrix elements of V_2 in each subspace $\mathscr{E}(\alpha LS)$ are the same as those of the operator $A(\mathbf{L}\cdot\mathbf{S})$, where A is a constant characteristic of the unperturbed level (αLS):

$$\langle \alpha LSM_L M_S | V_2 | \alpha LSM_L' M_S' \rangle = A \langle \alpha LSM_L M_S | (\mathbf{L}\cdot\mathbf{S}) | \alpha LSM_L' M_S' \rangle. \tag{XVI.45}$$

Proof: As the basis vectors are antisymmetrical, the contributions to the matrix element from the Z terms of V_2 [eq. (XVI.43)] are equal. Hence we need only to treat one of them, that of the term $(\boldsymbol{l}_1\cdot\boldsymbol{s}_1)g(r_1)$ say, (no confusion being possible, we drop the index 1 in what follows). From an arbitrary component l_m of $\boldsymbol{l}$ and an arbi-

---

[1]) V_1 is a fluctuation term whose effect on the energy of each electron increases roughly as $\sqrt{Z}$. The spin-orbit coupling energy of each electron increases roughly as Z^2. It can be evaluated using the Thomas–Fermi model; the coupling is proportional to the average value of $(1/r)(\mathrm{d}V_C/\mathrm{d}r)$ which with the notation of § XIV.13 is approximately given by

$$\frac{1}{r}\frac{\mathrm{d}V_C}{\mathrm{d}r} \simeq \frac{Z^2e^2}{b^3}\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\chi}{x}\right).$$

trary component s_μ of $\boldsymbol{s}$, one can form the operator $g(r)l_m s_\mu$ which is simultaneously component of a vector operator irreducible with respect to rotations of the orbital variables alone, and component of a vector operator irreducible with respect to rotations of the spin variables alone. The operator $L_m S_\mu$ has the same property. Applying the Wigner–Eckart theorem, we easily deduce (cf. Problem XIII.1)

$$\langle \alpha LSM_L M_S | g l_m s_\mu | \alpha LSM_L' M_S' \rangle = a \langle \alpha LSM_L M_S | L_m S_\mu | \alpha LSM_L' M_S \rangle,$$

where a is a constant independent of the magnetic quantum numbers M_L, M_L', M_S, M_S', m, μ. It follows that

$$\langle \alpha LSM_L M_S | g(r)(\boldsymbol{l}\cdot\boldsymbol{s}) | \alpha LSM_L' M_S' \rangle = a \langle \alpha LSM_L M_S | (\mathbf{L}\cdot\mathbf{S}) | \alpha LSM_L' M_S' \rangle.$$

One obtains (XVI.45) by multiplying each term by Z and putting $A = Za$.

In the subspace $\mathscr{E}(\alpha LS)$, V_2 is not diagonal in the $\{\alpha LSM_L M_S\}$ representation but being invariant under rotation it is obviously diagonal in the $\{\alpha LSJM\}$ representation whose basis vectors are the eigenvectors of $\boldsymbol{J}^2$ and J_z in $\mathscr{E}(\alpha LS)$. Since

$$\boldsymbol{J}^2 = \mathbf{L}^2 + \mathbf{S}^2 + 2\mathbf{L}\cdot\mathbf{S}$$

one has, with the aid of (XVI.45)

$$\langle \alpha LSJM | V_2 | \alpha LSJM \rangle = \tfrac{1}{2} A \langle \alpha LSJM | (\boldsymbol{J}^2 - \mathbf{L}^2 - \mathbf{S}^2) | \alpha LSJM \rangle = \tfrac{1}{2} A \hbar^2 [J(J+1) - L(L+1) - S(S+1)]. \quad \text{(XVI.46)}$$

Thus the unperturbed level (αLS) splits into as many levels as there are possible values for J ($J = L+S, \ldots, |L-S|$). The degeneracy of each of these levels is of order $(2J+1)$ and the perturbation energy is given by formula (XVI.46).

In Figure XVI.1, we have given the level-scheme for the ground-state configuration of carbon in LS coupling. Only the ^{3}P state is affected by the spin-orbit energy term; it gets split into three levels: $^3P_0$, $^3P_1$ and $^3P_2$. Since in this particular case $A > 0$, these levels occur in order of increasing J and the lowest of them is a $^3P_0$ state.

12. The Zeeman and Paschen–Back Effects

In the last three paragraphs we have examined the structure of atomic levels in the absence of an external field. Let us now consider an atom placed in a constant magnetic field $\mathscr{H}$. The Hamiltonian

is obtained from the Hamiltonian of the atom without external field, H_0, by adding the term [cf. eq. (XIII.96)]

$$W = -\frac{e}{2mc}[\mathscr{H}\cdot(\boldsymbol{L}+2\boldsymbol{S})] \tag{XVI.47}$$

(as well as a "diamagnetic" term in $\mathscr{H}^2$ which we neglect).

Zeeman Effect

For sufficiently small $\mathscr{H}$, W may be treated as a perturbation. Since H_0 is invariant under rotation, each level of H_0 corresponds to a definite value J of the total angular momentum. We shall suppose $J \neq 0$ [1]), in which case the level has a degeneracy of order $(2J+1)$. This degeneracy is removed by the perturbation W.

Let E_0 be one of the unperturbed levels, J its angular momentum and $|E_0JM\rangle$ the eigenvectors of $\boldsymbol{J}^2$ and J_z that span the subspace of the eigenvalue E_0. The level-shifts caused by the magnetic field are the eigenvalues of the matrix

$$\langle E_0JM|W|E_0JM'\rangle.$$

This matrix is automatically diagonalized if we take the z axis parallel to $\mathscr{H}$ since W then commutes with J_z.

Also, according to the Wigner–Eckart theorem, the matrix elements of the vector operators $\boldsymbol{L}+2\boldsymbol{S}$ and $\boldsymbol{J}$ in the subspace of E_0 are proportional:

$$\langle E_0JM|(\boldsymbol{L}+2\boldsymbol{S})|E_0JM'\rangle = g\,\langle E_0JM|\boldsymbol{J}|E_0JM'\rangle. \tag{XVI.48}$$

The proportionality factor, g, the so-called Landé factor, is equal to the ratio of the reduced matrix elements of these operators; it is a magnitude characteristic of the level E_0 considered. In particular,

$$\langle E_0JM|(L_z+2S_z)|E_0JM'\rangle = gM\hbar\delta_{MM'}.$$

With the above-mentioned choice of axes,

$$W = -\frac{e\mathscr{H}}{2mc}(L_z+2S_z)$$

---

[1]) If $J = 0$ (diamagnetic atom) the level shift vanishes to the first order in $\mathscr{H}$. In the calculation of this shift to the second order in $\mathscr{H}$ the contribution of the "diamagnetic term" cannot be neglected.

whence, if we denote the Bohr magneton by μ_B [eq. (XIII.74)]

$$\langle E_0 JM | W | E_0 JM' \rangle = - Mg\mu_B \mathscr{H} \delta_{MM'}.$$

Thus the perturbation completely removes the degeneracy and the $(2J+1)$ levels obtained are given by the Zeeman formula

$$E = E_0 - Mg\mu_B \mathscr{H} \qquad (M = -J, -J+1, \ldots, +J) \quad \text{(XVI.49)}$$

Calculation of the Landé factor in LS coupling

The Landé factor g is defined by equation (XVI.48). We suppose that the LS coupling scheme is valid for the atom considered and therefore, following the notation of § 11, the atomic levels in the absence of the external field will be characterized by the quantum numbers αLSJ.

To determine g we calculate in two different ways the average value

$$\langle \boldsymbol{J} \cdot (\boldsymbol{L} + 2\boldsymbol{S}) \rangle \equiv \langle \alpha LSJM | (\boldsymbol{J} \cdot (\boldsymbol{L} + 2\boldsymbol{S})) | \alpha LSJM \rangle.$$

(*a*) With the aid of the identity

$$\boldsymbol{J} \cdot (\boldsymbol{L} + 2\boldsymbol{S}) = \boldsymbol{J}^2 + \boldsymbol{S}^2 + \tfrac{1}{2}(\boldsymbol{J}^2 - \boldsymbol{L}^2 - \boldsymbol{S}^2)$$

we have

$$\langle \boldsymbol{J} \cdot (\boldsymbol{L} + 2\boldsymbol{S}) \rangle = \tfrac{1}{2}[3J(J+1) + S(S+1) - L(L+1)]\hbar^2.$$

(*b*) Directly multiplying the matrices in the $\{\alpha LSJM\}$ representation, and applying (XVI.48), we find (cf. Problem XIII.19)

$$\begin{aligned} \langle \boldsymbol{J} \cdot (\boldsymbol{L} + 2\boldsymbol{S}) \rangle &= \sum_{i=xyz} \sum_{M'} \langle \alpha LSJM | J_i | \alpha LSJM' \rangle \\ &\qquad \times \langle \alpha LSJM' | (L_i + 2S_i) | \alpha LSJM \rangle \\ &= g \langle \alpha LSJM | \boldsymbol{J}^2 | \alpha LSJM \rangle \\ &= g\, J(J+1)\, \hbar^2. \end{aligned}$$

Equating these two results gives

$$g = 1 + \frac{J(J+1) + S(S+1) - L(L+1)}{2J(J+1)}. \qquad \text{(XVI.50)}$$

Paschen–Back Effect

The above theory of the Zeeman Effect is justified only if the splitting of the levels by the perturbation W is small compared with the distance between the unperturbed levels. When LS coupling is

valid this supposes W to be small compared with the spin-orbit coupling V_2, or more precisely [cf. eq. (XVI.46) and (XVI.49)], that

$$|g|\mu_B\mathscr{H} \ll A\hbar^2.$$

Let us suppose that, on the contrary, the magnetic field is so strong that we have

$$A\hbar^2 \ll \mu_B\mathscr{H}.$$

In this case the spin-orbit coupling has a negligible effect as compared with that of the magnetic coupling. The latter can therefore be treated as a perturbation of the levels of the Hamiltonian

$$H_C + V_1.$$

Following the notations of § 11, each of these unperturbed levels is labelled by the quantum numbers (αLS) and has a degeneracy of order $(2S+1)\,(2L+1)$. The perturbation partly removes this degeneracy. The effect is known as the Paschen–Back effect. One obtains the displacements relative to an unperturbed level $E_{\alpha LS}$ by diagonalizing W in the corresponding subspace. Since W commutes with L_z and S_z, the perturbation matrix is diagonal in the $\{\alpha LSM_LM_S\}$ representation:

$$\langle \alpha LSM_LM_S|W|\alpha LSM_L'M_S'\rangle = -\mu_B\mathscr{H}(M_L+2M_S)\delta_{M_LM_L'}\delta_{M_SM_S'}.$$

Hence the level shifts are given by the formula

$$E = E_{\alpha LS} - \mu_B\mathscr{H}(M_L+2M_S)$$
$$(M_L = -L, \ldots, +L; \qquad M_S = -S, \ldots, +S).$$

If $L \neq 0$ and $S \neq 0$ certain of the levels obtained are still degenerate. The introduction of the small spin-orbit coupling term at least partially removes these residual degeneracies.

13. Symmetry of H and Removal of Degeneracy [1])

The preceding examples (§ 5, 7, 9, 10, 11, 12) show the importance of symmetry considerations in perturbation calculations of degenerate levels. The existence of degenerate eigenvalues can almost always be associated with the symmetries of the Hamiltonian; knowledge of

[1]) This paragraph, the only one of this chapter calling upon the results of Chapter XV and the few notions of group theory used there, may be omitted in a first reading.

the invariance groups of H_0 and H allows one to predict within what limits the degeneracy of a level E_a^0 can be removed by the perturbation, and considerably facilitates the calculation of the perturbation energies.

In all the examples previously studied, the unperturbed Hamiltonian H_0 was invariant with respect to a certain group G_0, and the perturbation λV invariant with respect to a certain subgroup G of G_0. Thus, for the case of the Stark effect for a rigid rotator (§ 7), G_0 was the rotation reflection group and G the group of reflections with respect to planes containing the z axis (cf. Ch. XV, § 14).

We return to the notation of § (XV. 11). With the operators of G we can form the observables J and M with eigenvalues labelled by the quantum numbers j and μ respectively. To each value of j there corresponds a definite irreducible representation of the group. Denote the degree of this representation by d_j; for a given j there are d_j possible values for μ distinguishing between the d_j basis vectors of the representation. To each pair of eigenvalues $(j\mu)$ there corresponds a certain subspace $\mathscr{E}(j\mu)$ of the state-vector space $\mathscr{E}$. Since H and H_0 are invariant with respect to G, and consequently commute with J and M, one can separately solve the corresponding eigenvalue problems in each of the subspaces $\mathscr{E}(j\mu)$. We obtain the same spectra with the same degeneracies in the d_j subspaces corresponding to a given value of j.

Suppose that E_a^0 is an eigenvalue of H_0 in $\mathscr{E}(j\mu)$ with a degeneracy of order p_j. According to the above,

$$g_a = \sum_j p_j d_j. \tag{XVI.51}$$

In the subspace $\mathscr{E}(j\mu)$, the introduction of the perturbation may more or less completely remove the degeneracy in the unperturbed energy (supposing $p_j > 1$); if H has no symmetries other than those of the group G, the degeneracy in question is in general completely removed and one obtains p_j distinct levels. *Hence, in the total space $\mathscr{E}$, the introduction of the perturbation splits the unperturbed level into at most $\sum_j p_j$ distinct levels; each of these levels corresponds to a specified value of j and has a degeneracy of order d_j when the degeneracy is completely removed in $\mathscr{E}(j\mu)$, and otherwise has a degeneracy which is a multiple of d_j.*

These results are based only on the fact that H has the symmetry

G and is a continuous function of λ tending to H_0 when $\lambda \to 0$. *They are exact in all orders of the perturbation calculation.*

If the subspace $\mathscr{E}_a^0$ is irreducible with respect to the group G, the sum (XVI.51) will contain just one term [1]), and the degeneracy of the level E_a^0 cannot be removed in any order. This is what generally occurs when H_0 has no symmetries other than those of H ($G_0 = G$). We have already had an example of this in the calculation of the Coulomb energy of the nucleus (§ 5).

In the case of § 7 on the other hand, G is definitely a subgroup of G_0. The subspace $\mathscr{E}_l^0$ of the eigenvalue E_l^0 is $(2l+1)$-fold degenerate. It is irreducible with respect to G_0 but reducible (if $l \neq 0$) with respect to G. In this particular case (cf. § XV.14), J is the operator l_z^2 with eigenvalues m^2 which we can label with the quantum number $|m|$ ($|m| = 0, 1, 2, \ldots$); the number of dimensions of the corresponding irreducible subspace is

$$d_{|m|} = \begin{cases} 2 & \text{if} \quad |m| \neq 0 \\ 1 & \text{if} \quad |m| = 0. \end{cases}$$

For this case (XVI.51) becomes:

$$g_l \equiv 2l + 1 = \sum_{|m|=0}^{l} d_{|m|}.$$

Thus the perturbation splits the level E_l^0 into at most $l+1$ distinct levels, one of which, $|m| = 0$, is non-degenerate, while all the rest are doubly degenerate. This is just what was found in § 7. It should be stressed that these symmetry considerations give only an upper limit for the removal of the degeneracy; this limit will be attained if the perturbation calculation is continued to sufficiently high orders. In the case of § 7 it was attained in the second order.

14. Quasi-Degeneracy

When two levels E_a^0, E_b^0 are so close together that corrections introduced by the perturbation V are large with respect to $|E_a^0 - E_b^0|$, the preceding methods lose their validity. One can nonetheless apply perturbation methods to this "quasi-degenerate" case if one appropriately modifies the definition of the unperturbed Hamiltonian and the perturbation.

---

[1]) $\mathscr{E}_a^0$ is an invariant subspace with respect to G and defines a certain representation of degree g_a of this group, which we denote $\mathbf{G}_a$. From (XVI.51), its decomposition in irreducible parts is: $\mathbf{G}_a \simeq \sum_j p_j \mathbf{G}^{(j)}$.

Let P_i^0 denote the projector onto the subspace of the eigenvalue E_i^0 of H_0. One has

$$H_0 = \sum_i E_i^0 P_i^0.$$

The modification mentioned above consists essentially in replacing the terms $E_a^0 P_a^0$ and $E_b^0 P_b^0$ in this expression by $E_\alpha^0(P_a^0 + P_b^0)$, where E_α^0 is a quantity intermediate between E_a^0 and E_b^0. One thereby obtains a new unperturbed Hamiltonian whose eigenvalue E_α^0 has a degeneracy of order $g_a + g_b$. The correction to E_α^0 due the perturbation

$$V + (E_a^0 - E_\alpha^0)P_a^0 + (E_b^0 - E_\alpha^0)P_b^0$$

can then be calculated.

In treating the Paschen–Back effect (§ 12) the method applied was essentially the one described here. There the field $\mathscr{H}$ was relatively strong so that the level shifts of a given LS term were no longer small as compared with their mutual separation. We then took $H_C + V_1$ as the unperturbed Hamiltonian rather than $H_C + V_1 + V_2$, which amounts to replacing the group of neighboring levels $E_{\alpha LSJ}$ ($J = |L-S|, \ldots, L+S$) by the single level $E_{\alpha LS}$. Normally one would then make a perturbation calculation of the corrections to $E_{\alpha LS}$ due to the perturbation $V_2 + W$. In the pure Paschen–Back effect, $\mathscr{H}$ is so strong that in a first approximation one may neglect V_2 in this calculation ($V_2 \ll W$), which is what was done in § 12. When V_2 and W are of the same order of magnitude this approximation is no longer valid and the perturbation calculation is considerably complicated (Problems XVI.7 and XVI.8).

III. EXPLICIT FORMS FOR THE PERTURBATION EXPANSION IN ALL ORDERS [1]

15. The Hamiltonian H and its Resolvent $G(z)$ [2]

In this section we briefly set forth a method for explicitly giving the perturbation expansion in all orders; the method, due to Kato [3],

[1] This section may be omitted in a first reading.

[2] Some authors use the term Green's function instead of resolvent.

[3] T. Kato, *Prog. Theor. Phys.* 4, 154 (1949). Kato is mainly interested in the conditions of convergence of the perturbation method, and his method lends itself particularly well to a discussion of this question. In most cases encountered, the perturbation expansion is an asymptotic expansion. In this section we leave aside these mathematical aspects of Kato's work.

consists in using the expansion of the resolvent $G(z)$ of the Hamiltonian H into a power series in the perturbation. In this paragraph we define $G(z)$ and describe some of its properties.

By definition, if H is an observable, its *resolvent* is the function

$$G(z) \equiv \frac{1}{z-H} \tag{XVI.52}$$

of the complex variable z [1]).

Figure XVI.2 shows the complex z-plane with the spectrum of H marked on the real axis; it has in general a discrete part and a continuous part, the latter being to the right of the discrete part and extending to infinity. $G(z)$ is an analytic function whose singularities constitute the spectrum of H.

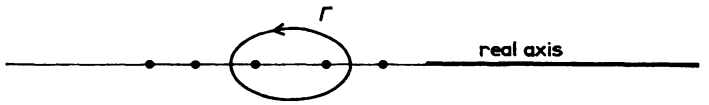


Fig. XVI.2. The complex z plane and the definition of the contour Γ. The spectrum of H is indicated in heavy print on the real axis.

For simplicity we shall suppose the spectrum of H to be entirely discrete. We denote the eigenvalues by $E_0, E_1, \ldots, E_i, \ldots$ and the projector onto the subspace of E_i by P_i

$$HP_i = E_i P_i. \tag{XVI.53}$$

---

[1]) By definition, an operator Q of Hilbert space is bounded if there exists a constant M such that

$$\frac{\langle u|Q^\dagger Q|u\rangle}{\langle u|u\rangle} \leqslant M$$

for any $|u\rangle$; the lowest value of M is by definition the upper bound or norm of the said operator and is denoted $\|Q\|$. In all rigor, the different operations of algebra and analysis, and in particular the notions of convergence of series, of differentiation, etc. can be extended without restriction only to the bounded operators of Hilbert space (cf. M. H. Stone, *Linear transformations in Hilbert space* (Amer. Math. Soc., New York, 1932).

From the mathematical point of view, the great interest of $G(z)$ is that it is *bounded* in the whole of the complex plane except for the eigenvalues of H: it is an analytic function of z whose singular points constitute the spectrum of H. If $\Delta(z)$ is the square of the distance from z to the closest eigenvalue of H, we have

$$\|G(z)\| = 1/\Delta(z).$$

One then has the orthogonality and closure relations:

$$P_i P_j = \delta_{ij} P_i \tag{XVI.54}$$

$$\sum_i P_i = 1. \tag{XVI.55}$$

From definition (XVI.52)

$$G(z)\, P_i = \frac{P_i}{z - E_i}$$

and therefore

$$G(z) = \sum_i \frac{P_i}{z - E_i}. \tag{XVI.56}$$

Each discrete eigenvalue E_i of H is a simple pole of $G(z)$ whose residue is the projector P_i, i.e.

$$P_i = \frac{1}{2\pi i} \oint_{\Gamma_i} G(z)\, dz, \tag{XVI.57}$$

where Γ_i is a closed contour in the complex plane enclosing E_i and excluding all other singularities of $G(z)$. More generally, if Γ is a closed contour in the complex plane not going through any of the eigenvalues of H (Fig. XVI.2) and if P_Γ is the sum of the projectors P_i relating to eigenvalues inside this contour, then

$$P_\Gamma = \frac{1}{2\pi i} \oint_\Gamma G(z)\, dz. \tag{XVI.58}$$

Multiplying (XVI.58) by H and taking into account the identity

$$(z - H)G \equiv G(z - H) \equiv 1$$

we obtain the important formula

$$HP_\Gamma = \frac{1}{2\pi i} \oint_\Gamma zG(z)\, dz. \tag{XVI.59}$$

16. Expansion of $G(z)$, P and HP into Power Series in λV

We now consider the perturbation problem itself.

The resolvents of H and of H_0 are respectively:

$$G \equiv \frac{1}{z - H_0 - \lambda V} \qquad G_0 \equiv \frac{1}{z - H_0}. \tag{XVI.60}$$

We note that

$$\frac{1}{z-H_0-\lambda V} = \frac{1}{z-H_0}\,[(z-H_0-\lambda V)+\lambda V]\,\frac{1}{z-H_0-\lambda V}$$

$$= \frac{1}{z-H_0} + \frac{1}{z-H_0}\,\lambda V\,\frac{1}{z-H_0-\lambda V}.$$

G is therefore the solution of the integral equation:

$$G = G_0(1+\lambda V G). \tag{XVI.61}$$

Iteration of this equation gives G as an expansion in powers of the perturbation [1]):

$$G = \sum_{n=0}^{\infty} \lambda^n G_0 (VG_0)^n. \tag{XVI.62}$$

Adopting the notation of § 8, we now look for the expansion of P as a power series in the perturbation. For sufficiently small λ, there certainly exists a contour in the complex z plane simultaneously enclosing the unperturbed eigenvalue $E_a{}^0$ and the eigenvalues of H that tend to $E_a{}^0$ in the limit when $\lambda \to 0$, and excluding all other eigenvalues of H and of H_0. Let such a contour be denoted by Γ_a. From (XVI.58),

$$P = \frac{1}{2\pi i} \oint_{\Gamma_a} G(z)\,\mathrm{d}z.$$

Substituting for G the expression in equation (XVI.62) and interchanging the order of summation and integration, we obtain the expansion of P as a power series in λ

$$P = P_0 + \sum_{n=1}^{\infty} \lambda^n A^{(n)}, \tag{XVI.63}$$

where

$$A^{(n)} = \frac{1}{2\pi i} \oint_{\Gamma_a} G_0 (VG_0)^n\,\mathrm{d}z. \tag{XVI.64}$$

The only singularity of the function $G_0\,(VG_0)^n$ inside the contour Γ_a is a pole of order $n+1$ in $E_a{}^0$. As seen from (XVI.64), $A^{(n)}$ is just the residue of that pole.

---

[1]) The radius of convergence for this expansion is $\|\lambda V G_0\|$. It is therefore absolutely convergent if $\|\lambda V\| < \Delta_0(z)$, where $\Delta_0(z)$ is the square of the distance from z to the nearest eigenvalue of H_0.

In order to calculate this residue we use the expansion of G_0 in a Laurent series about E_a^0. The coefficients are easily calculated using (XVI.56), and one finds

$$G_0 = \frac{P_0}{z - E_a^0} + \sum_{k=1}^{\infty} (-)^{k-1} (z - E_a^0)^{k-1} \frac{Q_0}{a^k}.$$

Following Kato we introduce the quantity S^k $(k > 0)$ defined by

$$S^k = \begin{cases} -P_0 & \text{if } k = 0 \\ \dfrac{Q_0}{a^k} & \text{if } k \geqslant 1 \end{cases} \tag{XVI.65}$$

so that our expansion becomes

$$G_0 = \sum_{k=0}^{\infty} (-)^{k-1} (z - E_a^0)^{k-1} S^k. \tag{XVI.66}$$

$A^{(n)}$ is the coefficient of $(z - E_a^0)^{-1}$ in the Laurent expansion of $G_0(VG_0)^n$. Taking (XVI.66) into account one easily finds

$$A^{(n)} = -\sum_{(n)} S^{k_1} V S^{k_2} V \ldots V S^{k_{n+1}}, \tag{XVI.67}$$

where $\sum_{(p)}$ extends over *all* sets of the non-negative integers

$$k_1, k_2, \ldots, k_{n+1}$$

such that:

$$k_1 + k_2 + \ldots + k_{n+1} = p. \tag{XVI.68}$$

As an illustration we give the first few terms in the expansion of P:

$$\begin{aligned} P = P_0 &+ \lambda \left(P_0 V \frac{Q_0}{a} + \frac{Q_0}{a} V P_0 \right) \\ &+ \lambda^2 \left(P_0 V \frac{Q_0}{a} V \frac{Q_0}{a} + \frac{Q_0}{a} V P_0 V \frac{Q_0}{a} + \frac{Q_0}{a} V \frac{Q_0}{a} V P_0 \right. \\ &\left. - P_0 V P_0 V \frac{Q_0}{a^2} - P_0 V \frac{Q_0}{a^2} V P_0 - \frac{Q_0}{a^2} V P_0 V P_0 \right) + \ldots \end{aligned} \tag{XVI.69}$$

Starting from (XVI.59), one obtains the expansion of HP in the same way. One finds:

$$(H - E_a^0) P = \sum_{n=1}^{\infty} \lambda^n B^{(n)}, \tag{XVI.70}$$

where

$$B^{(n)} = \sum_{(n-1)} S^{k_1} V S^{k_2} V \dots V S^{k_{n+1}}. \tag{XVI.71}$$

Thus

$$(H - E_a^0)P = \lambda P_0 V P_0 + \lambda^2 \left(P_0 V P_0 V \frac{Q_0}{a} + P_0 V \frac{Q_0}{a} V P_0 + \frac{Q_0}{a} V P_0 V P_0 \right) + \dots \tag{XVI.72}$$

17. Calculation of Eigenvalues and Eigenstates

The desired eigenvalues and eigenvectors of H are the eigenvalues and eigenvectors of the operator HP of the space $\mathscr{E}$ of which P is the projector. Since we know how to form P and HP, we have reduced the problem to that of diagonalizing a matrix in a g_a-dimensional space.

Non-degenerate case: $g_a = 1$

In this case the eigenvalue of H is $P|0\rangle$. The expansion of its norm, $\langle 0|P|0\rangle$, is easily obtained from that of the operator P. $P|0\rangle$ is a multiple of the eigenvector $|\psi\rangle$ formed in section I; the norm of the latter is easily shown to be $1/\langle 0|P|0\rangle$.

The eigenenergy E_a is given by the equation $HP = E_a P$. Since $\operatorname{Tr} P = \operatorname{Tr} P_0 = 1$, we have [eq. (XVI.70)]

$$E_a = \operatorname{Tr} HP = E_a^0 + \sum_{n=1}^{\infty} \lambda^n (\operatorname{Tr} B^{(n)}),$$

i.e.

$$\varepsilon_n = \operatorname{Tr} B^{(n)}. \tag{XVI.73}$$

Since each of the terms of $B^{(n)}$ contains P_0 at least once, and since, by the well-known property of the trace, one has the identity

$$\operatorname{Tr} MP_0N = \operatorname{Tr} P_0 NM = \langle 0|NM|0\rangle,$$

it is easy to cast ε_n into the form of the average value of a certain operator in the unperturbed state $|0\rangle$. The first few terms of the expansion of E_a are given by

$$E_a = E_a^0 + \lambda \langle 0|V|0\rangle + \lambda^2 \langle 0|V \frac{Q_0}{a} V|0\rangle + \lambda^3 \left(\langle 0|V \frac{Q_0}{a} V \frac{Q_0}{a} V|0\rangle - \langle 0|V \frac{Q_0}{a^2} V|0\rangle \langle 0|V|0\rangle \right) + \dots \tag{XVI.74}$$

in agreement with the results of section I [eq. (XVI.12) and (XVI.26)].

Degenerate case: $g_a \neq 1$

Rather than directly solve the eigenvalue problem for HP in $\mathscr{E}_a$, one can substitute a similar diagonalization problem in $\mathscr{E}_a^0$.

Let us suppose that any vector of $\mathscr{E}_a$ may be considered as the projection in $\mathscr{E}_a$ of a certain well-defined vector of $\mathscr{E}_a^0$ (this assumption was implicitly made in the non-degenerate case). Since $\mathscr{E}_a$ and $\mathscr{E}_a^0$ have the same number of dimensions, it follows that the projector P establishes a one-to-one correspondence between the vectors of $\mathscr{E}_a$ and the vectors of $\mathscr{E}_a^0$, and one can deduce that such a correspondence is also established by P_0. This property of "non-orthogonality" of the two subspaces is obviously realized for λ sufficiently small.

Thus any eigenvector of H in $\mathscr{E}_a$ can be put into the form $P|E_a^0\alpha\rangle$ and satisfies the equation:

$$HP|E_a^0\alpha\rangle = E_\alpha P|E_a^0\alpha\rangle.$$

A necessary and sufficient condition for this equality between vectors of $\mathscr{E}_a$ to be satisfied is that their projections on $\mathscr{E}_a^0$ be equal:

$$P_0HP|E_a^0\alpha\rangle = E_\alpha P_0P|E_a^0\alpha\rangle.$$

If we put

$$H_a = P_0HPP_0, \qquad K_a = P_0PP_0. \tag{XVI.75}$$

the preceding equation becomes

$$H_a|E_a^0\alpha\rangle = E_\alpha K_a|E_a^0\alpha\rangle. \tag{XVI.76}$$

K_a and H_a may be treated as operators in the space $\mathscr{E}_a^0$. K_a is a positive-definite Hermitian operator, H_a a Hermitian operator of this space. Equation (XVI.76) is a generalized eigenvalue equation. The eigenvalues E_α are the solutions of the secular equation (cf. § VII.17)

$$\det(H_a - xK_a) = 0.$$

They are the desired eigenenergies. The projections of the corresponding eigenvectors $|E_a^0\alpha\rangle$ in $\mathscr{E}_a$ are the eigenvectors of H.

The expansions of H_a and K_a are easily deduced from those of HP and P respectively [eq. (XVI.69) and (XVI.72)]:

$$K_a = P_0 - \lambda^2 P_0 V \frac{Q_0}{a^2} VP_0 + \ldots \tag{XVI.77}$$

$$H_a = E_a^0 K_a + \lambda P_0 VP_0 + \lambda^2 P_0 V \frac{Q_0}{a} VP_0 + \ldots \tag{XVI.78}$$

To a given order of approximation the eigenvalues E_α are obtained by stopping the expansions of H_a and K_a at that order. The results thereby obtained may differ from those of the elementary theory by quantities of higher order.

To the first order the results of the two methods are exactly the same. In making the comparison one must be careful to distinguish between $|E_a{}^0\alpha\rangle$ and the vector $|0\rangle$ of the elementary theory. $|0\rangle$ is the limit of $|E_a{}^0\alpha\rangle$ when $\lambda \to 0$.

A method very similar to that of Kato but leading to simpler expansions has been given by Bloch [1]). It is based on the fact that the operator $\mathscr{U}$ defined by

$$P P_0 \equiv \mathscr{U} K_a, \qquad \mathscr{U} P_0 \equiv \mathscr{U} \tag{XVI.79}$$

has a particularly simple expansion:

$$\mathscr{U} = \sum_{n=0}^{\infty} \lambda^n \, \mathscr{U}^{(n)} \tag{XVI.80}$$

$$\mathscr{U}^{(n)} = \sum_{(n)}{}' S^{k_1} V S^{k_2} V \ldots V S^{k_n} V P_0 \tag{XVI.81}$$

$\sum'_{(n)}$ being extended over all sets of non-negative integers $k_1, k_2, \ldots, k_n$ satisfying the conditions

$$\begin{aligned} &k_1 + k_2 + \ldots + k_p \geqslant p \quad (p = 1, 2, \ldots, n-1) \\ &k_1 + k_2 + \ldots + k_n = n. \end{aligned} \tag{XVI.82}$$

Observe that $P_0 \mathscr{U} = P_0$ and that $P \mathscr{U} = \mathscr{U}$. The action of $\mathscr{U}$ on a vector of $\mathscr{E}_a{}^0$ gives the vector of $\mathscr{E}_a$ of which it is the projection in $\mathscr{E}_a{}^0$.

According to the definition of $\mathscr{U}$, equation (XVI.76) is equivalent to

$$(P_0 H \mathscr{U} - E_a) \, K_a |E_a{}^0\alpha\rangle = 0. \tag{XVI.83}$$

This equation corresponds, in the method of Bloch, to equation (XVI.76) in the method of Kato. It is an ordinary eigenvalue equation for the *non-Hermitean* operator [2]) $P_0 H \mathscr{U} \equiv H_a K_a{}^{-1}$ of $\mathscr{E}_a{}^0$. The eigenvalues are the desired eigenenergies; the corresponding eigenvectors, $K_a |E_a{}^0\alpha\rangle$ are the projections of the eigenvectors of H onto

[1]) C. Bloch, Nuclear Physics, 6, (1958) 329.

[2]) Considered as an operator of the space $\mathscr{E}_a{}^0$, K_a has an inverse, as an operator of the entire Hilbert space it obviously has not.

$\mathscr{E}_a{}^0$; the eigenvectors themselves are obtained by application of the operator $\mathscr{U}$. The expansion of $P_0H\mathscr{U}$ is easily deduced from that of $\mathscr{U}$ if we recall that

$$P_0H\mathscr{U} = E_a{}^0P_0 + \lambda P_0V\mathscr{U}.$$

The first few terms are:

$$P_0H\mathscr{U} = E_a{}^0P_0 + \lambda P_0VP_0 + \lambda^2 P_0V\frac{Q_0}{a}VP_0 + \dots$$

These are seen to be simpler than those of H_a and K_a [eq. (XVI.77) and (XVI.78)].

In the non-degenerate case the energy is simply

$$E_a = \langle 0|H\mathscr{U}|0\rangle$$

from which we immediately extract the successive coefficients in its expansion:

$$\varepsilon_n = \sum'_{(n-1)} \langle 0|VS^{k_1}VS^{k_2}V \dots VS^{k_n}V|0\rangle.$$

EXERCISES AND PROBLEMS

1. The interaction $V(q)$ is added to the Hamiltonian $(p^2 + m^2\omega^2q^2)/2m$. Calculate with the perturbation theory the first- and second-order energy-level shifts for the following two cases:

(a) $\qquad V = \frac{1}{2}m\alpha^2q^2$

(b) $\qquad V = bq.$

In both cases the shifts can be exactly calculated. Compare the exact result with that of the perturbation theory.

2. Consider a hydrogen atom in the Schrödinger theory (Chapter XI). To what extent can the degeneracy of each level be removed by a constant electric field $\mathscr{E}$ (interaction $-e\mathscr{E}\cdot\mathbf{r}$)? Show that in the first order the level $n = 2$ splits into three equidistant levels (level distance $3\mathscr{E}\hbar^2/me$) and give the order of degeneracy of each.

3. Show that in the central field approximation the total angular momentum, the total orbital angular momentum and the total spin of the electrons of a complete shell are null. From this result deduce that the number of linearly independent vectors (LSM_LM_S) for a given configuration is the same as that obtained on taking into account only the electrons of incomplete shells.

4. What are the spectral terms of the ground state configuration of the carbon atom in jj coupling? Show that one obtains the same values for the total angular momentum J, with the same multiplicity, in jj and LS coupling (§ 11). Explain why.

5. The ground state configuration of the nitrogen atom ($Z = 7$) is $1s^2 2s^2 2p^3$. What are the different spectral terms in LS coupling?

6. Apply the considerations of § 13 to the study of the complex atoms of § 9, 10, 11. List the symmetries of H_C, V_1 and V_2 and show how and within what limits the successive introduction of the perturbations $V_i (i = 1, 2)$ removes the degeneracy of the levels of H_C: (a) in LS coupling (b) in jj coupling.

7. The ground-state configuration of a sodium atom ($Z = 11$) is $1s^2 2s^2 2p^6 3s^1$ and its first excited configuration is $1s^2 2s^2 2p^6 3p^1$. Therefore the ground-state of sodium is $^2S_{1/2}$ and its first two excited states are $^2P_{1/2}$ and $^2P_{3/2}$. Investigate the effect of a constant magnetic field $\mathcal{H}$ on these levels. Calculate the respective Landé factors. The intensity of $\mathcal{H}$ is gradually increased from the Zeeman region to the Paschen-Back. Give an expression for the ^2P levels as a function of the parameter $\varrho = \mu_B \mathcal{H}/A\hbar^2$ where μ_B is the Bohr magneton and A the spin-orbit coupling constant introduced in § 11 ($A > 0$); draw the corresponding curves.

8. Same question as in the preceding problem for the ground-state (^3P) of the carbon atom. [The levels are given as roots of secular equations, one of which is of the third degree; the main features of the curves can be obtained without solving these secular equations.]

9. The different levels $E_1, E_2, \ldots, E_n$ obtained from a g_a-fold degenerate level upon introduction of the perturbation V can be characterized by their center of gravity $\langle E \rangle$ and their root-mean-square deviation ΔE, defined respectively (with the notations of § 8) by

$$\langle E \rangle = \frac{1}{g_a} \sum_{i=1}^{n} g_i E_i \qquad (\Delta E)^2 = \frac{1}{g_a} \sum_{i=1}^{n} g_i (E_i - \langle E \rangle)^2.$$

Show that to the first order in V one has

$$\varepsilon \equiv \langle E \rangle - E_a{}^0 \simeq \frac{1}{g_a} \operatorname{Tr} P_0 V, \qquad (\Delta E)^2 \simeq \frac{1}{g_a} \operatorname{Tr} P_0 (V - \varepsilon) P_0 (V - \varepsilon).$$

Apply these formulas to the perturbation

$$V = A(\mathbf{L} \cdot \mathbf{S}) + \frac{e\hbar}{2mc} \mathcal{H} \cdot (\mathbf{L} + 2\mathbf{S})$$

of the two preceding problems. Show that in these cases $\varepsilon = 0$; give the expression for ΔE and examine its variation as a function of the parameter ϱ defined in Problem XVI.7.

APPROXIMATE SOLUTIONS OF THE TIME-DEPENDENT SCHRÖDINGER EQUATION

1. Change of "Representation" and Perturbation Treatment of a Part of the Hamiltonian

This chapter, with its two sections, is devoted to methods of obtaining approximate solutions to the time-dependent Schrödinger equation. Knowing that our quantum system is in a certain dynamical state at time t_0, we wish to determine its state at a later time t. The problem, therefore, is to determine as exactly as possible the operator $U(t, t_0)$ describing the evolution in time of the dynamical states of the system in the Schrödinger "representation".

Let us briefly recall the main properties of $U(t, t_0)$. This operator is completely determined once the Hamiltonian $H(t)$ of the system is given. It is defined by the integral equation

$$U(t, t_0) = 1 - i\hbar^{-1} \int_{t_0}^{t} H(\tau)\, U(\tau, t_0)\, d\tau \qquad \text{(XVII.1)}$$

or, equivalently, by the Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} U(t, t_0) = H(t)\, U(t, t_0) \qquad \text{(XVII.2)}$$

and the initial condition

$$U(t_0, t_0) = 1. \qquad \text{(XVII.3)}$$

Since $H(t)$ is Hermitean, U is unitary:

$$U(t, t')\, U^{\dagger}(t, t') = U^{\dagger}(t, t')\, U(t, t') = 1. \qquad \text{(XVII.4)}$$

In addition, we have the composition law [1])

$$U(t, t') = U(t, t'')\, U(t'', t') \qquad \text{(XVII.5)}$$

[1]) To demonstrate this formula, the physical sense of which is obvious, we note that if V is unitary and time-independent, then $U(t, t'')V$ is a solution of the Schrödinger equation (XVII.2). For it to be equal to $U(t, t')$ for any value of t, it is sufficient that it be equal for a particular value of t, say $t = t''$; this gives $V = U(t'', t')$.

whence

$$U^{\dagger}(t, t') = U(t', t). \tag{XVII.6}$$

An equivalent definition of U may be obtained by replacing (XVII.1) by its Hermitean conjugate. Taking (XVII.6) into account, we then obtain

$$U(t, t_0) = 1 - i\hbar^{-1} \int_{t_0}^{t} U(t, \tau)\, H(\tau)\, d\tau. \tag{XVII.7}$$

The methods described in this chapter are based on the following formal development. We suppose that the Hamiltonian H can be put in the form

$$H(t) \equiv H^{(0)}(t) + V(t), \tag{XVII.8}$$

where $H^{(0)}(t)$ is the Hamiltonian of a Schrödinger equation whose solution is known. Let $U^{(0)}(t, t_0)$ be the evolution operator corresponding to $H^{(0)}$:

$$i\hbar \frac{\partial}{\partial t} U^{(0)}(t, t_0) = H^{(0)}(t)\, U^{(0)}(t, t_0) \qquad U^{(0)}(t_0, t_0) = 1. \tag{XVII.9}$$

Since $U^{(0)}(t, t_0)$ is known, U will be determined if we can form the unitary operator

$$U_{\mathrm{I}}(t, t_0) \equiv U^{(0)\dagger}(t, t_0)\, U(t, t_0). \tag{XVII.10}$$

The physical significance of U_{I} has been given in § VIII.14: U_{I} is the evolution operator for the states in the intermediate "representation" derived from the Schrödinger representation by the unitary transformation $U^{(0)\dagger}(t, t_0)$. A simple calculation, the details of which were given in § VIII.14, shows that the time dependence of U_{I} is determined by the Hamiltonian

$$V_{\mathrm{I}}(t) \equiv U^{(0)\dagger}(t, t_0)\, V(t)\, U^{(0)}(t, t_0). \tag{XVII.11}$$

In other words, $U_{\mathrm{I}}(t, t_0)$ is defined by

$$i\hbar \frac{\partial}{\partial t} U_{\mathrm{I}}(t, t_0) = V_{\mathrm{I}}(t)\, U_{\mathrm{I}}(t, t_0) \qquad U_{\mathrm{I}}(t_0, t_0) = 1 \tag{XVII.12}$$

or equivalently, by

$$U_{\mathrm{I}}(t, t_0) = 1 - i\hbar^{-1} \int_{t_0}^{t} V_{\mathrm{I}}(\tau)\, U_{\mathrm{I}}(\tau, t_0)\, d\tau. \tag{XVII.13}$$

$U_{\mathrm{I}}(t, t_0)$ has all the properties of an evolution operator, and in

particular satisfies equations (XVII.1–7) with $H(t)$ everywhere replaced by $V_I(t)$.

The solution of integral equations (XVII.1), (XVII.7) and (XVII.13) can at least formally be obtained by iteration. Thus, substituting for $U_I(\tau, t_0)$ in the right-hand side of (XVII.13) the expression

$$1 - i\hbar^{-1} \int_{t_0}^{\tau} V_I(\tau')\, U_I(\tau', t_0)\, d\tau'$$

we get

$$U_I(t, t_0) = 1 - i\hbar^{-1} \int_{t_0}^{t} V_I(\tau)\, d\tau + (i\hbar)^{-2} \int_{t_0}^{t} d\tau \int_{t_0}^{\tau} d\tau'\, V_I(\tau)\, V_I(\tau')\, U(\tau', t_0). \quad \text{(XVII.14)}$$

Successive iterations give the expansion

$$U_I(t, t_0) = 1 + \sum_{n=1}^{\infty} U_I^{(n)}(t, t_0), \quad \text{(XVII.15)}$$

where $U_I^{(n)}$ is the integral

$$U_I^{(n)} \equiv (i\hbar)^{-n} \int_{t>\tau_n>\tau_{n-1}>\ldots>\tau_1>t_0} d\tau_n\, d\tau_{n-1} \ldots d\tau_1\, V_I(\tau_n)\, V_I(\tau_{n-1}) \ldots V_I(\tau_1). \quad \text{(XVII.16)}$$

From this result, with the aid of definitions (XVII.10) and (XVII.11), we get the following expansion for U:

$$U(t, t_0) = U^{(0)}(t, t_0) + \sum_{n=1}^{\infty} U^{(n)}(t, t_0) \quad \text{(XVII.17)}$$

$$U^{(n)}(t, t_0) = (i\hbar)^{-n} \int_{t>\tau_n>\tau_{n-1}>\ldots>\tau_1>t_0} d\tau_n\, d\tau_{n-1} \ldots d\tau_1\, U^{(0)}(t, \tau_n)\, V(\tau_n)\, U^{(0)}(\tau_n, \tau_{n-1})\, V(\tau_{n-1}) \ldots \times \ldots U^{(0)}(\tau_2, \tau_1)\, V(\tau_1)\, U^{(0)}(\tau_1, t_0). \quad \text{(XVII.18)}$$

Expansions (XVII.15) and (XVII.17) are power series in V which converge the more rapidly the closer $U^{(0)}(t, t_0)$ to $U(t, t_0)$. They are the basis of the methods set forth in this chapter. $U^{(0)}$ represents the zero order approximation, $U^{(1)}$, $U^{(2)}$, ..., $U^{(n)}$, ... respectively the corrections of order 1, 2, ..., n, ... to that approximation. In practice the calculation of these corrections becomes increasingly complicated with higher orders and one usually takes only those of the lowest order.

I. TIME DEPENDENT PERTURBATION THEORY

2. Definition and Perturbation Calculation of Transition Probabilities

The above method applies in particular when $H^{(0)}$ is *time-independ-*

ent. In this case the evolution operator $U^{(0)}(t, t')$ becomes simply

$$U^{(0)}(t, t') = \exp\left[-iH^{(0)}(t-t')/\hbar\right]. \tag{XVII.19}$$

We suppose the eigenvalue problem for $H^{(0)}$ to be solved. Unless otherwise specified, the spectrum will be assumed entirely discrete in order to simplify the writing. Let $|a\rangle, |b\rangle, \ldots, |k\rangle \ldots$ be a complete set of eigenvectors of $H^{(0)}$ and $E_a^0, E_b^0, \ldots, E_k^0, \ldots$ the corresponding eigenvalues. We shall use the notation

$$\omega_{kl} = (E_k^0 - E_l^0)/\hbar \tag{XVII.20}$$

$$V_{kl}(t) = \langle k|V(t)|l\rangle. \tag{XVII.21}$$

ω_{kl} is the "Bohr frequency" relative to the transition $l \to k$, $V_{kl}(t)$ the corresponding matrix element of $V(t)$.

We suppose that at the initial time t_0 the system is in an eigenstate of $H^{(0)}$, the state a say. We wish to calculate the probability that if a measurement is made at a later time t, the system will be found to be in a different eigenstate of $H^{(0)}$, the state b say. This quantity, by definition *the probability of transition from a to b*, will be denoted by $W_{a\to b}$. Clearly

$$W_{a\to b} = |\langle b|U(t, t_0)|a\rangle|^2. \tag{XVII.22}$$

If V were null, the vector representing the state of the system at time t would differ from the vector $|a\rangle$ representing the initial state only by a factor $\exp\left[-iE_a^0(t-t_0)/\hbar\right]$, and $W_{a\to b}$ would vanish. The expansion of the probability amplitude $\langle b|U(t, t_0)|a\rangle$ into a power series in V is obtained by substituting expansion (XVII.17) for $U(t, t_0)$:

$$\langle b|U(t, t_0)|a\rangle = \sum_{n=1}^{\infty} \langle b|U^{(n)}|a\rangle. \tag{XVII.23}$$

$U^{(n)}$ is given by (XVII.18).

In the $\{H^{(0)}\}$ representation defined above, the contributions of successive order to the amplitude are:

$$\langle b|U^{(1)}|a\rangle = -i\hbar^{-1} \int_{t_0}^{t} d\tau \left[e^{-iE_b^0(t-\tau)/\hbar}\, V_{ba}(\tau)\, e^{-iE_a^0(\tau-t_0)/\hbar}\right] \tag{XVII.24^1}$$

$$\langle b|U^{(2)}|a\rangle = (i\hbar)^{-2} \sum_k \int_{t_0}^{t} d\tau \int_{t_0}^{\tau} d\tau' \left[e^{-iE_b^0(t-\tau)/\hbar}\, V_{bk}(\tau)\, e^{-iE_k^0(\tau-\tau')/\hbar} \times V_{ka}(\tau')\, e^{-iE_a^0(\tau'-t_0)/\hbar}\right] \tag{XVII.24^2}$$

$$\langle b|U^{(3)}|a\rangle = (\mathrm{i}\hbar)^{-3} \sum_k \sum_l \int_{t_0}^{t} \mathrm{d}\tau \int_{t_0}^{\tau} \mathrm{d}\tau' \int_{t_0}^{\tau'} \mathrm{d}\tau'' \, [\mathrm{e}^{-\mathrm{i}E_b^0(t-\tau)/\hbar} V_{bk}(\tau) \times \mathrm{e}^{-\mathrm{i}E_k^0(\tau-\tau')/\hbar} V_{kl}(\tau') \, \mathrm{e}^{-\mathrm{i}E_l^0(\tau'-\tau'')/\hbar} V_{la}(\tau'') \, \mathrm{e}^{-\mathrm{i}E_a^0(\tau''-t_0)/\hbar}]. \quad \text{(XVII.24}^3\text{)}$$

In all of these expressions, the summations are extended over all of the basis vectors of $\{H^{(0)}\}$.

These contributions of successive order may be schematically represented by the diagrams of Figure XVII.1.

Diagram (*i*) represents the first-order contribution, or more exactly, it represents the product in brackets on the right-hand side of (XVII.24[1]). The full line describes the evolution of the system in the course of time. From t_0 to τ, this is determined by the unperturbed Hamiltonian $H^{(0)}$, and the system therefore remains in the state a, its state vector simply being multiplied by the factor

$$\exp\,[-\mathrm{i}E_a^0(\tau - t_0)/\hbar].$$

At time τ the system, under the influence of the perturbation, passes from a to b giving the transition matrix element $V_{ba}(\tau)$. From τ to t it evolves again according to $H^{(0)}$, and therefore remains in state b, giving the factor $\exp\,[-\mathrm{i}E_b(t-\tau)/\hbar]$. Thus, following the diagram from bottom to top, one obtains the three factors in the bracket in their order from right to left. The first-order contribution is obtained by integrating these products over τ.

Likewise, diagram (*ii*) represents the second-order correction. The system evolves according to $H^{(0)}$ from t_0 to τ', then under the influence

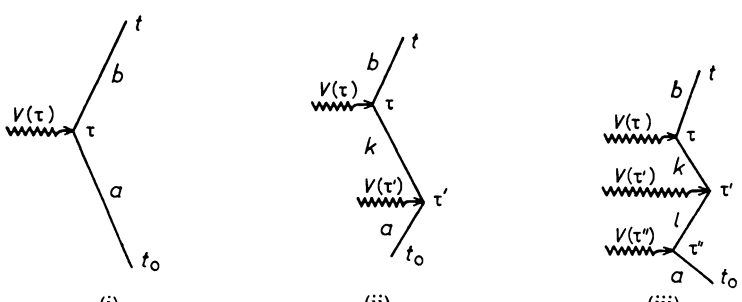


Fig. XVII.1. Diagrams representing the contributions of successive order tc the transition probability amplitude from a to b: (*i*) first order, (*ii*) second order, (*iii*) third order.

of $V(\tau')$, passes from the initial state a to the intermediate state k, then evolves according to $H^{(0)}$ from τ' to τ, then under the influence of $V(\tau)$ passes from state k to the final state b, and finally evolves according to $H^{(0)}$ from τ to t. We thereby obtain in their order of succession from right to left the five factors in the bracket of the right-hand side of (XVII.24²). The second-order contribution is obtained by integrating over τ' and $\tau(t_0 < \tau' < \tau < t)$ and summing over all of the intermediate states. Referring to this manner of representing the development of the second-order transitions in time, one often calls the state k a *virtual state*, as opposed to the real states a and b, and says that the second-order transition is effected through a virtual state.

In the same way, the third-order transitions, represented by diagram (*iii*), are effected through the two virtual states k and l. The perturbing potential comes into play three times in all, successively at times τ'', τ' and τ, taking the system from a into l, from l into k and from k into b. Similarily, the nth order transitions are made through $(n-1)$ virtual states.

The quantity obtained in taking the first n terms of expansion (XVII.23) is the sought for probability amplitude with a precision of order n. The square of its modulus gives, by definition, the transition probability to order n [1]), namely

$$W_{a\to b} \simeq |\langle b|U^{(1)}|a\rangle + \langle b|U^{(2)}|a\rangle + \ldots + \langle b|U^{(n)}|a\rangle|^2.$$

In particular, the *first-order transition probability* is given by

$$W_{a\to b} \simeq |\langle b|U^{(1)}|a\rangle|^2 = \hbar^{-2} \left|\int_{t_0}^{t} e^{i\omega_{ba}\tau} V_{ba}(\tau)\, d\tau\right|^2. \quad \text{(XVII.25)}$$

Observe that in this approximation

$$W_{b\to a} \simeq W_{a\to b}.$$

This relation ceases, in general, to be verified in approximations of higher order. It is not to be confused with the property of microreversibility, $W_{Kb\to Ka} = W_{a\to b}$, which is satisfied only if the Hamiltonian is

---

[1]) The expression obtained correctly represents the expansion of $W_{a\to b}$ in powers of V up to at least the order $n+1$. The precision obtained is of a higher order than $n+1$ if $\langle b|U^{(1)}|a\rangle = 0$.

invariant under time-reversal, in which case it is exactly true in all orders. (§ XV.20).

3. Semi-classical Theory of Coulomb Excitation of Nuclei

As an application of the theory we shall consider the Coulomb excitation of a nucleus by a charged particle, for example a proton [1]).

We suppose that a beam of mono-ergic protons collides with a nuclear target. In the course of the collision nuclei make transitions from their ground state into excited states. We propose to calculate the cross section for the transition to a given excited state β.

Denote by Ze the charge of the nucleus, by R its radius and by J_α and J_β, E_α and E_β, the spins and energies of states α and β respectively. There are $(2J_\alpha+1)$ linearly independent states α which may be distinguished one from another by the component M_α of the spin along a given quantization axis. They are represented by the vectors $|\alpha J_\alpha M_\alpha\rangle$. If H_N is the Hamiltonian for the nucleus, then

$$H_N|\alpha J_\alpha M_\alpha\rangle = E_\alpha|\alpha_\alpha J_\alpha M_\alpha\rangle \qquad (M_\alpha = -J_\alpha, \ldots, J_\alpha),$$

and similarly

$$H_N|\beta J_\beta M_\beta\rangle = E_\beta|\beta J_\beta M_\beta\rangle \qquad (M_\beta = -J_\beta, \ldots, J_\beta).$$

Denote by $E = \frac{1}{2}Mv^2$ the incident energy in the proton-nucleus center-of-mass system, by $\Delta E = E_\beta - E_\alpha$ the excitation energy of the nucleus, by Ω_α the direction of the incident proton, by Ω_β the direction of the inelastically scattered proton and by θ the angle between these two directions. The quantity to be calculated is the cross section

$$\mathrm{d}\sigma(\Omega_\alpha M_\alpha \to \Omega_\beta M_\beta)/\mathrm{d}\Omega$$

for the process where the proton is inelastically scattered into the direction Ω_β, the nucleus passing from the state $(\alpha J_\alpha M_\alpha)$ to the state $(\beta J_\beta M_\beta)$.

We now examine the proton-nucleus interaction. For a large distance r from the incident proton to the center of mass of the nucleus

---

[1]) On Coulomb exitation and its applications in the study of nuclear structure, see Alder *et al.*, Rev. Mod. Phys. 22 (1956) 432; the semi-classical treatment given here is due to K. A. Ter-Martyrosian, J. Exptl. Theoret. Phys. URSS 28 (1952) 284.

($r \gg R$) it reduces to the pure Coulomb interaction Ze^2/r. As r becomes smaller it increasingly deviates from this simple form. So long as $r > R$ the deviation is of purely electromagnetic origin and reduces essentially to the difference between the exact Coulomb interaction and the term Ze^2/r, namely

$$V = e^2 \sum_{i=1}^{Z} \left(\frac{1}{|\boldsymbol{r} - \boldsymbol{r}_i|} - \frac{1}{r} \right) \tag{XVII.26}$$

($\boldsymbol{r}_i$, position vector of the ith proton in the nucleus).

Once the proton has "penetrated" into the nucleus ($r < R$) the nuclear interactions proper enter into play and prevail largely over the electromagnetic interactions.

If the energy E is sufficiently low, the Coulomb repulsion Ze^2/r prevents the proton from approaching the nucleus and therefore remains the dominant interaction throughout the collision. The proton-nucleus motion is then governed in a first approximation by the Hamiltonian

$$H^{(0)} \equiv H_N + \left(\frac{\boldsymbol{p}^2}{2M} + \frac{Ze^2}{r} \right),$$

where $\boldsymbol{p}^2/2M$ is the kinetic energy of the proton. The motion of the proton separates completely from that of the nucleus. The latter remains in its ground state while the proton undergoes an elastic scattering process with a differential cross section given by the Rutherford formula (VI.29)

$$\mathrm{d}\sigma_R/\mathrm{d}\Omega = \tfrac{1}{4} a^2 \sin^{-4} \tfrac{1}{2}\theta$$

a is half the distance of closest approach in the classical motion:

$$a = \tfrac{1}{2} \frac{Ze^2}{E}. \tag{XVII.27}$$

The above approximation is justified so long as

$$a \gg R. \tag{XVII.28}$$

In addition, we shall suppose that

$$\frac{\Delta E}{E} \ll 1 \tag{XVII.29}$$

$$\gamma \gg 1, \tag{XVII.30}$$

where

$$\gamma \equiv \frac{a}{\lambda} = \frac{Ze^2}{\hbar v}. \tag{XVII.31}$$

Due to the deviation of the proton-nucleus interaction from Ze^2/r, inelastic collisions may occur. Since by condition (XVII.28) the proton "penetrates" very little into the nucleus, these deviations reduce essentially to the term V. Because of condition (XVII.30), the Coulomb scattering may be treated classically (§ VI.5). The motion of the proton is that of a wave-packet of negligible dimensions whose center $\boldsymbol{r}(t)$ obeys the corresponding classical equations of motion. In a given inelastic collision ($\Omega_\alpha M_\alpha \to \Omega_\beta M_\beta$) we can also treat the motion of the proton classically. The solution to the classical equations of motion is known if V is neglected. This implies that the energy ΔE transferred from the proton to the nucleus during the collision is also neglected. This approximation is legitimate if condition (XIII.29) is met. The trajectory of the proton $\boldsymbol{r}(t)$ being thus determined, *V is a time-dependent perturbation* acting on the dynamical variables of the nucleus:

$$V(t) = e^2 \sum_{i=1}^{Z} \left[\frac{1}{|\boldsymbol{r}(t) - \boldsymbol{r}_i|} - \frac{1}{r(t)} \right]$$

and can induce the transition $(\alpha J_\alpha M_\alpha) \to (\beta J_\beta M_\beta)$. The transition probability $W_{\alpha\to\beta}$ being small (one can verify *a posteriori* that $W_{\alpha\to\beta} \ll 1$), we take only the first-order approximation. From (XVII.25),

$$W_{\alpha\to\beta} = \hbar^{-2} \left| \int_{-\infty}^{+\infty} e^{i(\Delta E)t/\hbar} \langle \beta J_\beta M_\beta | V(t) | \alpha J_\alpha M_\alpha \rangle \, dt \right|^2. \tag{XVII.32}$$

The cross section in question is the product of this probability and the Rutherford cross section

$$d\sigma(\Omega_\alpha M_\alpha \to \Omega_\beta M_\beta) = (\tfrac{1}{4}a^2 \sin^{-4} \tfrac{1}{2}\theta) W_{\alpha\to\beta} \, d\Omega. \tag{XVII.33}$$

It remains to calculate $W_{\alpha\to\beta}$. We shall give only the outline of this calculation, omitting details. The polar coordinates of $\boldsymbol{r}_i$ and $\boldsymbol{r}(t)$ are denoted by (r_i, Ω_i) and (r, Ω) respectively. The latter are functions of t. If $|\boldsymbol{r}(t) - \boldsymbol{r}_i|^{-1}$ is expanded into spherical harmonics [eq. (B.99)], one finds

$$V(t) = \sum_{l=1}^{\infty} \sum_{m=-l}^{+l} (-)^m Q_l^m T_l^{-m}, \tag{XVII.34}$$

where

$$Q_l^m = \sum_{i=1}^{Z} e\, r_i^l Y_l^m(\Omega_i) \tag{XVII.35}$$

$$T_l^m(t) = \frac{4\pi e}{2l+1} \frac{Y_l^m(\Omega)}{r^{l+1}}. \tag{XVII.36}$$

This expansion is valid only if $r_i < r$, condition always fulfilled since the proton does not "penetrate" into the nucleus. Substituting into formula (XVII.32), we obtain

$$W_{\alpha\to\beta} = |\sum_{lm} (-)^m S_l^{-m} \langle \beta J_\beta M_\beta | Q_l^m | \alpha J_\alpha M_\alpha \rangle|^2 \tag{XVII.37}$$

$$S_l^m = \int_{-\infty}^{+\infty} e^{i(\Delta E)t/\hbar} T_l^m(t)\, dt/\hbar. \tag{XVII.38}$$

The coefficients S_l^m depend on the classical trajectory of the proton alone, and can be determined by numerical integration.

The $(2l+1)$ operators $Q_l^m (m = -l, \ldots, +l)$ are the standard components of the *electric* 2^l*-pole* moment $\mathbf{Q}^{(l)}$ (cf. § XIII.33). Therefore the only non-zero matrix elements figuring in formula (XVII.37) are those satisfying the angular momentum and parity selection rules:

$$\begin{aligned} &|J_\alpha - J_\beta| \leqslant l \leqslant J_\alpha + J_\beta \qquad m = M_\beta - M_\alpha \\ &\Pi_\alpha \Pi_\beta = (-)^l \end{aligned} \tag{XVII.39}$$

(Π_α, Π_β parities of states α and β respectively). Also, by the Wigner–Eckart theorem,

$$\langle \beta J_\beta M_\beta | Q_l^m | \alpha J_\alpha M_\alpha \rangle = (2J_\beta + 1)^{-\frac{1}{2}} \langle J_\alpha l M_\alpha m | J_\beta M_\beta \rangle \langle \beta \| \mathbf{Q}^{(l)} \| \alpha \rangle.$$

Owing to the selection rules (XVII.39), the sum in (XVII.37) is limited to a finite number of values of l of well-defined parity, and to a single value of m. A rough calculation shows that, all things being equal, the "$l+2$" contribution is about $(R/a)^2$ times the "l" contribution. We are justified, therefore, in keeping only the term corresponding to the smallest value of l allowed by the selection rules, either $|J_\alpha - J_\beta|$ or $|J_\alpha - J_\beta| + 1$ as the case may be. Calling this value l_0 we have

$$W_{\alpha\to\beta} \simeq \frac{\langle J_\alpha l_0 M_\alpha m | J_\beta M_\beta \rangle^2}{(2J_\beta + 1)} |S_{l_0}^{-m}|^2 |\langle \beta \| \mathbf{Q}^{(l_0)} \| \alpha \rangle|^2 \qquad (m = M_\beta - M_\alpha).$$

Substituting this expression in (XVII.33) we obtain a theoretical

expression for the cross section which may be compared with experiment.

In an experiment where the target nuclei are not oriented, and where the polarization of the excited nuclei is not observed, the measured cross section is that obtained by averaging the cross section defined above over the $2J_\alpha+1$ possible values of M_α, and summing it over the $2J_\beta+1$ possible values of M_β. Taking into account the orthogonality relations for the Clebsch–Gordon coefficients, this gives

$$\frac{\mathrm{d}\sigma_{\alpha\to\beta}}{\mathrm{d}\Omega} = \frac{a^2}{4(2J_\alpha+1)\,(2l_0+1)}\,|\langle\beta\|\mathbf{Q}^{(l_0)}\|\alpha\rangle|^2\left[\sin^{-4}\tfrac{1}{2}\theta\,\textstyle\sum_m |S^m_{l_0}|^2\right]$$

The angular dependence of this cross section is given by the expression in brackets and must be calculated numerically. Observe that the initial and final states of the nucleus enter into this formula only through their spin, their parity and the square of the modulus of the electric 2^{l_0}-pole transition moment, $\langle\beta\|\mathbf{Q}^{(l_0)}\|\alpha\rangle$, and that this last mentioned quantity appears only as a simple proportionality constant. The comparison of this formula with experiment is therefore a very direct way to determine these characteristic quantities of nuclear structure.

4. Case when V is Independent of Time. Conservation of Unperturbed Energy

When V does not depend explicitly on the time, the integrals over the time in formulas (XVII.24) are easily effected and certain simple properties appear. Let us limit the discussion to the *first-order transitions.*

We take $t_0=0$. Then, from (XVII.25)

$$W_{a\to b} \simeq |V_{ba}|^2\, f(t, \omega_{ba})/\hbar^2, \tag{XVII.40}$$

where

$$f(t, \omega) \equiv \left|\int_0^t \mathrm{e}^{\mathrm{i}\omega\tau}\,\mathrm{d}\tau\right|^2 = 2\,(1-\cos\omega t)/\omega^2. \tag{XVII.41}$$

$f(t, \omega)$ as a function of ω is shown in Fig. XVII.2. Note the very sharp peak about the value $\omega=0$; its width is equal to $2\pi/t$. Note in passing that

$$\int_{-\infty}^{+\infty} f(t, \omega)\,\mathrm{d}\omega = 2\pi t \tag{XVII.42}$$

as may easily be shown using the method of residues, and that, from (A.15*b*),

$$f(t, \omega) \underset{t\to\infty}{\sim} 2\pi t\delta(\omega). \qquad \text{(XVII.43)}$$

For a given value of t, $W_{a\to b}$ depends in a simple way on the final state b. To within a constant it is the squared modulus of the perturbation matrix element $\langle b|V|a\rangle$ *weighted* by the factor $f(t, \omega_{ba})$ which depends on the Bohr frequency for the transition $a \to b$. Since this weight factor has a very pronounced peak of width $2\pi/t$ at the point $\omega_{ba}=0$, the transition will be made preferentially towards states whose energy is situated in a band of width

$$\delta E_0 \simeq 2\pi\hbar/t$$

about the energy of the initial state. In other words, *the transitions conserve the unperturbed energy to within* $2\pi\hbar/t$.

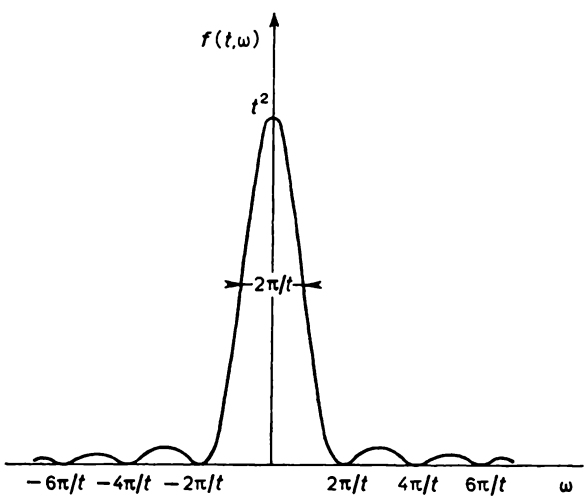


Fig. XVII.2. The function $f(t, \omega) = 2(1 - \cos\omega)/\omega^2$.

This result is in some ways analogous to the time-energy uncertainty relation (§ IV.10 and VIII.13). However, one must take note that here the energy in question is $H^{(0)}$ and not the total energy of the system, including the perturbation, and the time t is the time after which the measurement of $H^{(0)}$ is made and not the time characterizing the evolution of the system.

For a given state b, the variation of $W_{a\to b}$ with t is also given by the factor $f(t, \omega_{ba})$. If the transition exactly conserves the unperturbed energy ($\omega_{ba}=0$) it simply increases as t^2. If not, it is a function that oscillates between 0 and $4/\omega_{ba}^2$ with period $2\pi/\omega_{ba}$. $W_{a\to b}$ oscillates with the same period about the average value $2|V_{ba}|^2/(E_b-E_a)^2$ and has the t^2 behavior only for values of t small compared with that period.

Rather than consider transitions to a particular state, we may consider *transitions to a group of states of neighboring energies*. This is what is always done for transitions to states of the continuous spectrum. One may then, under certain limiting conditions to be presently specified, define a *probability of transition per unit time*.

Let us therefore consider a particular sequence of eigenvectors of $H^{(0)}$ belonging to the continuous spectrum. $|b\rangle$ will denote a particular one of these, and $E(b)$ the corresponding eigenvalue of $H^{(0)}$.

In the definition of transition probabilities, attention must be paid to the normalization of the $|b\rangle$. We shall suppose them normalized such that

$$\langle b|b'\rangle = \delta(b-b')/n(b),$$

where $n(b)$ is some real positive function. The projector onto the states of a certain domain B of the variable b is (cf. § VIII.13):

$$P_{\mathrm{B}} = \int_{\mathrm{B}} |b\rangle\, n(b)\, \mathrm{d}b\, \langle b|.$$

If we take $E(b)$ as a new variable and denote by $\mathrm{B}(E)$ the corresponding domain of integration, we find

$$P_{\mathrm{B}} = \int_{\mathrm{B}(E)} |b\rangle\, \varrho_b(E)\, \mathrm{d}E\, \langle b| \qquad \text{(XVII.44)}$$

where

$$\varrho_b(E) = n(b)\, \mathrm{d}b/\mathrm{d}E. \qquad \text{(XVII.45)}$$

$\varrho_b(E)$ is known as the *density of the levels b at energy E*. According to (XVII.44), it is indeed the number of vectors $|b\rangle$ per unit interval of energy; it will be observed that $\varrho_b(E)$ depends on the normalization of the $|b\rangle$.

The transition probability $W_{a\to B}$ into one of the states of the domain B is then

$$\begin{aligned} W_{a\to B} &\equiv \langle a|U^{\dagger}(t,0)\, P_B\, U(t,0)|a\rangle \\ &= \int_{B(E)} W_{a\to b}\, \varrho_b(E)\, \mathrm{d}E. \end{aligned} \tag{XVII.46}$$

Formula (XVII.46) is obtained by replacing P_B by expression (XVII.44) and putting

$$W_{a\to b} = |\langle b|U(t,0)|a\rangle|^2. \tag{XVII.47}$$

Here $W_{a\to b}$ formally represents the probability for the transition $a \to b$ defined by equation (XVII.22). The various steps in the calculation of that quantity are all valid here. In particular, by substituting (XVII.40) into the right-hand side of (XVII.44), one obtains the transition probability $W_{a\to B}$ to the first order in the perturbation:

$$W_{a\to B} \simeq \int |V_{ba}|^2\, \varrho_b\, f(t, \omega_{ba})\, \mathrm{d}E/\hbar^2 \tag{XVII.48}$$

here $V_{ba} \equiv \langle b|V|a\rangle$ depends on E through the parameter b.

As a definite example consider the transitions to the levels b inside the interval $(E_1 - \frac{1}{2}\varepsilon, E_1 + \frac{1}{2}\varepsilon)$, and suppose the width ε sufficiently small that V_{ba} and ϱ_b are practically constant over the interval so that they can be taken outside the integral sign. We also suppose t sufficiently large for ε to be much greater than the period of oscillation of the function f:

$$\varepsilon \gg 2\pi\hbar/t. \tag{XVII.49}$$

The integral on the right-hand side of (XVII.49) is then easily evaluated. Two cases are to be considered:

(*i*) The central peak of f is outside the domain of integration (transitions not conserving the energy). One may then replace f by its value averaged over several oscillations, which leads to the time-independent expression

$$W_{a\to B} \simeq 2\varepsilon\, \varrho_b(E_1)\, |V_{ba}|^2/(E_1 - E_a)^2.$$

(*ii*) The central peak of f is in the domain of integration (energy-conserving transitions); the main contribution comes from this peak, and only a small error is involved in extending the limits of the integral to infinity, whence [eq. (XVII.42)]:

$$W_{a\to B} \simeq 2\pi\hbar^{-1} |V_{ba}(E_a)|^2\, \varrho_b(E_a)\, t.$$

Owing to inequality (XVII.49), this probability is greater than the sum of all the others.

We now introduce the probability of transition per unit time:

$$w_{a\to B} \equiv \mathrm{d}W_{a\to B}/\mathrm{d}t.$$

From the preceding results we can conclude that this quantity vanishes for transitions not conserving the energy, and that for those transitions which do conserve the energy it is given by the important formula

$$\boxed{w_{a\to B} \simeq \frac{2\pi}{\hbar} |V_{ba}|^2 \varrho_b} \qquad \text{(XVII.50)}$$

in which the matrix element $V_{ba} \equiv \langle b|V|a\rangle$ and the level density ϱ_b relate to states b whose energy is equal to that of the initial state.

For formula (XVII.50) to be valid, t must be sufficiently large for conditions (XVII.49) to be met, and at the same time be sufficiently small for the first-order approximation to be justified ($w_{a\to B}\, t \ll 1$).

The above demonstration has the advantage of bringing out the significance and the conditions of validity of formula (XVII.50). The latter can be more simply obtained by replacing f by its asymptotic expression (XVII.43) in the right-hand side of (XVII.40), which gives

$$W_{a\to b} \underset{t\to\infty}{\sim} 2\pi\hbar^{-1} |V_{ba}|^2 \, \delta(E_b - E_a)\, t, \qquad \text{(XVII.51)}$$

and then substituting this into the definition of $W_{a\to B}$ [eq. (XVII.46)].

5. Application to the Calculation of Cross-sections in the Born Approximation

From formula (XVII.50) one may deduce an expression for the collision cross sections in the so-called *Born approximation*, that is, to the first order in the interaction potential between the particle and the target. Here we shall give a simple derivation, of this formula somewhat at the expense of rigor. A rigorous derivation of the same formula will be given in Chapter XIX.

We consider the simplest possible case, the scattering of a particle by a potential $V(\boldsymbol{r})$. The latter is treated as a perturbation. The unperturbed Hamiltonian is that of a free particle:

$$H = H^{(0)} + V(\boldsymbol{r}) \qquad H^{(0)} = \boldsymbol{p}^2/2m.$$

The eigenstates of $H^{(0)}$ are the plane waves $e^{i\mathbf{k}\cdot\mathbf{r}}$. Such a wave represents a state of momentum $\mathbf{p}=\hbar\mathbf{k}$ *normalized to unit density.* We shall denote the corresponding ket vectors by $|\mathbf{k}\rangle$: they obey the orthogonality and closure relations

$$\langle\mathbf{k}|\mathbf{k}'\rangle = (2\pi)^3\,\delta(\mathbf{k}-\mathbf{k}') \qquad \int |\mathbf{k}\rangle \frac{d\mathbf{k}}{(2\pi)^3} \langle\mathbf{k}| = 1.$$

In the space of the $\mathbf{k}$ vectors, the density of the states thus normalized is therefore constant and equal to $(2\pi)^{-3}$: the number of states in the interval $(\mathbf{k}, \mathbf{k}+d\mathbf{k})$ is $d\mathbf{k}/(2\pi)^3$. We are interested in the states of momentum in a particular direction Ω, and denote their density, as defined in the previous paragraph, by $\varrho(E)$ [eq. (XVII.45)] [*a priori* this function could depend on Ω; we shall see that in fact it does not]: $\varrho(E)\,d\Omega\,dE$ is the number of states whose momentum is in the solid angle $(\Omega, \Omega+d\Omega)$ and whose energy $E=p^2/2m$ is in the band $(E, E+dE)$. We have, therefore

$$\varrho(E)\,d\Omega\,dE = d\mathbf{k}/(2\pi)^3 = d\mathbf{p}/(2\pi\hbar)^3$$

and since $d\mathbf{p}=p^2\,d\Omega\,dp$, this gives

$$\varrho(E)\,dE = p^2\,dp/(2\pi\hbar)^3,$$

and therefore

$$\varrho(E) = \frac{p^2}{(2\pi\hbar)^3}\frac{dp}{dE} = \frac{mp}{(2\pi\hbar)^3}. \tag{XVII.52}$$

Let us now proceed to calculate the scattering cross section from a mono-ergic incident beam of energy $E=\frac{1}{2}mv_a^2$ into a given direction Ω_b. Let $\hbar\,\mathbf{k}_a$ be the momentum of the incident particles and $\hbar\,\mathbf{k}_b$ the momentum corresponding to the same energy E but with direction Ω_b. Starting from the initial state $|\mathbf{k}_a\rangle$, we know that the probability per unit time $w_{a\to B}\,d\Omega$ that the system make a transition into one of the states b with momentum in the solid angle $(\Omega_b, \Omega_b+d\Omega)$ and energy close to E is given to the first order by formula (XVII.50):

$$w_{a\to B}\,d\Omega \simeq 2\pi\hbar^{-1}\,|\langle\mathbf{k}_b|V|\mathbf{k}_a\rangle|^2\,\varrho(E)\,d\Omega. \tag{XVII.53}$$

Let $d\sigma_{a\to b}/d\Omega$ be the differential scattering cross section. $d\sigma_{a\to b}$ is the number of particles scattered into the solid angle $(\Omega_b, \Omega_b+d\Omega)$ per unit time and per unit incident flux. Since $|\mathbf{k}_a\rangle$ represents a wave of flux v_a, we have

$$d\sigma_{a\to b} = w_{a\to B}/v_a.$$

Substituting for $w_{a\to B}$ from the approximate expression (XVII.53) we have

$$\frac{d\sigma_{a\to b}}{d\Omega} \simeq \frac{2\pi}{\hbar v_a} |\langle \boldsymbol{k}_b | V | \boldsymbol{k}_a \rangle|^2 \varrho_b(E), \tag{XVII.54}$$

where $\varrho_b(E)$ is the density of final states [eq. (XVII.52)], and

$$\langle \boldsymbol{k}_b | V | \boldsymbol{k}_a \rangle \equiv \int e^{i(\boldsymbol{k}_a - \boldsymbol{k}_b)\boldsymbol{r}} V(\boldsymbol{r})\, d\boldsymbol{r}$$

is the matrix element of the potential responsible for the transition.

6. Periodic Perturbation. Resonances

To the first order [formula (XVII.25)], the probability of transition $W_{a\to b}$ is proportional to the squared modulus of the amplitude of the frequency ω_{ba} in the harmonic analysis of the function $V_{ba}(t)$, where by convention we have put $V_{ba}=0$ outside the interval (t_0, t). This harmonic analysis is particularly simple for V independent of time and gives rise, as we have seen, to the "conservation of the unperturbed energy". It is also very simple in the more general case where V is a periodic function of the time, and here we have a phenomenon of great practical importance — the phenomenon of *resonance*.

Suppose that V is a sinusoidal function of time with frequency ω. Since it is a Hermitian operator it can be put in the form

$$V = A\, e^{i\omega t} + A^\dagger e^{-i\omega t},$$

where A is a certain time-independent operator. To the first order, the probability of transition $W_{a\to b}$ is given by (taking $t_0=0$):

$$W_{a\to b} \simeq \hbar^{-2} \left| \langle b|A|a\rangle \int_0^t e^{i(\omega_{ba}+\omega)\tau}\, d\tau + \langle b|A^\dagger|a\rangle \int_0^t e^{i(\omega_{ba}-\omega)\tau}\, d\tau \right|^2$$

which may be compared with expression (XVII.40).

The amplitude of transition is here made up of two terms. For t sufficiently large the first term is small unless $\omega_{ba}+\omega$ is close to 0, that is, unless E_b is to be found in a band (of width $2\pi\hbar/t$) about the point

$$E_b = E_a - \hbar\,\omega, \tag{XVII.55}$$

while the second term is small except in a band (of the same width) about the point

$$E_b = E_a + \hbar\,\omega. \tag{XVII.55'}$$

In practice, t is sufficiently large ($t \gg 2\pi/\omega$) for these regions not to overlap. $W_{a\to b}$ is then small except for transitions in which *the unperturbed system emits or absorbs the quantity of energy* $\hbar\omega$, as indicated by equations (XVII.55) and (XVII.55′) respectively.

In the first case only the first term is involved in the transition amplitude and the transition probability reduces to the expression

$$W_{a\to b} \simeq |A_{ba}|^2 f(t, \omega_{ba} + \omega)/\hbar^2.$$

The essential difference between this result and (XVII.40) is that here ω_{ba} is replaced by $\omega_{ba}+\omega$. In complete analogy with the work of § 4 we can now consider transitions to a group of levels in an energy band ΔE ($\gg 2\pi\hbar/t$) about the point $E_a-\hbar\omega$ and define, under suitable conditions, a probability of transition per unit time; this is again given by formula (XVII.50) with the sole difference that V_{ba} and ϱ_b now relate to states b with energy $\hbar\omega$ less than the initial state. The same considerations apply to transitions in which the system absorbs a quantity of energy $\hbar\omega$ (cf. Problem XVII.2).

Consider now the more general case when V is any periodic function of the time, of frequency ω. Here we merely state the results; the demonstrations are left to the reader. One then has the Fourier expansion

$$V = \sum_{s=1}^{\infty} (A_s\, \mathrm{e}^{\mathrm{i}s\omega t} + A_s^{\dagger}\, \mathrm{e}^{-\mathrm{i}s\omega t}).$$

If $t \gg 2\pi/\omega$ the contributions to the first-order transition probability from different terms of this series do not interfere since each causes transitions corresponding to a different energy transfer. In the "A_s transitions" the system loses, to within $2\pi\hbar/t$, the energy $s\hbar\omega$; in the "$A_s^{\dagger}$ transitions" it gains, to within $2\pi\hbar/t$, the energy $s\hbar\omega$.

II. SUDDEN OR ADIABATIC CHANGE OF THE HAMILTONIAN

7. The Problem and the Results

A frequently occurring problem is that of finding the modification in the state of the system when we change the external field. The classic example is that of an atomic system put into a magnetic field. As a general rule, the results obtained depend critically on the time T during which the modification of the Hamiltonian takes place.

In this section we examine the limiting cases when T is very small (sudden change) and very large (adiabatic change).

We suppose the Hamiltonian to change-over in a continuous way from a certain initial value H_0 at time t_0 to a certain final value H_1, at time t_1. We put

$$T = t_1 - t_0 \qquad s = (t - t_0)/T$$

and denote by $H(s)$ the value taken by the Hamiltonian at time $t = t_0 + sT$. $H(s)$ is a continuous function of s and we have:

$$H(0) = H_0 \qquad H(1) = H_1.$$

The evolution of the system from t_0 to t_1 now depends only on the parameter T measuring the speed of the passage from H_0 to H_1. It is convenient to put

$$U(t, t_0) = U_T(s).$$

Our problem is essentially to determine $U(t_1, t_0)$, that is, $U_T(1)$, and to examine its dependence on T.

The results are remarkably simple for the two limiting cases mentioned above.

In the limit when $T \to 0$, i.e. in the case of an infinitely rapid passage, the dynamical state of the system remains unchanged:

$$\lim_{T \to 0} U_T(1) = 1. \tag{XVII.56}$$

In the limit when $T \to \infty$, i.e. in the case of an infinitely slow, or adiabatic passage, if the system is initially in an eigenstate of H_0 it will, at time t_1, under certain conditions to be specified later, *have passed into the eigenstate of H_1, that derives from it by continuity.* This important result is known as the *Adiabatic Theorem* [1]).

8. Rapid Passage and the Sudden Approximation

The first of the above results follows immediately from equation (XVII.1) for the operator describing the evolution of the system.

---

[1]) It is also called Ehrenfest's theorem. The work of Ehrenfest on the subject relates to Classical Mechanics and the Old Quantum Theory. The extension of the theorem to Quantum Mechanics is due essentially to M. Born and V. Fock, Zeit. f. Phys. **51** (1928) 165. See also T. Kato, Journ. Phys. Soc. Jap. **5** (1950) 435; K. O. Friedrichs, *On the Adiabatic Theorem in Quantum Theory*, Report IMM.NYU-218, (New York, 1955).

With the notations of § 7 it reads

$$U_T(s) = 1 - i\hbar^{-1} T \int_0^s H(s)\, U_T(s)\, ds.$$

In the limit when $T \to 0$ the second term on the right-hand side goes to zero and we obtain the desired result (XVII.56).

For T sufficiently small we may, in a first approximation, suppose that $U_T(1) \simeq 1$. This is called the *sudden approximation.*

Let $|0\rangle$ denote the state vector of the system at time t_0 and Q_0 the projector onto the space of the vectors orthogonal to $|0\rangle$. Supposing $|0\rangle$ to be of norm 1, we have

$$Q_0 = 1 - |0\rangle\langle 0|.$$

The sudden approximation consists in writing

$$U(t_1, t_0)|0\rangle \simeq |0\rangle.$$

A measure of the error involved in this approximation is given by the probability ϖ of finding the system in a state other than the initial state:

$$\begin{aligned} \varpi &= \langle 0|U^\dagger(t_1, t_0)\, Q_0\, U(t_1, t_0)|0\rangle \\ &= \langle 0|U_T{}^\dagger(1)\, Q_0\, U_T(1)|0\rangle. \end{aligned} \qquad \text{(XVII.57)}$$

The corrections to this approximation can be calculated by the perturbation method set forth in § 1. In the present case

$$(H^{(0)} = 0,\ V = H),$$

the zero order approximation is simply the operator 1 and expansion (XVII.17) gives

$$U_T(1) = 1 - i\hbar^{-1} T \int_0^1 H(s)\, ds + (i\hbar)^{-2} T^2 \int_0^1 ds_1 \int_0^{s_1} ds_2\, H(s_1)\, H(s_2) + \dots . \qquad \text{(XVII.58)}$$

In particular, substituting this expansion into the right-hand side of (XVII.57), one obtains the expansion of ϖ in powers of T. As $Q_0|0\rangle = 0$, the lowest-order term is in T^2, and is obtained by substituting in the first two terms of (XVII.58). Put

$$\bar{H} = \int_0^1 H(s)\, ds = \frac{1}{T} \int_{t_0}^{t_1} H\, dt. \qquad \text{(XVII.59)}$$

We then have

$$\varpi = \frac{T^2}{\hbar^2} \langle 0|\bar{H}\, Q_0\, \bar{H}|0\rangle + O(T^3).$$

And since

$$\langle 0|\bar{H}\, Q_0\, \bar{H}|0\rangle = \langle 0|\bar{H}^2|0\rangle - \langle 0|\bar{H}|0\rangle^2 = (\Delta\bar{H})^2$$

where $\Delta\bar{H}$ is the root mean square deviation of the observable $\bar{H}$ in the state $|0\rangle$, one has

$$\varpi = \frac{T^2(\Delta\bar{H})^2}{\hbar^2} + O(T^3). \tag{XVII.60}$$

Thus the condition for the validity of the sudden approximation, $\varpi \ll 1$, requires that [1])

$$\boxed{T \ll \hbar/\Delta\bar{H}} \tag{XVII.61}$$

Condition (XVII.61) is nothing else but a particular form of the time-energy uncertainty relation. Following definition (XVII.59), $\bar{H}$ is the Hamiltonian of the system averaged over the interval (t_0, t_1). The evolution of the system is therefore roughly governed by the Hamiltonian $\bar{H}$ during that time interval. According to the time-energy uncertainty relation, the state of a system obeying such an equation of motion cannot appreciably be modified before a time $\hbar/\Delta\bar{H}$ has elapsed. Condition (XVII.61), therefore, is indeed the condition that the modification of the state after a time T will be negligible.

9. Sudden Reversal of a Magnetic Field

As an application we consider what happens to an atom in a constant magnetic field when the direction of the field is suddenly reversed. We suppose the atom to be one for which the LS coupling scheme holds, and the field to be sufficiently strong to completely decouple the total angular momentum $\boldsymbol{L}$ from the total spin $\boldsymbol{S}$ (Paschen–Back effect). For simplicity we shall also suppose the magnetic field to be

---

[1]) The condition $\varpi \ll 1$ does not necessarily imply that $U_1(t_1, t_0)\,|0\rangle \simeq |0\rangle$ but only that these two vectors differ essentially only by a phase factor. However, formula (XVII.60) supposes a rapid convergence of the expansion (XVII.58), and in general (XVII.61) is a sufficient condition for this to be so.

always parallel to the z axis, and to pass from the value $-\mathscr{H}_0$ to the value $\mathscr{H}_0$ according to the linear law

$$\mathscr{H}(t) = \mathscr{H}_0[2s-1] = \mathscr{H}_0[2(t-t_0)/T-1]. \qquad \text{(XVII.62)}$$

According to the study of Chapter XVI (§ 9–12), the Hamiltonian of the system is of the form

$$H = H^{(0)} + A(\mathbf{L}\cdot\mathbf{S}) - \frac{e}{2mc}(L_z + 2S_z)\,\mathscr{H}(t) \qquad \text{(XVII.63)}$$

$H^{(0)}$ is the unperturbed Hamiltonian in the LS coupling scheme. Since according to law (XVII.62) the time average of the magnetic field is zero, the time average of H is equal to the Hamiltonian of the atom without external field:

$$\bar{H} = H^{(0)} + A(\mathbf{L}\cdot\mathbf{S}).$$

We shall now suppose that initially the system is in an eigenstate of H. Since at that time we have the conditions of the Paschen–Back effect, the state vector will be very nearly equal to one of the vectors $|\alpha LSM_LM_S\rangle$ defined in § XVI.11. We therefore put: $|0\rangle = |\alpha LSM_LM_S\rangle$. The calculation of $\Delta\bar{H}$ is straightforward. Since $|\alpha LSM_LM_S\rangle$ is an eigenvector of $H^{(0)} + AL_zS_z$ the only contribution comes from the difference $\frac{1}{2}A(L_+S_- + L_-S_+)$. The calculation gives

$$\Delta\bar{H} = \tfrac{1}{2}A\hbar^2[2(L(L+1) - M_L{}^2)\,(S(S+1) - M_S{}^2) - 2M_LM_S]^{\frac{1}{2}}.$$

The bracket is a numerical factor of the order of 1 (it vanishes in the two extreme cases $M_L = \pm L$, $M_S = \pm S$). The deviation $\Delta\bar{H}$ is therefore of the order of magnitude of $A\hbar^2$, i.e. of the order of magnitude of the splitting of the LS levels by spin-orbit coupling.

In consequence, if

$$T \ll \frac{1}{A\hbar} \qquad \text{(XVII.64)}$$

the rapid passage condition (XVII.61) is met: the state vector remains practically constant when we reverse the field. For the dynamical state of the atom to remain constant it is sufficient that the state vector be modified only by a phase factor; inequality (XVII.64) is therefore a sufficient but not a necessary condition for the atom to remain in the same state when the field is reversed. We shall return to this point at the end of § 14.

10. Adiabatic Passage. Generalities. Trivial Case

In the rest of this section we shall examine the other extreme case, that of a very slow modification of the Hamiltonian. We follow the notations of § 7.

We first wish to establish the Adiabatic Theorem, which is a property of the states of the discrete spectrum of $H(s)$. Although it is in no way an essential assumption, for the sake of simplicity we shall suppose the spectrum of H to be entirely discrete [1]).

Denote the eigenvalues of H by $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_j \ldots$, and the projectors onto their respective subspaces by $P_1, P_2, \ldots, P_j, \ldots$. These quantities are all supposed to be *continuous functions of* s. In addition we suppose that:

(*i*) *the eigenvalues remain distinct* throughout the whole transition period $0 \leqslant s \leqslant 1$:

$$\varepsilon_j(s) \neq \varepsilon_k(s), \qquad \text{whatever } j \text{ and } k \tag{XVII.65}$$

(*ii*) *the derivatives* $\mathrm{d}P_j/\mathrm{d}s$, $\mathrm{d}^2P_j/\mathrm{d}s^2$ are well-defined and piece-wise continuous in the whole interval.

The evolution operator $U_T(s)$ satisfies the Schrödinger equation

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}s} U_T(s) = TH(s)\, U_T(s), \tag{XVII.66}$$

the Hamiltonian $H(s)$ being given by the expression

$$H(s) = \sum_j \varepsilon_j(s)\, P_j(s). \tag{XVII.67}$$

The Adiabatic Theorem states that $U_T(s)$ *has the asymptotic property* [2]):

$$\lim_{T\to\infty} U_T(s)\, P_j(0) = P_j(s) \lim_{T\to\infty} U_T(s) \qquad (j = 1, 2, \ldots). \tag{XVII.68}$$

---

[1]) It is sufficient that the discrete eigenvalues and their respective subspaces satisfy the continuity, "non-crossing" and differentiability conditions indicated below. See T. Kato, *loc. cit.*, note, p. 740.

[2]) This property is equivalent to the one given in § 7, for if $|j\rangle$ is an eigenvector of $H(0)$ belonging to the eigenvalue $\varepsilon_j(0)$, then $P_j(0)\,|j\rangle = |j\rangle$ and (XVII.68) gives

$$\lim_{T\to\infty} U_T(s)\,|j\rangle = P_j(s) \lim_{T\to\infty} U_T(s)\,|j\rangle.$$

Thus the vector $U_T(s)\,|j\rangle$ tends toward a vector of the subspace of $\varepsilon_j(s)$ when $T \to \infty$ as stated in § 7.

Let us first suppose that the subspace of each eigenvalue of $H(s)$ remains unchanged

$$P_j(s) = P_j(0) \equiv P_j \qquad (j = 1, 2, \ldots).$$

In this case $H(s)$ takes the simple form

$$H(s) = \sum_j \varepsilon_j(s)\, P_j$$

and commutes with each projector P_j whatever s; each P_j is therefore a constant of the motion:

$$U_T(s)\, P_j\, U_T^{\dagger}(s) = P_j. \tag{XVII.69}$$

Relation (XVII.69) is verified for any T and *a fortiori* for $T \to \infty$.

In addition, equation (XVII.66) is exactly integrable in this particular case and gives

$$\begin{aligned} U_T(s) &= \exp\left(-\mathrm{i}T \int_0^s H(\sigma)\, \mathrm{d}\sigma/\hbar\right) \\ &= \sum_j \mathrm{e}^{-\mathrm{i}T\varphi_j(s)/\hbar}\, P_j. \end{aligned} \tag{XVII.70}$$

where we use the notation

$$\varphi_j(s) = \int_0^s \varepsilon_j(\sigma)\, \mathrm{d}\sigma. \tag{XVII.71}$$

We therefore have the result that if at time t_0 the state vector of the system is an eigenvector of H_0 belonging to the eigenvalue $\varepsilon_j(0)$, at time t_1 it differs from this only by the phase factor $\mathrm{e}^{-\mathrm{i}T\varphi_j(1)/\hbar}$.

11. "Rotating Axis Representation"

In the general case, the exact integration of the Schrödinger equation is no longer possible as the eigenvectors of the Hamiltonian $H(s)$ are effecting a certain rotational motion in Hilbert space. The first step in treating the general case will be to eliminate this rotational motion as far as possible through an appropriate change of "representation".

To this effect, we introduce a unitary operator $A(s)$ having the property

$$P_j(s) = A(s)\, P_j(0)\, A^{\dagger}(s) \qquad (j = 1, 2, \ldots). \tag{XVII.72}$$

The unitary transformation $A(s)$ takes any set of basis vectors of $H(0)$ over into a set of basis vectors of $H(s)$, each eigenvector of $H(0)$

being carried over into one of the eigenvectors of $H(s)$ that derive from it by continuity.

$A(s)$ is unambiguously defined by the initial condition

$$A(0) = 1 \tag{XVII.73}$$

and the differential equation

$$i\hbar \, dA/ds = K(s) \, A(s) \tag{XVII.74}$$

where $K(s)$ is an appropriate Hermitian operator. For (XVII.72) to be satisfied it is necessary and sufficient that $K(s)$ obey the commutation relations

$$[K(s), P_j(s)] = i\hbar \, dP_j/ds \qquad (j = 1, 2, \ldots). \tag{XVII.75}$$

These are necessary since they are immediately obtained upon differentiating both sides of (XVII.72) with respect to s. They are sufficient since, $A(s)$ and $P_j(s)$ satisfying equations (XVII.74) and (XVII.75) respectively, the expression

$$A^{\dagger}(s) \, P_j(s) \, A(s)$$

has a vanishing derivative with respect to s, and is therefore equal to its initial value $P_j(0)$.

$K(s)$ is not completely defined by relations (XVII.75) which are still satisfied if we add to $K(s)$ the operator $\sum_k P_k(s) f_k(s) P_k(s)$, where the $f_k(s)$ are arbitrary operators depending on s. In other words the projections $P_j(s)K(s)P_j(s)$ $(j = 1, 2, \ldots)$ may be arbitrarily fixed. For reasons that will presently become clear, we remove the arbitrary by imposing the additional condition

$$P_j(s) \, K(s) \, P_j(s) = 0 \qquad (j = 1, 2, \ldots). \tag{XVII.76}$$

This gives (Problem XVII.5)

$$K(s) = i\hbar \sum_j (dP_j/ds) \, P_j(s).$$

The unitary transformation $A^{\dagger}(s)$ carries the vectors and operators of the Schrödinger "representation" over into the vectors and operators of a new "representation", the "rotating axis representation". The observable $H(s)$ transforms into

$$H^{(A)}(s) = A^{\dagger}(s) \, H(s) \, A(s)$$

giving, with the aid of (XVII.67) and (XVII.72),

$$H^{(A)}(s) = \sum_j \varepsilon_j(s)\, P_j(0). \tag{XVII.77}$$

Similarly $K(s)$ becomes

$$K^{(A)}(s) = A^\dagger(s)\, K(s)\, A(s). \tag{XVII.78}$$

The evolution operator in this new "representation" is

$$U^{(A)}(s) \equiv A^\dagger(s)\, U_T(s). \tag{XVII.79}$$

It is defined [cf. § 1, eq. (XVII.12), where one must take $V = TH - K$] by

$$i\hbar\, \mathrm{d}U^{(A)}/\mathrm{d}s = [TH^{(A)}(s) - K^{(A)}(s)]\, U^{(A)}(s), \tag{XVII.80}$$

$$U^{(A)}(0) = 1. \tag{XVII.81}$$

12. Proof of the Adiabatic Theorem

Equation (XVII.80) would be easily integrable if one could neglect $K^{(A)}$ as compared with $TH^{(A)}$. We would then have the trivial case studied in § 10. Let us denote by $\Phi_T(s)$ the solution of the resulting Schrödinger equation:

$$i\hbar\, \mathrm{d}\Phi_T/\mathrm{d}s = TH^{(A)}(s)\, \Phi_T(s) \tag{XVII.82}$$

$$\Phi_T(0) = 1. \tag{XVII.83}$$

We find [eq. (XVII.70)]

$$\Phi_T(s) = \sum_j \mathrm{e}^{-\mathrm{i}T\varphi_j(s)/\hbar}\, P_j(0) \tag{XVII.84}$$

the φ_j being still defined by (XVII.71).

However, referring to (XVII.77) and (XVII.78) we see that $H^{(A)}(s)$ and $K^{(A)}(s)$ are independent of T. We may therefore expect that in the limit $T \to \infty$ the effect of $K^{(A)}$ on the right-hand side of (XVII.80) will be completely masked by that of $TH^{(A)}$, and that $U^{(A)}(s)$ will tend towards $\Phi_T(s)$. As we shall see, this is indeed the case, and therefore [eq. (XVII.79)]

$$U_T(s) \underset{T\to\infty}{\sim} A(s)\, \Phi_T(s). \tag{XVII.85}$$

To prove this result we begin by effecting a new unitary transformation and putting

$$W \equiv \Phi_T{}^\dagger\, U^{(A)} = \Phi_T{}^\dagger\, A^\dagger\, U_T. \tag{XVII.86}$$

We easily deduce the equation satisfied by this new unitary operator from (XVII.80) and (XVII.82). In its integral form it reads

$$W(s) = 1 + \frac{\mathrm{i}}{\hbar} \int_0^s \bar{K}(\sigma)\, W(\sigma)\, \mathrm{d}\sigma, \tag{XVII.87}$$

where

$$\bar{K}(s) \equiv \Phi_T^{\dagger}(s)\, K^{(A)}(s)\, \Phi_T(s) \tag{XVII.88}$$

$$= \Phi_T^{\dagger} A^{\dagger} K A \Phi_T. \tag{XVII.89}$$

We are going to show that the kernel $\bar{K}(s)$ is a sum of oscillating functions whose frequencies increase indefinitely with T, and that in consequence the integral on the right-hand side of the Volterra equation (XVII.87) tends to zero when $T \to \infty$.

Any operator Q can be decomposed as follows [1]):

$$Q = \sum_j \sum_k P_j(0)\, Q\, P_k(0).$$

We shall henceforth employ the notation

$$Q_{jk} \equiv P_j(0)\, Q\, P_k(0).$$

From (XVII.72), (XVII.84) and (XVII.89)

$$\begin{aligned} \bar{K}_{jk} &= \mathrm{e}^{\mathrm{i}T(\varphi_j - \varphi_k)/\hbar}\, K_{jk}^{(A)} \\ K_{jk}^{(A)} &= A^{\dagger}(s)\, P_j(s)\, K(s)\, P_k(s)\, A(s). \end{aligned} \tag{XVII.90}$$

According to condition (XVII.76), the $K_{jj}^{(A)}$ $(j = 1, 2, \ldots)$ all vanish, and therefore the diagonal parts $\bar{K}_{jj}$ of the decomposition of $\bar{K}$ are all exactly null:

$$\bar{K}_{jj} = 0 \qquad (j = 1, 2, \ldots). \tag{XVII.91}$$

The non-diagonal parts $\bar{K}_{jk}$ $(j \neq k)$ contain the oscillating factor

$$\mathrm{e}^{\mathrm{i}T(\varphi_j - \varphi_k)/\hbar} \equiv \exp\left[\mathrm{i}\hbar^{-1} T \int_0^s (\varepsilon_j(\sigma) - \varepsilon_k(\sigma))\, \mathrm{d}\sigma\right].$$

The circular frequency of the oscillations is obtained by differentiating the phase of the exponential with respect to s, which gives

$$T\, |\varepsilon_j(s) - \varepsilon_k(s)|/\hbar.$$

---

[1]) The procedure employed here amounts to using a representation in which $H(0)$ is diagonal and at the same time avoiding the various complications arising from the possible degeneracy of the levels of $H(0)$ and the arbitrariness in the phase of the basis vectors.

According to the hypothesis (XVII.65) the difference $\varepsilon_j - \varepsilon_k$ never vanishes, and therefore the frequency increases as T when $T \to \infty$.

Consider now the operator

$$F(s) \equiv \int_0^s \bar{K}(\sigma)\, \mathrm{d}\sigma. \tag{XVII.92}$$

According to (XVII.91) its diagonal elements are all null:

$$F_{jj} = 0.$$

The non-diagonal elements are given by the expression

$$F_{jk} = \int_0^s \mathrm{e}^{\mathrm{i}T(\varphi_j - \varphi_k)/\hbar} K_{jk}^{(A)}\, \mathrm{d}\sigma \qquad (j \neq k). \tag{XVII.93}$$

$K_{jk}^{(A)}$ is a continuous function of s independent of T. The phase of the exponential, however, does depend on T and F_{jk} is therefore of the form $\int_0^s \mathrm{e}^{\mathrm{i}T\alpha(\sigma)} f(\sigma)\, \mathrm{d}\sigma$, where $f(\sigma)$ is a continuous function and $\alpha(\sigma)$ is a continuous monotonic function. Such an integral is known to go to zero when $T \to \infty$. Indeed, integrating by parts, it becomes

$$F_{jk}(s) = \frac{\hbar}{\mathrm{i}T}\left[\mathrm{e}^{\mathrm{i}T(\varphi_j - \varphi_k)/\hbar} \frac{K_{jk}^{(A)}}{\varepsilon_j - \varepsilon_k}\right]_0^s - \int_0^s \mathrm{e}^{\mathrm{i}T(\varphi_j - \varphi_k)/\hbar} \left[\frac{\mathrm{d}}{\mathrm{d}\sigma}\left(\frac{K_{jk}^{(A)}}{\varepsilon_j - \varepsilon_k}\right)\right] \mathrm{d}\sigma\Bigg]. \tag{XVII.94}$$

Clearly the expression in brackets remains finite if $K_{jk}^{(A)}$ and the derivatives of $K_{jk}^{(A)}$, ε_j and ε_k with respect to s remain finite. F_{jk} therefore tends asymptotically to zero as $1/T$. In conclusion, when $T \to \infty$ we have

$$F(s) = \mathrm{O}\left(\frac{1}{T}\right).$$

Now the integral on the right-hand side of (XVII.87) can be written, after integration by parts,

$$F(s)\, W(s) - \int_0^s F(\sigma) \frac{\mathrm{d}W}{\mathrm{d}\sigma}\, \mathrm{d}\sigma,$$

or again,

$$F(s)\, W(s) - \mathrm{i}\hbar^{-1} \int_0^s F(\sigma)\, \bar{K} W(\sigma)\, \mathrm{d}\sigma \tag{XVII.95}$$

where we have taken into account the equation $\mathrm{d}W/\mathrm{d}s = \mathrm{i}\bar{K}W/\hbar$. Both

terms in (XVII.95) contain the factor F, and hence tend to zero as $1/T$ when $T \to \infty$. In consequence [1])

$$W = 1 + O\left(\frac{1}{T}\right). \tag{XVII.96}$$

Substituting (XVII.96) into the definition (XVII.86) of W, one finds

$$U_T(s) \underset{T\to\infty}{\sim} A(s)\, \Phi_T(s) \left[1 + O\left(\frac{1}{T}\right)\right] \tag{XVII.97}$$

which completes the proof of (XVII.85).

Since $\Phi_T(s)$ commutes with the projectors $P_j(0)$ [cf. eq. (XVII.84)] and since the unitary operator $A(s)$ has the property (XVII.72), one has

$$A(s)\, \Phi_T(s)\, P_j(0) = P_j(s)\, A(s)\, \Phi_T(s).$$

Therefore the asymptotic expression (XVII.97) does indeed verify relations (XVII.68). Q.E.D.

13. Adiabatic Approximation

If T is sufficiently large or, better, *if the basis vectors of $H(t)$ rotate sufficiently slowly*, we can, in a first approximation, replace $U_T(1)$ by its asymptotic form:

$$U(t_1, t_0) \equiv U_T(1) \simeq A(1)\, \Phi_T(1). \tag{XVII.98}$$

This is called *the adiabatic approximation.*

Let $|0\rangle$ be the vector of norm 1 representing the system at time t_0 and let Q_0 be the projector onto the complementary space. The adiabatic approximation consists in writing

$$U(t, t_0)|0\rangle \simeq A(1)\, \Phi_T(1)|0\rangle.$$

---

[1]) In all rigor, the above argument is only valid if the operator $K(s)$ is bounded over the whole interval (0, 1) (cf. note, p. 713). It can then be shown that $\|F(s)\|$ tends uniformly to zero with respect to s as $1/T$. Denote by $\varkappa$ and ε the upper limits of $\|K(s)\|$ and of $\|F(s)\|$ respectively. Since W, Φ and A are unitary, the norms of K, $K^{(A)}$, $\bar{K}$ and KW are equal; similarly the norms of F and FW are equal. It follows that

$$\|FW\| \leqslant \varepsilon, \quad \left\|\int_0^s F\bar{K}W\, d\sigma\right\| < \varepsilon\varkappa s.$$

And since $W - 1$ is given by (XVII.95),

$$\|W - 1\| \leqslant \varepsilon(1 + \varkappa s).$$

A measure of the error involved in this approximation is given by the probability η of finding the system at time t_1 in a state different to $A(1)\Phi_T(1)|0\rangle$. The projector onto the space orthogonal to this vector is

$$Q_1 \equiv A(1)\,\Phi_T(1)\,Q_0\,\Phi_T^{\dagger}(1)\,A^{\dagger}(1),$$

so that

$$\begin{aligned}\eta &\equiv \langle 0|U^{\dagger}(t_1, t_0)\,Q_1\,U(t_1, t_0)|0\rangle \\ &= \langle 0|W^{\dagger}(1)\,Q_0\,W(1)|0\rangle.\end{aligned}$$

Corrections to the adiabatic approximation may be calculated by the perturbation method of § 1. Here $A(1)\Phi_T(1)$ plays the role of $U^{(0)}$ and $W(1)$ that of U_{I}. The method consists in retaining only the initial terms in the expansion of W obtained by iteration of equation (XVII.87). If we go only to first order we find

$$W(1) \simeq 1 + (\mathrm{i}F/\hbar), \tag{XVII.99}$$

where F is the operator $F(1)$ defined by (XVII.92).

This expansion may also be used to evaluate η. A treatment analogous to the calculation of ϖ in the case of the sudden approximation gives

$$\eta \simeq \hbar^{-2}\,\langle 0|F\,Q_0\,F|0\rangle = (\Delta F/\hbar)^2, \tag{XVII.100}$$

where ΔF is the root-mean-square-deviation of the observable F in the state $|0\rangle$. The condition $\eta \ll 1$, condition for the validity of the adiabatic approximation [1]), becomes

$$\Delta F \ll \hbar.$$

In the general form given here this result is much less useful than the corresponding result for the sudden approximation [relation (XVII.61)]. The observable F involved here is much more difficult to construct than the observable $\bar{H}$ of the sudden approximation. $\bar{H}$ is the result of a simple integration of $H(t)$; to calculate F one must solve the eigenvalue problem for $H(t)$ at each point of the interval (t_0, t_1) and construct the operator $A(s)$ for each of these points.

Let us examine the validity conditions for the adiabatic approxi-

---

[1]) The same remarks can be made in this connection as were made in the note, p. 742 in connection with the criterion for the validity of the sudden approximation.

mation in the case when the system is initially in an eigenstate of $H(0)$. This is, in fact, the only case of practical interest.

Henceforth we shall use the variable t itself rather than the substitute s. We thus denote the eigenvalues of the Hamiltonian at time t by $\varepsilon_j(t)$, the corresponding projectors by $P_j(t)$, the operator giving our "rotating axes" by $A(t)$ and the operator $\Phi_T(s)$ by $\Phi(t)$. Definitions (XVII.67) and (XVII.84) are therefore superceded by the definitions:

$$H(t) = \sum_j \varepsilon_j(t)\, P_j(t) \tag{XVII.101}$$

$$\Phi(t) = \sum_j \exp\left[-\mathrm{i} \int_{t_0}^{t} \varepsilon_j(\tau)\, \mathrm{d}\tau/\hbar\right] P_j(0). \tag{XVII.102}$$

Equations (XVII.73) and (XVII.74) defining A become

$$\mathrm{i}\hbar\, \mathrm{d}A/\mathrm{d}t = K'(t)\, A(t) \qquad A(0) = 1 \tag{XVII.103}$$

with

$$K'(t) \equiv K(s)/T = \mathrm{i}\hbar \sum_j (\mathrm{d}P_j/\mathrm{d}t)\, P_j(t). \tag{XVII.104}$$

From its definition, (XVII.92), we get

$$F \equiv F(t_1) = \int_{t_0}^{t_1} \Phi^\dagger(t)\, A^\dagger(t)\, K'(t)\, A(t)\, \Phi(t)\, \mathrm{d}t. \tag{XVII.105}$$

It is further assumed, for simplicity, that the spectrum of $H(t)$ is completely non-degenerate. We choose a set of basis vectors $|1\rangle_0, |2\rangle_0, \ldots, |j\rangle_0, \ldots$ of the Hamiltonian H_0 and denote the set of basis vectors of $H(t)$ obtained from these by application of the transformation $A(t)$ by $|1\rangle_t, |2\rangle_t, \ldots |j\rangle_t, \ldots$. For any t,

$$\left.\begin{aligned} |j\rangle_t &= A(t)\, |j\rangle_0 \\ H(t)\, |j\rangle_t &= \varepsilon_j(t)\, |j\rangle_t \\ P_j(t) &= |j\rangle_t\, {}_t\langle j| \end{aligned}\right\} \quad (j = 1, 2, \ldots). \tag{XVII.106}$$

Let $|i\rangle_0$ be the state vector of the system at time t_0. In the adiabatic approximation, its state vector at time t_1 will be equal to within a phase factor to $|i\rangle_1 \equiv A(t_1)|i\rangle_0$:

$$U(t_1, t_0)|i\rangle_0 \simeq \exp\left[-\mathrm{i} \int_{t_0}^{t_1} \varepsilon_i(\tau)\, \mathrm{d}\tau/\hbar\right] |i\rangle_1. \tag{XVII.107}$$

The probability $p_{i\to j}$ $(j \neq i)$ of finding the system in another eigenstate of $H(t)_1$, say $|j\rangle_1$, is by definition

$$p_{i\to j} = |{}_1\langle j|U(t_1, t_0)|i\rangle_0|^2.$$

The perturbation calculation described above gives

$$p_{i\to j} \simeq |{}_0\langle j|F|i\rangle_0|^2/\hbar^2. \tag{XVII.108}$$

In agreement with the result (XVII.100), we have

$$\eta_i = \sum_{j\neq i} p_{i\to j} \simeq \hbar^{-2} \sum_{j\neq i} {}_0\langle i|F|j\rangle_0\, {}_0\langle j|F|i\rangle_0$$

With the aid of equations (XVII.102), (XVII.104) and (XVII.105), and the properties of $|j\rangle_t$ as given by (XVII.106), we find

$${}_0\langle j|F|i\rangle_0 = i\hbar \int_{t_0}^{t_1} \alpha_{ji}(t) \exp\left[i \int_{t_0}^{t} \omega_{ji}(\tau)\, d\tau\right] dt,$$

where

$$\begin{aligned} \alpha_{ji} &\equiv {}_0\langle j|A^\dagger(t)\, (dP_t/dt)\, A(t)|i\rangle_0 \\ &= {}_t\langle j|(d|i\rangle_t/dt) \end{aligned} \tag{XVII.109}$$

$$\omega_{ji}(t) = [\varepsilon_j(t) - \varepsilon_i(t)]/\hbar. \tag{XVII.110}$$

Therefore, from (XVII.108)

$$p_{i\to j} \simeq \left| \int_{t_0}^{t_1} \alpha_{ji}(t) \exp\left[i \int_{t_0}^{t} \omega_{ji}(\tau)\, d\tau\right] dt \right|^2. \tag{XVII.111}$$

The physical interpretation of the quantities α_{ji} and ω_{ji} can be read directly from definitions (XVII.109) and (XVII.110). $\alpha_{ji}(t)$ is a measure of the speed of rotation of the eigenvectors of $H(t)$. It is the component along $|j\rangle_t$ of the velocity of the axis $|i\rangle_t$. $\omega_{ji}(t)$ is the "Bohr frequency" of the transition $i \to j$.

The integrand on the right-hand side of (XVII.111) is the product of the function $\alpha_{ji}(t)$ by an oscillating exponential of frequency $\omega_{ji}(t)$. If α_{ji} and ω_{ji} are time-independent it is easily integrated to give

$$p_{i\to j} \simeq \left|\frac{\alpha_{ji}}{\omega_{ji}}\right|^2 2(1 - \cos \omega_{ji} T).$$

$p_{i\to j}$ is therefore a quantity of the order of $|\alpha_{ji}/\omega_{ji}|^2$. If α_{ji} and ω_{ji}, instead of being just constant, exhibit a sufficiently smooth variation in time, then $p_{i\to j}$ will at most be of the order of magnitude of the maximum value attained by the ratio $|\alpha_{ji}/\omega_{ji}|^2$ in the interval (t_0, t_1):

$$p_{i\to j} \lesssim \max \left|\frac{\alpha_{ji}(t)}{\omega_{ji}(t)}\right|^2. \tag{XVII.112}$$

Similarly η_i is at most of the order of magnitude of the quantity obtained on summing the right-hand side of (XVII.112) over all states j different to i. This sum will generally be smaller than the expression $|\alpha_i{}^{\max}/\omega_i{}^{\min}|^2$ in which $\omega_i{}^{\min}$ is the minimum value of the Bohr frequency for the transition from i to its nearest neighbor and $\alpha_i{}^{\max}$ is the maximum value of the positive quantity $\alpha_i(t)$ defined by

$$\alpha_i{}^2(t) = \sum_{j \neq i} |\alpha_{ji}(t)|^2.$$

Returning to the definition (XVII.109) of α_{ji} and noting that [1])

$$_t\langle i| \left(\frac{\mathrm{d}}{\mathrm{d}t} |i\rangle_t\right) = 0 \qquad \text{(XVII.113)}$$

we see that α_i is the length of the vector $\mathrm{d}|i\rangle_t/\mathrm{d}t$, that is, the "angular velocity" of the eigenvector $|i\rangle_t$. The condition $\eta_i \ll 1$ is therefore, in most cases, certainly satisfied if

$$\boxed{\left|\frac{\alpha_i{}^{\max}}{\omega_i{}^{\min}}\right|^2 \equiv \left|\frac{\text{maximum angular velocity of } |i\rangle_t}{\text{minimum Bohr frequency of } |i\rangle_t}\right|^2 \ll 1} \qquad \text{(XVII.114)}$$

Condition (XVII.114) may be taken as a criterion for the validity of the adiabatic approximation. It is in fact too restrictive, but has the advantage of being relatively easy to handle. It is to be particularly noted that it is not necessary to determine $A(t)$ in order to calculate either $\alpha_i(t)$ or the quantity $|\alpha_{ji}(t)|$ figuring in the right-hand side of (XVII.112); we need only to solve the eigenvalue problem for $H(t)$ and to determine to within a phase factor the vectors $|i\rangle_t, |j\rangle_t$.

---

[1]) From (XVII.104) and (XVII.106) we have

$$\frac{\mathrm{d}}{\mathrm{d}t} |j\rangle_t = \frac{\mathrm{d}}{\mathrm{d}t} A(t) |j\rangle_0 = \frac{\mathrm{d}P_j}{\mathrm{d}t} |j\rangle_t \qquad (j = 1, 2, \ldots).$$

Confronting this result and the one we obtain from term by term differentiation of the identity $P_j(t) |j\rangle_t = |j\rangle_t$ we find $P_j(\mathrm{d}|j\rangle_t/\mathrm{d}t) = 0$; therefore

$$_t\langle j| \left(\frac{\mathrm{d}}{\mathrm{d}t} |j\rangle_t\right) = 0 \qquad (j = 1, 2, \ldots). \qquad \text{(XVII.113')}$$

The vectors $|j\rangle_t$ are defined to within a phase factor by the condition that they are eigenvectors of $H(t)$. This phase factor is fixed by condition (XVII.113').

Furthermore, we can show that

$$\alpha_{ji}(t) = -\,{}_t\langle j|\,\frac{\mathrm{d}H}{\mathrm{d}t}\,|i\rangle_t/\hbar\omega_{ji}(t)$$

(cf. Problem (XVII.6).

14. Adiabatic Reversal of a Magnetic Field

Let us now return to the problem of § 9, following the same notation. We start from the same initial conditions, but now suppose condition (XVII.64) no longer to be satisfied.

The Hamiltonian being given by (XVII.63), we see that α, L, S and $M_J \equiv M_L + M_S$ are good quantum numbers ($H(t)$ commutes with $J_z \equiv L_z + S_z$, whatever t). Therefore, if $|\alpha L S M_L M_S\rangle$ is the initial state-vector, its subsequent evolution will be in the space of vectors having the same values of α, L, S and $M_L + M_S$.

We shall examine in detail the case when the system is initially in a ^{2}P state. There are in all 6 different ^{2}P states, linear combinations of the 6 basis vectors $|\alpha 1 \tfrac{1}{2} M_L M_S\rangle$ ($M_L = 1, 0, -1$; $M_S = \tfrac{1}{2}, -\tfrac{1}{2}$) which will henceforth be denoted $|M_L M_S\rangle$. Since they are all eigenstates of $H^{(0)}$, we can take the corresponding eigenvalue as the zero of the energy. The levels corresponding to the energy H are easily determined (Problem XVI.7). They are functions of the parameter $\varrho = \mu_B \mathscr{H}/A\hbar^2$ ($\mu_B \equiv e\hbar/2mc$ = Bohr magneton) and are plotted in Figure XVII.3. Each level corresponds to a well-defined value of $M_J = M_L + M_S$, and the corresponding eigenvector tends toward a particular $|M_L M_S\rangle$ in each of the limits $\varrho \to \pm\infty$, as indicated in the figure. We shall successively examine the cases $M_J = \frac{3}{2}$ and $M_J = \frac{1}{2}$.

We first consider the case when $M_J = \frac{3}{2}$. There is just one state for this value of J_z, and the corresponding vector is $|1\ \tfrac{1}{2}\rangle$. It is necessarily an eigenvector of H for any t. A simple calculation gives

$$H|1\ \tfrac{1}{2}\rangle = A\hbar^2(\tfrac{1}{2} - 2\varrho(t))|1\ \tfrac{1}{2}\rangle.$$

If the atom is initially in the state $|1\ \tfrac{1}{2}\rangle$, we will have a case of the trivial integration of the Schrödinger equation described in § 10. In the course of time the atom remains in the same state, its state vector simply being multiplied by the phase factor

$$\exp\left[-\mathrm{i}A\hbar \int_{t_0}^{t} (\tfrac{1}{2} - 2\varrho(\tau))\,\mathrm{d}\tau\right].$$

In the case of a linear variation of the magnetic field [law (XVII.62)], the state vector at the end of a time T is equal to

$$\exp(-iA\hbar T)|1\ \tfrac{1}{2}\rangle.$$

This is true whatever T. In particular, in the limit when $A\hbar T \ll 1$, we get the result given by the sudden approximation, i.e. the vector $|1\ \frac{1}{2}\rangle$ itself.

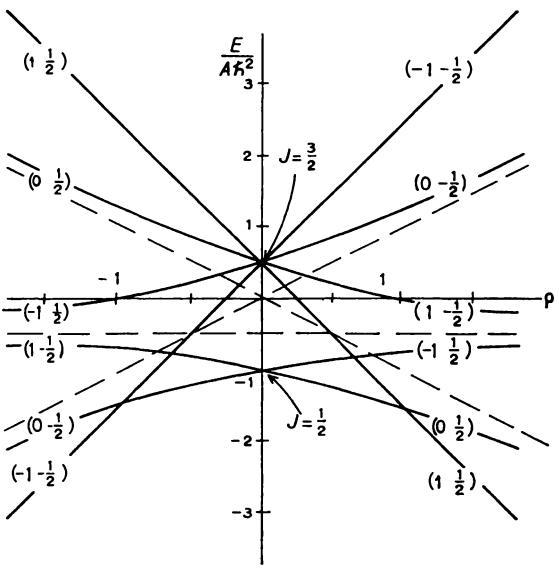


Fig. XVII.3. Position of the ^2P levels as a function of the intensity of the magnetic field $\mathscr{H}$ ($\varrho = \mu_B \mathscr{H}/A\hbar^2$). The numbers in parenthesis at the end of each curve are the quantum numbers ($M_L M_S$) of the eigenstates in the two limits $\varrho \to \pm\infty$.

We next consider the case when $M_J = \frac{1}{2}$. To this value of J_z there correspond two eigenstates of H, each a linear combination of the vectors $|0\ \frac{1}{2}\rangle$ and $|1\ -\frac{1}{2}\rangle$. In order to solve the eigenvalue problem for $H(t)$ in the subspace spanned by these two vectors we take the two of them as basis vectors, in which case $H(t)$ is represented by the matrix

$$A\hbar^2 \begin{pmatrix} -\varrho & \frac{1}{2}\sqrt{2} \\ \frac{1}{2}\sqrt{2} & -\frac{1}{2} \end{pmatrix}.$$

If we introduce the Pauli matrices $\boldsymbol{\sigma} \equiv (\sigma_x, \sigma_y, \sigma_z)$ this matrix can be written in the particularly convenient form:

$$H(t) = \tfrac{1}{4}A\hbar^2[(-2\varrho - 1) + \boldsymbol{b}\cdot\boldsymbol{\sigma}] \qquad \text{(XVII.115)}$$

the vector $\boldsymbol{b}$ having the components

$$b_x = 2\sqrt{2}, \qquad b_y = 0, \qquad b_z = 1 - 2\varrho.$$

We also introduce $\boldsymbol{u}$, the unit vector along $\boldsymbol{b}$:

$$\boldsymbol{b} = b\boldsymbol{u} \qquad b = \sqrt{8 + (1-2\varrho)^2}$$
$$\boldsymbol{u} \equiv \left(\frac{2\sqrt{2}}{b}, 0, \frac{1-2\varrho}{b}\right). \qquad \text{(XVII.116)}$$

Note that the vector $\boldsymbol{b}$ and the vector operator $\boldsymbol{\sigma}$ belong to a three-dimensional space having nothing whatever to do with ordinary three-dimensional space; we have merely employed a simple mathematical artifice that permits certain properties to be deduced by geometrical arguments analogous to those made in ordinary space.

From (XVII.115) and (XVII.116),

$$H(t) = \tfrac{1}{4}A\hbar^2[(-1 - 2\varrho) + b(\boldsymbol{\sigma}\cdot\boldsymbol{u})] \qquad \text{(XVII.117)}$$

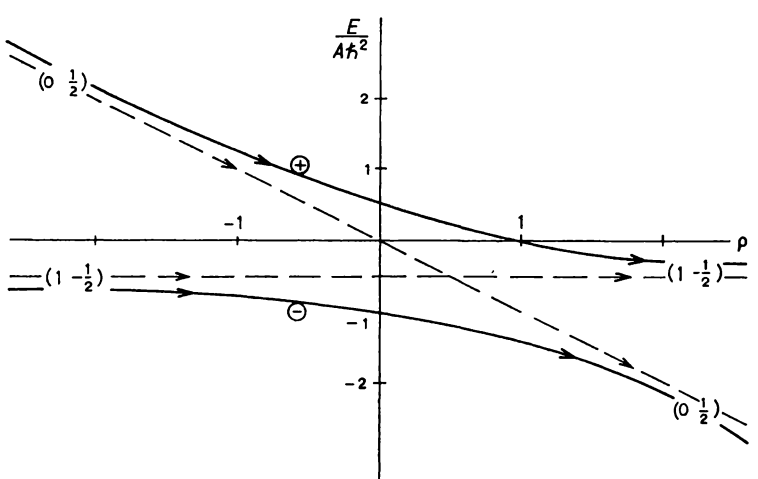


Fig. XVII.4. Evolution of the two ^{2}P levels with $M_J = \frac{1}{2}$ when the magnetic field is reversed (same notations as Figure XVII.3). The full line represents adiabatic reversal, the broken line rapid reversal.

$H(t)$ is therefore a function of the operator $(\boldsymbol{\sigma}\cdot\boldsymbol{u})$. It results that the corresponding eigenvalue problem is easily solved — we obtain the eigenvalues $\frac{1}{4}A\hbar^2(-1-2\varrho \pm b)$ and the projectors $P_\pm = \frac{1}{2}(1 \pm \boldsymbol{\sigma}\cdot\boldsymbol{u})$. We shall denote the corresponding eigenvectors by $|+\rangle$ and $|-\rangle$. They are defined to within a phase that may be fixed by the first of conditions (XVII.106), but since that phase does not come into what follows we need not be concerned with it here. It is easy to follow the continuous evolution of these two levels and their projectors as functions of the parameter $\varrho(t)$ (Figs. XVII.3 and XVII.4). When ϱ goes from $-\infty$ to ∞, the eigenvector $|+\rangle$ goes (except for a phase factor) from $|0\,\frac{1}{2}\rangle$ to $|1\,-\frac{1}{2}\rangle$, and the corresponding energy level follows the upper branch of the hyperbola in Figure 4; at the same time the eigenvector $|-\rangle$ goes from $|1\,-\frac{1}{2}\rangle$ into $|0\,\frac{1}{2}\rangle$, and the corresponding level follows the lower branch of the hyperbola.

Suppose, for example, that initially the system is in the state $|0\,\frac{1}{2}\rangle$. If the field is reversed slowly enough, the state vector of the system will always be equal to the vector $|+\rangle$ (to within a phase factor) and the system will effectively be in the state $|1\,-\frac{1}{2}\rangle$ once the field is reversed. The criterion for the validity of this adiabatic passage can be determined by applying the considerations of the preceding paragraph. Using the same notation we find, [relation (XVII.112)],

$$\eta_+ = p_{+\to-} \lesssim \max\left|\frac{\alpha(t)}{\omega(t)}\right|^2 \ll 1, \qquad \text{(XVII.118)}$$

where $\omega(t)$ is the Bohr frequency for the transition $+\to-$,

$$\omega(t) = \tfrac{1}{2}A\hbar b = \tfrac{1}{2}A\hbar\sqrt{8+(1-2\varrho)^2},$$

and $\alpha(t)$ is the projection onto $|-\rangle$ of the speed of $|+\rangle$, whence

$$|\alpha(t)| = |\langle -|(\mathrm{d}P_+/\mathrm{d}t)|+\rangle| = \tfrac{1}{2}|\langle -|(\boldsymbol{\sigma}\cdot\mathrm{d}\boldsymbol{u}/\mathrm{d}t|+\rangle|.$$

Since $\mathrm{d}\boldsymbol{u}/\mathrm{d}t$ is a vector perpendicular to $\boldsymbol{u}$ and since $|+\rangle$ and $|-\rangle$ are eigenvectors of $(\boldsymbol{\sigma}\cdot\boldsymbol{u})$, we have

$$|\alpha(t)| = \tfrac{1}{2}\left|\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}t}\right| = \frac{\sqrt{8}}{8+(1-2\varrho)^2}\left|\frac{\mathrm{d}\varrho}{\mathrm{d}t}\right|.$$

If the field is reversed according to the linear law (XVII.62), the maximum of $|\alpha/\omega|$ is attained when $\varrho = +\frac{1}{2}$, i.e.

$$\max\left|\frac{\alpha}{\omega}\right| = \tfrac{1}{2}\frac{\mu_B\mathscr{H}_0}{A\hbar^2} \times \frac{1}{A\hbar T}.$$

Condition (XVII.118) is thus realized if

$$T' \equiv \frac{2A\hbar^2}{\mu_B \mathscr{H}_0} T \gg (1/A\hbar). \qquad \text{(XVII.119)}$$

T' is the time necessary for the magnetic coupling energy $\mu_B \mathscr{H}$ to go from $-2A\hbar^2$ to $2A\hbar^2$. It is essentially during this period that $|+\rangle$ makes its rotation from the position $|0\ \tfrac{1}{2}\rangle$ to the position $|1\ -\tfrac{1}{2}\rangle$. Eq. (XVII.119) expresses that this period must be long as compared with $1/A\hbar$, period characteristic of the transition $+ \rightarrow -$.

It is interesting to compare this condition of adiabatic passage with the rapid-passage condition (XVII.64). The latter, in fact, is too restrictive. It is a necessary condition for the state-vector to remain practically unchanged during the whole time T that the field is being reversed, but, with the exception of the interval T' defined above, the eigenvectors of the Hamiltonian remain practically fixed during this time, and the state-vector of the system is simply multiplied by a phase factor. For the dynamical state itself to remain unchanged, that is, for the state vector to remain unchanged except for a *phase factor* during the reversal of the field, it is sufficient that the rapid passage condition be realized for the time T' during which the rotation of the eigenvectors of H takes place, i.e.

$$T' \ll \frac{1}{A\hbar} \qquad \text{(XVII.120)}$$

(cf. Problem XVII.8).

EXERCISES AND PROBLEMS

1. Let u_1 and u_2 be two orthogonal eigenstates corresponding to a doubly degenerate level of the Hamiltonian H_0 of a system. The introduction of a constant perturbation V removes the degeneracy and splits the level into two levels a distance ε apart. Suppose that the system is initially in the state u_1 and that the perturbation V is introduced during a time T. If $W_{1\rightarrow 2}$ is the probability of finding the system in the state u_2 after the perturbation has been "turned off", show that $W_{1\rightarrow 2}$ is a periodic function of T with angular frequency $\varepsilon/\hbar$ and verify that in the limit when $\varepsilon T \ll \hbar$ we obtain the result given by the first-order perturbation theory. What is necessary in order that $W_{1\rightarrow 2}$ vanish whatever T?

2. A hydrogen atom is subject to an oscillating electric field $\mathscr{E} = \mathscr{E}_0 \cos \omega t$ whose circular frequency ω is greater than its ionization frequency $me^4/2\hbar^3$. If the atom is initially in its ground state, what is the probability per unit time of a transition to an ionized state (suppose that we may use plane waves to

represent ionized states)? What is the angular distribution of the electron emitted in this exitation process? [N.B. The process described here is that of the photoelectric effect for which one thus obtains a semi-classical treatment in which the electromagnetic field is not quantized. The results are the same as those given by the correct treatment with the quantized electromagnetic field (cf. Problem XXI.12).]

3. In β decay an atomic nucleus emits an electron with a velocity that is usually close to c, and its charge goes from Ze to $(Z+1)e$. Show that the effect of this transition on the other electrons may be treated in the sudden approximation. Verify that the method applies to the disintegration of triton H^3 ($\equiv$ 1 proton + 2 neutrons) into He^3 ($\equiv$ 1 neutron + 2 protons), where the average kinetic energy of the disintegration electron is only 16 keV (mc^2 = 500 keV). The triton atom is initially in its ground state; what is the probability of finding, after the decay, an He^+ ion in the 1s state? In the 2s state? In a state $l \neq 0$?

4. Let $H(t)$ be the Hamiltonian of a non-conservative system. We suppose that there exists a time-independent vector $|u\rangle$ satisfying the equation $H(t)\,|u\rangle = \varepsilon(t)\,|u\rangle$. Show that the vector

$$\exp\left(-\mathrm{i}\int_{t_0}^{t}\varepsilon(\tau)\,\mathrm{d}\tau/\hbar\right)|u\rangle$$

satisfies the Schrödinger equation for the system.

5. Let $P_1, P_2, \ldots, P_j, \ldots$ be a complete set of orthogonal projectors. We suppose each of them to be a continuous, differentiable function of a parameter s, and that the variations with respect to s conserve the orthogonality and closure relations:

$$P_j P_k = \delta_{jk} P_k \qquad \sum_j P_j = 1.$$

Show that the operator

$$K(s) \equiv \mathrm{i}\hbar \sum_j (\mathrm{d}P_j/\mathrm{d}s)\,P_j \equiv -\mathrm{i}\hbar \sum_j P_j(\mathrm{d}P_j/\mathrm{d}s)$$

is Hermitean and that it obeys the commutation relations (XVII.75). Show also that it obeys the identities:

$$P_j K P_j = 0 \qquad (j = 1, 2, \ldots)$$
$$P_j K P_k = \mathrm{i}\hbar P_j(\mathrm{d}P_k/\mathrm{d}s)\,P_k = -\mathrm{i}\hbar P_j(\mathrm{d}P_j/\mathrm{d}s)\,P_k.$$

6. The $P_j(s)$ ($j = 1, 2 \ldots$) being defined as in the preceding problem, show that the derivative of the operator

$$H(s) = \sum_j \varepsilon_j(s)\,P_j(s),$$

where the $\varepsilon_j(s)$ are differentiable functions of s, obeys the equation

$$P_j(\mathrm{d}H/\mathrm{d}s)\,P_k = (\varepsilon_k - \varepsilon_j)\,P_j(\mathrm{d}P_k/\mathrm{d}s)\,P_k + \delta_{jk}(\mathrm{d}\varepsilon_k/\mathrm{d}s)\,P_k.$$

Deduce from this that the "angular velocity" $\alpha_{ji}(t)$ defined by equation (XVII.109) is also given by the relation:

$$\alpha_{ji}(t) = \frac{-1}{\hbar\omega_{ji}(t)}\, {}_t\langle j| \frac{\mathrm{d}H}{\mathrm{d}t} |i\rangle_t \qquad (j \neq i).$$

7. A uniform magnetic field $\mathscr{H}$ of fixed magnitude rotates with a constant angular velocity α about an axis making a given angle θ with it. Into this field is put an infinitely heavy particle of spin $\boldsymbol{J}$. We take $\hbar = 1$, we denote a unit vector parallel to the magnetic field by $\boldsymbol{u}(t)$ and we put $\gamma = \mu\mathscr{H}$, where μ is the gyromagnetic ratio of the particle. The motion of the spin $\boldsymbol{J}$ is therefore governed by the Hamiltonian $H(t) = -\gamma(\boldsymbol{J}\cdot\boldsymbol{u})$. Let $t = 0$ be the initial time, J_0 the component of $\boldsymbol{J}$ along $\boldsymbol{u}(0)$, J_z its component along the axis of rotation of the field.

Construct the unitary operator giving the "rotating axes", then show that the evolution operator in the Schrödinger representation is exactly given by the formula:

$$U(t) = \exp(-\mathrm{i}\alpha J_z t)\exp[\mathrm{i}(\gamma J_0 + \alpha J_z)t].$$

Verify the adiabatic theorem for this example; show by direct calculation and by the method of § 13 that the criterion for the validity of the adiabatic approximation is that $(\alpha \sin\theta/\gamma)^2 \ll 1$.

8. We consider the Schrödinger equation of the system defined in § 14 in the case when $M_J = \frac{1}{2}$. Let $\begin{pmatrix} u \\ v \end{pmatrix}$ be the components of its solution in the representation defined in that paragraph. We take $t_0 = -\frac{1}{2}T$ and put

$$T' = 2A\hbar^2 T/\mu_{\mathrm{B}}\mathscr{H}_0 \ (\ll T) \qquad \varkappa = 4/A\hbar T'.$$

Show that:

$$u = y\exp(-\mathrm{i}\varkappa\xi^2) \qquad v = \mathrm{i}(\mathrm{d}y/\mathrm{d}\xi)\exp(-\mathrm{i}\varkappa\xi^2),$$

where y, considered as a function of the variable $\xi = A\hbar t/\sqrt{2}$, obeys the equation

$$y'' - \tfrac{1}{2}\mathrm{i}(\sqrt{2} + 4\varkappa\xi)\, y' + y = 0.$$

If we put $x = \xi + (\sqrt{2}/4\varkappa)$, the general solution to this equation can be written in the form

$$y = A_0 F\left(\frac{\mathrm{i}}{4\varkappa}\middle|\frac{1}{2}\middle|\mathrm{i}\varkappa x^2\right) + A_1 x F\left(\frac{1}{2} + \frac{\mathrm{i}}{4\varkappa}\middle|\frac{3}{2}\middle|\mathrm{i}\varkappa x^2\right).$$

The initial state being $|0\ \tfrac{1}{2}\rangle$, show, using the asymptotic form of the hypergeometric functions, that the probability ϖ for it to be in the same state at the end of a time T, is given by the formula [valid if $T \gg (T'/A\hbar)^{\frac{1}{2}}$]

$$\varpi = \left(\frac{1-\gamma}{1+\gamma}\right)^2, \qquad \gamma = \tanh\frac{\pi}{4\varkappa}.$$

Verify that conditions (XVII.119) and (XVII.120) do in fact correspond to adiabatic and rapid passage respectively.

THE VARIATIONAL METHOD AND ASSOCIATED PROBLEMS

1. The Ritz Variational Method

Apart from the WKB method, which can only be applied in a very narrow domain, there are two principal methods for the approximate determination of the energy levels and wave functions of the discrete spectrum: the perturbation method (Chapter XVI) and the variational method. The present chapter is devoted to the second of these methods.

The variational method is a very general one that can be used whenever the equations can be put into variational form. The principle involved is the following. The desired solutions belong to a certain function space $\mathscr{F}$. Let Ψ be an arbitrary function of that space. Let us suppose that the solutions of the equation being studied are the functions of $\mathscr{F}$ for which a certain functional $Q[\Psi]$ is stationary. That equation is therefore equivalent to the variational equation

$$\delta Q = 0. \tag{XVIII.1}$$

The Ritz variational method consists in seeking the solutions of equation (XVIII.1) among the functions of a space $\mathscr{F}'$ more restricted than $\mathscr{F}$.

Suppose, for example, that $\mathscr{F}$ is the set of all wave functions of a given quantum system. For $\mathscr{F}'$ we may take a subset of particular wave functions $\Phi(a, b, c)$ labelled with a certain number of continuous parameters $a, b, \ldots$, and spanning only a part of $\mathscr{F}$. Considered as a functional of the Φ, the quantity Q reduces to a simple function of the variational parameters $a, b, \ldots$,

$$q(a, b, \ldots) \equiv Q[\Phi(a, b, \ldots)].$$

Each set of values $a_0, b_0, \ldots$, for which this function is stationary, defines an approximate solution $\Phi_0(a_0, b_0, c_0,)$ of equation (XVIII.1).

The success of the method depends essentially on the choice of the trial-function space $\mathscr{F}'$. The trial function must be simple enough to lend itself easily to the calculation, but must vary in a sufficiently large or well-chosen domain for the solutions obtained to be close to the exact ones.

In practice the stationary values of Q have a definite physical significance. One of the chief virtues of the variational method is to automatically furnish a precise evaluation of these quantities. Clearly the difference between $Q[\Phi_0]$ and $Q[\Psi_0]$ is smaller the closer the approximate solution Φ_0 to the exact solution Ψ_0; but in addition, since $Q[\Psi]$ is stationary at the point $\Psi = \Psi_0$ this difference is an infinitesimal of higher order than the difference between Φ_0 and Ψ_0. Thus the variational method is especially useful for evaluating quantities that can be put in the form of stationary functionals. This is notably the case for the energy levels of bound states. We shall see in Chapter XIX that it can also be applied to the calculation of collision amplitudes.

The calculation of the levels of the discrete spectrum by the variational method is described in section 1 of this chapter. In the remaining two sections we treat two important problems by methods more or less directly related to the variational method: the determination of the wave functions of complex atoms in the central field approximation by the methods of Hartree and of Fock–Dirac (Section II) and the treatment of molecules in the Born–Oppenheimer adiabatic approximation (Section III).

I. VARIATIONAL METHOD FOR BOUND STATES

2. Variational Form of the Eigenvalue Problem

The stationary functional involved in the determination of bound states by the variational method is the average value of the energy. We shall prove the following theorem [1]):

THEOREM. *Let H denote the Hamiltonian of a quantum system and $E[\Psi]$ the average value of its energy*

$$E[\Psi] \equiv \frac{\langle \Psi | H | \Psi \rangle}{\langle \Psi | \Psi \rangle}. \tag{XVIII.2}$$

Any state-vector for which this average value, considered as a functional of the vectors of state-vector space, is stationary, is an eigenvector of the

---

[1]) This theorem is a general result relative to the discrete spectra of Hermitean operators in Hilbert space. Only the Hermitean character of H is involved in the following demonstration.

discrete spectrum of H, and conversely. The corresponding eigenvalue is the stationary value of the functional $E[\Psi]$.

It is to be noted that the vectors $|\Psi\rangle$ introduced here have a *finite norm*: the function space $\mathscr{F}$ (definition of § 1) is the *Hilbert space* of the dynamical states of the system. The theorem therefore states that the eigenfunctions of H in this Hilbert space are the solutions of the variational equation

$$\delta E = 0. \tag{XVIII.3}$$

We also note that the functional $E[\Psi]$ is independent both of the norm and of the phase of $|\Psi\rangle$. Therefore, any supplementary condition that may be imposed on these quantities will not effect the validity of the theorem. In particular, it is sometimes convenient to limit the domain of variation of $|\Psi\rangle$ to vectors of norm 1, as is done in some of the examples of this chapter.

Proof of the theorem

Calculating the variation of $E[\Psi]$, we get

$$\begin{aligned}\langle\Psi|\Psi\rangle\,\delta E &= \delta(\langle\Psi|H|\Psi\rangle) - E\delta(\langle\Psi|\Psi\rangle)\\ &= \langle\delta\Psi|(H-E)|\Psi\rangle + \langle\Psi|(H-E)|\delta\Psi\rangle.\end{aligned}$$

So long as $\langle\Psi|\Psi\rangle$ remains finite and non-null, equation (XVIII.3) is therefore equivalent to

$$\langle\delta\Psi|(H-E)|\Psi\rangle + \langle\Psi|(H-E)|\delta\Psi\rangle = 0. \tag{XVIII.4}$$

The ket $|\delta\Psi\rangle$ is the variation of $|\Psi\rangle$, the bra $\langle\delta\Psi|$ the variation of the bra conjugate to $|\Psi\rangle$. The variations $|\delta\Psi\rangle$ and $\langle\delta\Psi|$ are therefore not independent. They may, however, be treated as such, for, equation (XVIII.4) being satisfied by any infinitesimal ket $|\delta\Psi\rangle$, we may replace $|\delta\Psi\rangle$ by $\mathrm{i}|\delta\Psi\rangle$, thus obtaining a second equation

$$-\mathrm{i}\,\langle\delta\Psi|(H-E)|\Psi\rangle + \mathrm{i}\,\langle\Psi|(H-E)|\delta\Psi\rangle = 0, \tag{XVIII.4'}$$

and by taking linear combinations of (XVIII.4) and (XVIII.4') we can deduce the two equivalent equations:

$$\langle\delta\Psi|(H-E)|\Psi\rangle = 0, \qquad \langle\Psi|(H-E)|\delta\Psi\rangle = 0.$$

These are equivalent to the single equation (XVIII.4) if we agree

to consider the variations $|\delta\Psi\rangle$ and $\langle\delta\Psi|$ as arbitrary and independent [1]).

We then have the two relations:

$$(H-E)|\Psi\rangle = 0, \qquad \langle\Psi|(H-E) = 0,$$

i.e.

$$(H-E[\Psi])|\Psi\rangle = 0 \tag{XVIII.5a}$$

$$(H^\dagger - E^*[\Psi])|\Psi\rangle = 0. \tag{XVIII.5b}$$

Now H is Hermitean ($H=H^\dagger$), so that relations (XVIII.5*a*) and (XVIII.5*b*) are identical. Equation (XVIII.3) is therefore equivalent to equation (XVIII.5*a*): any ket $|\Psi_1\rangle$ for which E is stationary is an eigenket of H belonging to the eigenvalue $E[\Psi_1]$.

Conversely, let $|\Psi_1\rangle$ be an eigenket of finite norm and E_1 the corresponding eigenvalue:

$$H|\Psi_1\rangle = E_1|\Psi_1\rangle.$$

Scalar multiplication on the left by $\langle\Psi_1|$ gives

$$E_1 = E[\Psi_1].$$

Therefore $|\Psi_1\rangle$ satisfies equation (XVIII.5*a*), and also equation (XVIII.5*b*) since H is Hermitean and E_1 real. Consequently, $E[\Psi]$ is indeed stationary for $\Psi=\Psi_1$. Q.E.D.

The fundamental theorem is complemented by the following lemma:

LEMMA. *Whatever may be the dynamical state of the system, the average value of its energy is equal to or greater than the energy of the ground state*:

$$E[\Psi] \geqslant E_0. \tag{XVIII.6}$$

To prove this inequality it suffices to calculate the difference between the right- and left-hand sides in a representation where H is diagonal. To simplify the writing we suppose that the spectrum of H is entirely discrete. Let $E_0, E_1, \ldots, E_n, \ldots$, be the energy levels in increasing order and $P_0, P_1, \ldots, P_n, \ldots$ the projectors onto their respective subspaces. Using the decomposition of unity, we find

$$E[\Psi] - E_0 \equiv \frac{\langle\Psi|(H-E_0)|\Psi\rangle}{\langle\Psi|\Psi\rangle} = \sum_{n=1}^{\infty} (E_n - E_0)\frac{\langle\Psi|P_n|\Psi\rangle}{\langle\Psi|\Psi\rangle}.$$

[1]) This is a general rule. It is due to the fact that (XVIII.4) is a linear relation between $|\delta\Psi\rangle$ and $\langle\delta\Psi|$ while the correspondence between bras and kets is antilinear.

Since each term in this sum is either positive or null, so is the sum itself, proving (XVIII.6).

3. Variational Calculation of Discrete Levels

It was seen in the general discussion of § 1 that one can obtain approximate solutions to the variational equation (XVIII.3) by restricting the domain of variation of the vectors $|\Psi\rangle$ to only a part of state-vector space. If this more restricted domain $\mathscr{F}'$ is well chosen, we thereby obtain certain eigenvectors of H to a good approximation, and the corresponding energy eigenvalues to a still better approximation.

The method takes a particularly simple form *when the trial function depends linearly on the variational parameters*, i.e. when $\mathscr{F}'$ is also a vector space. $\mathscr{F}'$ is then a subspace of $\mathscr{F}$, the term subspace having its usual meaning (§ VII.2).

We denote by P the projector onto $\mathscr{F}'$, by Φ an arbitrary vector of $\mathscr{F}'$ and by H_P the projection of the Hamiltonian onto $\mathscr{F}'$:

$$H_P \equiv PHP. \tag{XVIII.7}$$

$E[\Phi]$ [definition (XVIII.2)] is equal to the average value of H_P. However, H_P is Hermitean and transforms the vectors of $\mathscr{F}'$ linearly into vectors of $\mathscr{F}'$; thus it can be treated as a Hermitean operator of $\mathscr{F}'$ to which the fundamental theorem of § 2 is applicable. The variational equation

$$\delta E[\Phi] = 0 \tag{XVIII.8}$$

is therefore equivalent to the eigenvalue equation

$$H_P \Phi = E\Phi. \tag{XVIII.9}$$

Thus the variational approximation here consists in replacing the eigenvalue problem for H by a problem of the same type but *a priori* easier to solve since defined in a more restricted space.

Note the analogy with the perturbation method (§ XVI.8). In particular, if $\mathscr{F}'$ is the subspace of a given eigenvalue of the unperturbed Hamiltonian, the levels given by the variational method are identical with those given by a first-order perturbation calculation.

4. A Simple Example: The Hydrogen Atom

In order to gain some familiarity with the method it will now be applied to the calculation of the ground state of the hydrogen atom and the results compared with the exact ones of Chapter XI.

We put:

$$a_0 = \hbar^2/me^2 \qquad E_H = \tfrac{1}{2}(e^2/\hbar c)^2 mc^2 \qquad \varrho = r/a_0.$$

Since we are looking for eigenstates of well-defined angular momentum (lm), we take trial functions of the form

$$\Phi = a_0^{-3/2} \frac{u(\varrho)}{\varrho} Y_l^m(\theta, \varphi).$$

A simple calculation gives

$$E[\Phi] = -E_H \frac{\int_0^\infty u^* \left(\frac{d^2}{d\varrho^2} - \frac{l(l+1)}{\varrho^2} + \frac{2}{\varrho} \right) u \, d\varrho}{\int_0^\infty |u|^2 \, d\varrho}.$$

We limit ourselves to s-states $(l = m = 0)$ and calculate the stationary values of the energy for the three different trial functions;

$$u_1 = \varrho e^{-b\varrho}, \qquad u_2 = \frac{\varrho}{b^2 + \varrho^2}, \qquad u_3 = \varrho^2 e^{-b\varrho}.$$

Each of these trial functions depends on only the one parameter b. In each case, therefore, $E[\Phi]$ reduces to a function of b and the variational calculation consists in finding the minimum of that function. The calculations are not difficult and the results are given in table XVIII.1. First are listed, as functions of b, the analytical expressions for the norm of the trial function $N^2 = \langle \Phi | \Phi \rangle$, and for the average value of the energy, then the position $b_{\min}$ of the minimum and its value E_{var}. It is interesting to compare E_{var} with the ground state energy $E_0 = -E_H$, and also the approximate solution Ψ_{var} with the ground state eigenfunction Ψ_0. For this purpose the normalized radial parts $(u/N)_{\text{var}}$ are given in table XVIII.1 and the corresponding curves shown in Figure XVIII.1; they are to be compared with the exact radial function $2\varrho e^{-\varrho}$. The table also gives the average value $\langle r \rangle_{\text{var}}$ corresponding to each of the three approximate solutions Ψ_{var}, as well as the quantity $\varepsilon \equiv 1 - |\langle \Psi_0 | \Psi_{\text{var}} \rangle|^2$ (Ψ_0 and Ψ_{var} being supposed normalized to unity); ε is a good measure of the difference

between Ψ_{var} and the ground state (it is the *square* of the modulus of the amplitude of the component of Ψ_{var} orthogonal to Ψ_0).

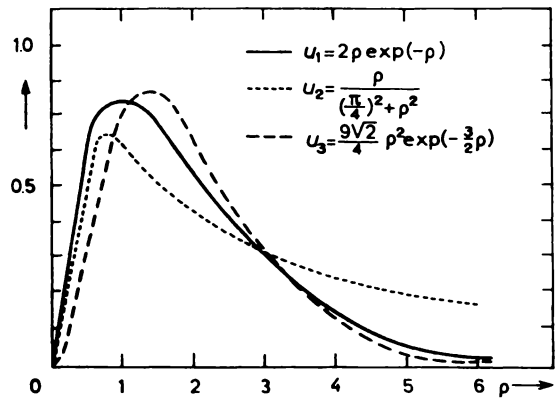


Fig. XVIII.1. Radial functions (normalized to unity) given by the variational calculation of § 4 for the ground state of the hydrogen atom.

All three trial functions have in common with the ground state wave function that they have no radial nodes. Therefore, they are expected to resemble more this function than those of the excited states, and the values E_{var} are expected to be closer to the ground state energy $-E_{\text{H}}$ than to any of the other levels (first excited level: $E_1 = -\frac{1}{4}E_{\text{H}}$). To show that these predictions are indeed verified we have given under each value of E_{var} in Table XVIII.1 the corresponding value of the ratio $(E_{\text{var}} - E_0)/(E_1 - E_0)$ which is a good measure of the error involved in the variational calculation of the ground state energy. The best result is given by the trial function u_1, which gives the exact ground-state wave function and the exact eigenvalue: note that u_1 has the same behavior at the origin ($\propto \varrho$) and the same exponentially-decreasing asymptotic behavior as the s-state eigenfunctions (with an attractive potential other than the Coulomb potential we would not obtain the exact wave function but the result would still be excellent). The function u_2 has the proper behavior at the origin but an asymptotic behavior totally different from that of the exact solution; the value that it gives, $E_{\text{var}} = -0.81E_{\text{H}}$ is still very satisfactory. The function u_3, which on the contrary has a very inexact behavior at the origin, ($\propto \varrho^2$), but a correct exponential falling-

TABLE XVIII.1

Variational calculation of the ground state of the hydrogen atom

		1	2	3
$u(b, \varrho)$	=	$\varrho\, e^{-b\varrho}$	$\dfrac{\varrho}{b^2 + \varrho^2}$	$\varrho^2\, e^{-b\varrho}$
N^2	=	$1/4b^3$	$\pi/4b$	$3/4b^5$
$E(b)/E_H$	=	$b^2 - 2b$	$(\pi - 8b)/2\pi b^2$	$\frac{1}{3}b^2 - b$
b_{min}	=	1	$\frac{1}{4}\pi$	$\frac{3}{2}$
E_{var}	=	$-E_H$	$-0.81\, E_H$	$-0.75\, E_H$
$\dfrac{E_{var} - E_0}{E_1 - E_0}$	=	0	0.25	0.33
$(u/N)_{var}$	=	$2\varrho\, e^{-\varrho}$	$\varrho[(\frac{1}{4}\pi)^2 + \varrho^2]^{-1}$	$\frac{1}{4} \times 9\sqrt{2}\, \varrho^2\, e^{-3/2\varrho}$
$\langle r \rangle_{var}$	=	$1.5\, a_0$	∞	$1.66\, a_0$
$\varepsilon = 1 - \lvert\langle \Psi_0 \rvert \Psi_{var} \rangle\rvert^2$	=	0	0.21	0.05

off at infinity, gives a poorer result. Consideration of the quantities $\langle r \rangle_{var}$ and ε shows that u_3, on the whole, resembles more the exact solution than does u_2; the fact that it gives a poorer result underlines the importance of behavior at the origin for the calculation of the energy, importance essentially due to the attractive character of the potential.

5. Discussion. Application to the Calculation of Excited Levels

The variational method is a very powerful one, but it is difficult to assess the significance of the results.

There is no infallible method for knowing to what level it is giving an approximate value, nor, *a fortiori*, for estimating the error. However, it is often possible to give a reply to the first of these questions by comparing the general form of the Ψ_{var} obtained (number of nodes, behavior at the origin and at infinity) with that of the exact solution, or at least with our *a priori* knowledge of the exact solution. In practice, one chooses trial functions with a simple analytical form and a very limited number of oscillations (or nodes). They have therefore every chance to be close to the ground state.

The above considerations show that *the variational method is especially suitable for the calculation of the ground-state level*, for which it gives an upper limit [Lemma (XVIII.6)]. Unfortunately, there is no sure method for evaluating the order of magnitude of the error (cf. Problem XVIII.1). Everything depends on the choice and on the flexibility of the trial function, that is, on the choice and on the extension of the function space $\mathscr{F}'$.

The necessity for more complex trial functions, the difficulty in interpreting the results and particularly in knowing the size and sign of the error, make the application of the variational method to the calculation of excited levels more hazardous. There are, however, two situations in which it can be of use in such calculations.

In the first place, when the ground state wave function Ψ_0 is known, the trial function Φ can be chosen from among wave functions that are orthogonal to Ψ_0. In this case, the functional $E[\Phi]$ is at least equal to the energy E_1 of the first excited state

$$E[\Phi] \geqslant E_1 \tag{XVIII.10}$$

and the variational method gives an upper limit for E_1 (cf. Problem XVIII.2). When the exact ground state solution is unknown, it may happen that we have an approximate one, Φ_0 (determined, for example, by a variational calculation). In this case, we may apply the variational method to the calculation of E_1 starting from trial functions orthogonal to Φ_0 providing that the difference between Φ_0 and Ψ_0 is sufficiently small, i.e. if

$$\varepsilon_0 = 1 - |\langle \Psi_0 | \Phi_0 \rangle|^2 \ll 1$$

(Ψ_0 and Φ_0 being of norm 1). The stationary solution Φ_1 (which we suppose of norm 1) is no longer orthogonal to Ψ_0 and inequality (XVIII.10) may be violated, but necessarily (Problem XVIII.3)

$$|\langle \Phi_1 | \Psi_0 \rangle|^2 < \varepsilon_0 \tag{XVIII.11}$$

from which it follows that

$$E[\Phi_1] \geqslant E_1 - \varepsilon_0 (E_1 - E_0). \tag{XVIII.12}$$

The second favorable situation occurs when H has certain symmetry properties. Suppose, for example, that H is invariant under rotation. The eigenfunctions and eigenvalue can then be classified according

to the quantum numbers j, m. Let the functions of angular momentum (jm) span a space $\mathscr{E}(jm)$. Choosing a trial function in $\mathscr{E}$ we can make a variational calculation of the levels (jm) and notably of the lowest of them, for which the variational method automatically gives an upper limit (Problem XVIII.4).

6. Ground State of the Helium Atom

In this paragraph we shall apply the variational method to the calculation of the ground state energy of the He atom and, more generally, of $(Z-2)$-times ionized atoms like Li^+, Be^{++}, etc. The same problem has already been treated by the perturbation method in § XVI.4; unless otherwise specified we shall follow the same notation.

As trial function we take the function given by the zero order perturbation theory:

$$\Phi_a(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\pi a^3} e^{-(r_1+r_2)/a}$$

except that here a will be treated as a variational parameter rather than being given the value a_0/Z.

The trial function being of norm 1, the average value of the energy is given by

$$E(a) = \langle \varphi_a | H | \varphi_a \rangle = \iint \varphi_a^*(H\varphi_a)\, d\mathbf{r}_1\, d\mathbf{r}_2.$$

The Hamiltonian of the system can be put in the form

$$H = k_1 + k_2 + v_1 + v_2 + V_{12},$$

where $k_i (i = 1, 2)$ denotes the kinetic energy of the ith electron, $v_i = Ze^2/r_i$ its interaction with the nucleus, and V_{12} the interaction of the two electrons. $E(a)$ is therefore the sum of the average values of these five operators. The calculation of these quantities is greatly simplified by the observation that the wave function φ_a can be put in the form $\varphi_a = f_a(\mathbf{r}_1) f_a(\mathbf{r}_2)$, where $f_a(\mathbf{r})$ is the eigenfunction for the ground state of an electron in a Coulomb field of charge $Z'e$ with

$$Z' = a_0/a.$$

It follows that the total energy of such an electron is $-Z'^2 E_H$, the average value of its kinetic energy $+Z'^2 E_H$ and the average value

of its potential energy $-2Z'^2 E_H$ (Problem XI.1). Consequently

$$\langle \varphi_a | k_i | \varphi_a \rangle = Z'^2 E_H$$
$$\langle \varphi_a | v_i | \varphi_a \rangle = -2Z'^2 E_H(Z/Z') = -2ZZ' E_H.$$

The calculation of § XVI.4 (eq. XVI.17–20) is still valid and gives

$$\langle \varphi_a | V_{12} | \varphi_a \rangle = \tfrac{5}{4} Z' E_H,$$

whence

$$E(a) = 2E_H(Z'^2 - 2(Z - \tfrac{5}{16}) Z').$$

This expression, considered as a function of a, or, equivalently, as a function of Z', has a minimum for

$$Z' = Z - \tfrac{5}{16}. \qquad \text{(XVIII.13)}$$

The value at this minimum is

$$E_{\text{var}} = -2(Z - \tfrac{5}{16})^2 E_H. \qquad \text{(XVIII.14)}$$

The numerical values of E_{var} corresponding to the atoms He, Li^+ and Be^{++} are given in table 1 of § XVI.4 (column 5). It is interesting to compare these values with the value given by the first-order perturbation calculation. Note that

$$E_{\text{var}} = -2Z^2 E_H + \tfrac{5}{4} Z E_H - \tfrac{25}{128} E_H$$
$$= E_{\text{pert}} - \tfrac{25}{128} E_H,$$

and therefore that the value found is smaller than the value given by the perturbation calculation by the Z-independent quantity:

$$\tfrac{25}{128} E_H = 2.64 \text{ eV}.$$

It is, therefore, a better approximation, as one might have expected. E_{var} is nonetheless greater than the experimental value E_{exp}, in agreement with inequality (XVIII.6).

The eigenfunction given by this variational calculation has a simple physical interpretation. It represents two independent particles moving in a Coulomb field whose charge $Z'e$ is given by equation (XVIII.13); this charge is smaller than the charge of the nucleus by $\tfrac{5}{16} e$; the difference represents, in this approximation, the screening effect felt by each electron in its motion in the Coulomb field of the nucleus due to the presence of the other electron.

By increasing the flexibility of the trial function one obtains a value E_{var} still closer to the exact eigenvalue. In particular, rather

than having the trial function φ_a depending on just the one variational parameter a, we can take the product of φ_a by a polynomial of given degree in r_1, r_2, and r_{12} whose coefficients are also treated as variational parameters. The value found for E_{var} progressively decreases and approaches the exact value as the complexity of the polynomial is increased. Preceding in this way, Hylleraas has obtained a theoretical value in excellent agreement with experiment [1]).

II. THE HARTREE AND FOCK–DIRAC ATOMS

7. The Self-consistent Field Method

We have already given the broad outline of the quantum mechanical treatment of complex atoms. One starts from the independent particle approximation in which each electron moves independently of the others in a potential representing the attraction of the nucleus and the average repulsive effect of the other electrons. In this approximation one represents the atomic wave function by a certain Slater determinant which should be chosen as near as possible to the exact solution of the Schrödinger equation of the atom. The best choice is the function obtained by the variational method using an arbitrary Slater determinent Φ as trial function. This very important particular case of the variational method is called the self-consistent field method. Here we shall describe the main steps without going into the details of calculation [2]).

The application of this method is not limited to atoms. It can also be applied to the electrons in a molecule or in a solid and, more generally, to systems of identical particles in any external field whatever. Although atoms are specifically quoted throughout this section, the following treatment is readily applicable to these more general cases.

---

[1]) Hylleraas, Zeit. f. Phys. **65** (1930) 209: the details of Hylleraas' treatment are given in Condon and Shortly, *Theory of Atomic Spectra* (Cambridge, 4th ed., 1957), p. 345. Hylleraas' method converges very rapidly. With eight variational parameters the calculated value is slightly below the experimental one, in apparent contradiction with the inequality (XVIII.6). In fact E_{exp} is slightly above the eigenvalue E_0 of the ground state of H due to relativistic effects which can be estimated. The Hylleraas method is a (very precise) method for calculating E_0, and not E_{exp}.

[2]) For a detailed treatment of the method and its practical applications, cf. D. R. Hartree, *The Calculation of Atomic Structures* (Wiley, 1957). Cf. also Condon and Shortly, *loc. cit.*, note 1, this page.

8. Calculation of $E[\Phi]$

The Hamiltonian of the system of Z electrons is written

$$H = H_1 + H_2 \tag{XVIII.13}$$

$$H_1 = \sum_{i=1}^{Z} h^{(i)}, \qquad h^{(i)} = \frac{\boldsymbol{p}^{(i)2}}{2m} + V(\boldsymbol{r}^{(i)}) \tag{XVIII.14}$$

$$H_2 = \sum_{i<j} w^{(ij)}. \tag{XVIII.15}$$

The first term, H_1, includes the kinetic energy and the potential energy due to the external field (the electric field of the nucleus); it is the sum of Z identical, individual Hamiltonians. The second term, H_2, is the interaction energy of the electrons, i.e. a sum of $\frac{1}{2}Z(Z-1)$ identical terms representing respectively the interaction of each pair of electrons; $w^{(ij)}$ is the potential between electron i and electron j. If we neglect the spin dependent forces, $w^{(ij)}$ is simply the repulsive electrostatic potential

$$w^{(ij)} = e^2/r_{ij} \qquad (r_{ij} \equiv |\boldsymbol{r}_i - \boldsymbol{r}_j|). \tag{XVIII.16}$$

The following treatment does not depend on this particular form for $w^{(ij)}$; it supposes simply that $w^{(ij)}$ is a function of the dynamical variables of the ith and jth electrons alone, and that this function is symmetrical in the permutation (ij).

As $E[\Phi]$ is independent of the normalization of the trial function Φ, the latter may always be normalized to unity. With the notation of Chapter XIV, we write it in the form

$$|\Phi\rangle \equiv (Z!)^{\frac{1}{2}} A|\hat{\Phi}\rangle, \tag{XVIII.17}$$

where A is the antisymmetrizer defined by equation (XIV.26):

$$A \equiv \frac{1}{Z!} \sum_P (-)^p P, \tag{XVIII.18}$$

and where $|\hat{\Phi}\rangle$ is the tensor product of Z arbitrary, orthonormal, individual kets:

$$|\hat{\Phi}\rangle \equiv |\alpha\rangle^{(1)}|\beta\rangle^{(2)} \ldots |\zeta\rangle^{(Z)} \equiv |\alpha^{(1)}\beta^{(2)} \ldots \zeta^{(Z)}\rangle \tag{XVIII.19}$$

$$\langle\lambda|\mu\rangle \equiv \delta_{\lambda\mu} \qquad (\lambda, \mu = \alpha, \beta, \ldots, \zeta). \tag{XVIII.20}$$

The normalization condition is then automatically satisfied:

$$\langle\Phi|\Phi\rangle = 1. \tag{XVIII.21}$$

$E[\Phi]$ is the sum of the average values of H_1 and H_2. Their calculation is made easier by the fact that H_1 and H_2, being invariant under permutation, commute with A, and that A is a projector ($A^2 = A$).

The average value of H_1 becomes successively:

$$\begin{aligned}\langle H_1\rangle \equiv \langle\Phi|H_1|\Phi\rangle &= Z!\,\langle\hat{\Phi}|H_1 A|\hat{\Phi}\rangle \\ &= \sum_{i=1}^{Z}\sum_{P}(-)^p\,\langle\hat{\Phi}|h^{(i)}P|\hat{\Phi}\rangle \\ &= \sum_{i=1}^{Z}\langle\hat{\Phi}|h^{(i)}|\hat{\Phi}\rangle.\end{aligned}$$

Replacing $|\hat{\Phi}\rangle$ by its definition, (XVIII.19), the last line becomes:

$$\langle H_1\rangle = \textstyle\sum_\lambda \langle\lambda|h|\lambda\rangle \qquad (\lambda = \alpha, \beta, \ldots, \zeta). \qquad \text{(XVIII.22)}$$

Thus $\langle H_1\rangle$ is just the sum of the average values of the individual Hamiltonian h relative to the Z individual quantum states occupied by the electrons.

Similarly, $\langle H_2\rangle$ can be put into the form of a sum of matrix elements of the operator w between two-electron states. We have, successively:

$$\begin{aligned}\langle H_2\rangle \equiv \langle\Phi|H_2|\Phi\rangle &= Z!\,\langle\hat{\Phi}|H_2 A|\hat{\Phi}\rangle \\ &= \sum_{i<j}\sum_{P}(-)^p\,\langle\hat{\Phi}|w^{(ij)}P|\hat{\Phi}\rangle \\ &= \sum_{i<j}\langle\hat{\Phi}|w^{(ij)}\,(1-P_{(ij)})|\hat{\Phi}\rangle\end{aligned}$$

i.e.

$$\langle H_2\rangle = \textstyle\sum'(\langle\lambda^{(1)}\mu^{(2)}|w^{(12)}|\lambda^{(1)}\mu^{(2)}\rangle - \langle\lambda^{(1)}\mu^{(2)}|w^{(12)}|\mu^{(1)}\lambda^{(2)}\rangle) \quad \text{(XVIII.23)}$$

the sum $\sum'$ being extended over the $\frac{1}{2}Z(Z-1)$ pairs of individual states λ, μ that can be formed from the Z individual states $\alpha, \beta, \ldots, \zeta$. The first term in the brackets is the average value of the interaction energy relative to the ket $|\lambda^{(1)}\,\mu^{(2)}\rangle$ for which electron 1 is in the state λ and electron 2 in the state μ; the second term is the *exchange term*, that is, the matrix element of w between the states $|\lambda^{(1)}\,\mu^{(2)}\rangle$ and $|\mu^{(1)}\,\lambda^{(2)}\rangle$ [Note that the exchange term is real and that we have: $\langle\lambda^{(1)}\mu^{(2)}|w^{(12)}|\mu^{(1)}\lambda^{(2)}\rangle = \langle\mu^{(1)}\lambda^{(2)}|w^{(12)}|\lambda^{(1)}\mu^{(2)}\rangle$. This property is due to $w^{(12)}$ being Hermitean and invariant in the permutation (1 2)]. $\langle H_2\rangle$ can also be written

$$\langle H_2\rangle = \tfrac{1}{2}\textstyle\sum_\lambda\sum_\mu(\langle\lambda^{(1)}\mu^{(2)}|w^{(12)}|\lambda^{(1)}\mu^{(2)}\rangle - \langle\lambda^{(1)}\mu^{(2)}|w^{(12)}|\mu^{(1)}\lambda^{(2)}\rangle) \quad \text{(XVIII.24)}$$
$$(\lambda, \mu = \alpha, \beta, \ldots, \zeta).$$

And finally

$$E[\Phi] = \langle H_1 \rangle + \langle H_2 \rangle, \tag{XVIII.25}$$

where $\langle H_1 \rangle$ and $\langle H_2 \rangle$ are given by equations (XVIII.22) and (XVIII.24).

9. The Fock-Dirac equations

The variational solution of the Schrödinger equation in the self-consistent field approximation is the solution that makes $E[\Phi]$ stationary with respect to the Z orthonormal vectors $|\lambda\rangle$ ($\lambda = \alpha, \beta, \ldots, \zeta$). To say that E is stationary with respect to variations of Z vectors obeying the Z^2 conditions (XVIII.20) is equivalent to saying (method of Lagrange multipliers) that there exist Z^2 constants $\varepsilon_{\lambda\mu}$ ($\lambda, \mu = \alpha, \beta, \ldots, \zeta$) such that the variational equation

$$\delta E - \sum_{\lambda} \sum_{\mu} \varepsilon_{\lambda\mu} \, \delta\langle\mu|\lambda\rangle = 0 \tag{XVIII.26}$$

is satisfied.

The Z^2 constants $\varepsilon_{\lambda\mu}$ can be regarded as the elements of a certain $Z \times Z$ matrix, ε. ε is a Hermitean matrix, for, E being real, δE is real too, and we can subtract equation (XVIII.26) from its complex conjugate to get

$$\textstyle\sum_\lambda \sum_\mu (\varepsilon_{\lambda\mu} - \varepsilon^*_{\mu\lambda}) \, \delta\langle\mu|\lambda\rangle = 0,$$

which requires that

$$\varepsilon_{\lambda\mu} = \varepsilon^*_{\mu\lambda}.$$

The Z vectors $|\alpha\rangle, |\beta\rangle, \ldots, |\zeta\rangle$ constitute an orthonormal basis in a certain subspace $\mathscr{E}_\Phi$ of the space of the individual states. A change of basis in this subspace results in $|\Phi\rangle$ being multiplied by a phase factor. To see this, denote by S the $Z \times Z$ unitary matrix defining the change of basis and by $|\alpha'\rangle, |\beta'\rangle, \ldots, |S'\rangle$ the new basis vectors:

$$|\lambda'\rangle = \textstyle\sum_\lambda |\lambda\rangle \, S_{\lambda\lambda'}.$$

According to a well-known property of the products of determinants, the Slater determinant of these Z new vectors is equal to the product of that of the Z old ones by det S. Consequently,

$$|\Phi'\rangle = (\det S)\,|\Phi\rangle.$$

Since S is unitary $|\det S| = 1$. It follows that the functional $E[\Phi]$ is invariant under a change of basis and the variational equation (XVIII.26) defines the sequence $|\alpha\rangle, |\beta\rangle, \ldots, |\zeta\rangle$ to within such a change of basis.

It is easy to show from (XVIII.26) that the analogous equation

$$\delta E - \sum_{\lambda} \sum_{\mu} \varepsilon'_{\lambda\mu} \, \delta \langle \mu' | \lambda' \rangle = 0$$

is also true, where the matrix ε' is the transform of ε in the transformation $S^{\dagger}$:

$$\varepsilon'_{\lambda\mu} = (S^{\dagger} \varepsilon S)_{\lambda\mu}.$$

In particular, S can be chosen to make ε' diagonal. Since the variational problem is not modified by limiting the choice of basis, we shall henceforth require that ε be diagonal. The variational equation (XVIII.26) then becomes

$$\delta E - \sum_{\lambda} e_{\lambda} \, \delta \langle \lambda | \lambda \rangle = 0. \qquad \text{(XVIII.26')}$$

Returning to equations (XVIII.22), (XVIII.24) and (XVIII.25) we easily calculate δE, and the left-hand side of (XVIII.26) then takes the form of a homogeneous linear combination of the $2Z$ variations $\langle \delta\lambda |$ and $|\delta\lambda\rangle$ ($\lambda = \alpha, \beta, \ldots, \zeta$). Writing that it vanishes whatever these variations, considered as being independent [cf. note, p. 765] and making use of the fact that H is Hermitean, one obtains – the details will not be given here – a set of Z equations for the Z orthogonal vectors

$$|\alpha\rangle, |\beta\rangle, \ldots, |\zeta\rangle,$$

namely the equations

$$h^{(1)}|\lambda\rangle^{(1)} + \textstyle\sum_{\mu}{}^{(2)}\langle\mu|w^{(12)}|\mu\rangle^{(2)}|\lambda\rangle^{(1)} - \sum_{\mu}{}^{(2)}\langle\mu|w^{(12)}|\lambda\rangle^{(2)}|\mu\rangle^{(1)} = e_{\lambda}|\lambda\rangle^{(1)} \quad (\lambda = \alpha, \beta, \ldots, \zeta). \qquad \text{(I)}$$

Note the absence of factors $\frac{1}{2}$ in the last two terms on the left-hand side. Scalar multiplication of both sides by ${}^{(1)}\langle\lambda|$ and summation over λ gives

$$\textstyle\sum_{\lambda} e_{\lambda} = \langle H_1 \rangle + 2\langle H_2 \rangle = E[\Phi] + \langle H_2 \rangle. \qquad \text{(XVIII.27)}$$

This relation will be discussed further on.

In practice, one represents the individual kets by their wave functions

$$u_\lambda(q) \equiv \langle q|\lambda\rangle,$$

where $q \equiv (\mathbf{r}, m_s)$ denotes the space and spin coordinates.

It is convenient to introduce the "electron density":

$$\varrho(q, q') \equiv \langle q|\varrho|q'\rangle = \sum_\mu u_\mu(q)\, u_\mu^*(q'). \qquad \text{(XVIII.28)}$$

It is the matrix representative of the projector onto the space $\mathscr{E}_\Phi$ defined above:

$$\varrho = \sum_\mu |\mu\rangle\langle\mu|.$$

The diagonal element

$$\varrho(q) \equiv \varrho(q, q) = \sum_\mu |u_\mu(q)|^2 \qquad \text{(XVIII.29)}$$

is the probability density for finding an electron at the point q.

The interaction $w^{(ij)}$ is a certain real, symmetrical function of the variables $q^{(i)}$, $q^{(j)}$ which will henceforth be denoted $w(q^{(i)}, q^{(j)})$. Put:

$$W_{\text{exc}}(q, q') = \varrho(q, q')\, w(q, q') \qquad \text{(XVIII.30)}$$

$$W(q) = \int \varrho(q')\, w(q, q')\, \mathrm{d}q' \qquad \text{(XVIII.31)}$$

$\int \mathrm{d}q'$ denoting, by convention, integration over the spatial coordinates and summation over the spin coordinate. With these notations, (I) becomes the set of integrodifferential equations:

$$\left[-\frac{\hbar^2}{2m}\Delta + V(q)\right] u_\lambda(q) + W(q)\, u_\lambda(q) - \int W_{\text{exc}}(q, q')\, u_\lambda(q')\, \mathrm{d}q' = e_\lambda\, u_\lambda(q) \qquad \text{(II)}$$

$$(\lambda = \alpha, \beta, \ldots, \zeta).$$

These are called *the Fock–Dirac integrodifferential equations.*

This set of equations can be solved by iteration. Starting from an approximate value ϱ_0 for the density, one obtains, by substitution in (XVIII.30) and (XVIII.31), approximate values for W and W_{exc}. Substituting these in (II) one obtains an eigenvalue equation whose first Z solutions $u_\alpha^{(1)}, \ldots, u_\zeta^{(1)}$ lead to a new value ϱ_1 for the density. Starting from ϱ_1 and repeating these operations we get a new value ϱ_2, and so on. If it converges, the sequence $\varrho_0, \varrho_1, \varrho_2, \ldots$ tends to the exact solution ϱ. The criteria for convergence and the rate of convergence will not be discussed here. They obviously depend on the choice of ϱ_0.

10. Discussion

Each of equations (II) resembles a Schrödinger equation determining one of the Z individual states occupied by the Z electrons of the atom. These are not, however, true eigenvalue equations, since operators W and W_{exc} depend on the electron density and consequently the eigensolutions $u_\alpha, u_\beta, \ldots, u_\zeta$ enter into the definition of the Hamiltonian. Nevertheless, it is interesting to examine this individual Hamiltonian, and to try to give a physical significance to its various terms.

For this purpose, denote the density of electrons in the $(Z-1)$ occupied states other than the state λ by $\varrho^{(\lambda)}$:

$$\varrho^{(\lambda)} \equiv \sum_{\mu \neq \lambda} |\mu\rangle \langle\mu| \equiv \varrho - |\lambda\rangle \langle\lambda|, \qquad \varrho^{(\lambda)}(q, q') \equiv \varrho(q, q') - u_\lambda(q)\, u_\lambda^*(q')$$

and the expressions obtained by replacing ϱ by $\varrho^{(\lambda)}$ in definitions (XVIII.30) and (XVIII.31) by $W_{\text{exc}}^{(\lambda)}$ and $W^{(\lambda)}$ respectively. We also introduce the average potential $X^{(\lambda)}$ created by the electron in the state $|\lambda\rangle$:

$$X^{(\lambda)}(q) = \int |u_\lambda(q')|^2\, w(q, q')\, \mathrm{d}q'. \qquad \text{(XVIII.32)}$$

$W^{(\lambda)}$ is the average potential created by the electrons situated in the $(Z-1)$ other occupied states, while W is the average potential created by all of the electrons:

$$W(q) = W^{(\lambda)}(q) + X^{(\lambda)}(q).$$

The equation of system (II) relating to the state λ can now be put in the form

$$\left[-\frac{\hbar^2}{2m}\Delta + V(q)\right] u_\lambda(q) + W^{(\lambda)}(q)\, u_\lambda(q) - \int W_{\text{exc}}^{(\lambda)}(q, q')\, u_\lambda(q')\, \mathrm{d}q' = e_\lambda\, u_\lambda(q) \qquad \text{(III)}$$

$$(\lambda = \alpha, \beta, \ldots, \zeta)$$

since $W^{(\lambda)}(q)\, u_\lambda(q)$ differs from $W(q)\, u_\lambda(q)$ by the "self-energy term" $X^{(\lambda)}(q)\, u_\lambda(q)$, and further,

$$\int W_{\text{exc}}^{(\lambda)}(q, q')\, u_\lambda(q')\, \mathrm{d}q' = \int W_{\text{exc}}(q, q')\, u_\lambda(q')\, \mathrm{d}q' - X^{(\lambda)}(q)\, u_\lambda(q).$$

In the form (III), the "Schrödinger equation" for the electron in the state λ is easily interpreted. The Hamiltonian represents the energy of the electron in the field consisting of the field of the nucleus and the

average field of the other electrons. It is made up of four terms, the kinetic energy $-\hbar^2\Delta/2m$, the nuclear potential $V(q)$, the average potential $W^{(\lambda)}(q)$ from the presence of the $(Z-1)$ other electrons, and a fourth term, also due to the other electrons, representing the *exchange effect* between the state λ and the $(Z-1)$ other occupied states. We see that the exchange effect leads to a non-local potential defined by the kernel $W^{(\lambda)}_{\text{exc}}(q, q')$.

This interpretation suggests that the eigenvalue e_λ represents the energy of the electron in the state λ. The Z quantities $e_\alpha, e_\beta, \ldots, e_\zeta$ given by the Fock–Dirac equations do in fact represent to a good approximation the *ionization energy* of the Z electrons of the atom. It is to be observed, however, that the sum of these individual energies does not represent the total energy of the system of Z electrons. In summing the individual energies, we count once their kinetic energy, once their energy of interaction with the nucleus, but twice their energy of mutual interaction; to obtain the total energy, we must therefore subtract the average value of the mutual-interaction energy, that is, the quantity $\langle H_2 \rangle$. This result has already been obtained above [eq. (XVIII.27)].

11. The Hartree Equations

If one neglects the exchange terms in system (III), one obtains a much simpler set of equations:

$$\left[-\frac{\hbar^2}{2m}\Delta + V(q) + W^{(\lambda)}(q)\right] u_\lambda(q) = e_\lambda\, u_\lambda(q) \qquad (\lambda = \alpha, \beta, \ldots, \zeta). \tag{IV}$$

These equations, proposed by Hartree in virtue of intuitive arguments alone, can also be obtained by the variational method if we take a simple product of individual kets like the $|\hat{\Phi}\rangle$ of equation (XVIII.19) as the trial function, rather than the antisymmetrized product $|\Phi\rangle$ of equation (XVIII.17) [1].

The Hartree system, like the Fock–Dirac, can be solved by iteration. Owing to the absence of exchange terms, the calculations are considerably shorter. However, this system is less symmetrical than the

[1] This does not constitute a valid justification of these equations, since unlike the kets $|\Phi\rangle$, the $|\hat{\Phi}\rangle$ do not belong to the space of the ket vectors of the system.

Fock–Dirac since the Hartree Hamiltonian, $h + W^{(\lambda)}$, is not the same for all of the individual states. This results in the eigensolutions of the Z Hartree equations not being mutually orthogonal, and leads to some difficulties in the application of the method. These difficulties will not be discussed here.

III. THE STRUCTURE OF MOLECULES

12. Generalities. Separation of the Electronic and Nuclear Motions

A molecule is a bound state of several atoms, and therefore consists of a certain number of atomic nuclei with electrons circulating about them. To find the stationary states of such a complex system is a very difficult problem. It has, however, a simplifying feature: *the mass of the electrons is much smaller than the mass of the nuclei* while the forces to which they are subjected are of comparable magnitude. As a consequence, the motion of the nuclei is very much slower than that of the electrons, and, to a very good approximation, these two motions, electronic and nuclear, can be treated separately. Indeed, in a first approximation the electrons "see" the nuclei as fixed force centers and their dynamical state is that of a system of electrons circulating about fixed nuclei. Since the latter move slowly, the dynamical state of the electrons adiabatically follows this gradual evolution in the potential to which they are subject (cf. Chapter XVII, section II). Conversely, since the electrons describe many revolutions during any appreciable displacement of the nuclei, the latter are essentially subject to only the average effect. The motion of nuclei is obtained to a very good approximation by replacing their interaction with electrons by its average value over several electron revolutions. Application of this procedure leads to a Schrödinger equation in which the electron variables have completely disappeared. The approximation upon which this method of separation of variables is based is called the *adiabatic approximation.*

The object of this section is to give an outline of the method and to discuss the conditions for its validity. We first complete the foregoing discussion with a semi-classical analysis of the nuclear motion, and by a rough evaluation of the different effects.

The potential in the Schrödinger equation for the nuclei depends only on the mutual distances of the latter. Since the molecule exists, this potential must have a minimum for certain well-defined, finite

values of the internuclear distances: this minimum corresponds to the point of stable equilibrium of the system about which the nuclei can make small oscillations. To these internal *vibrations* of the nuclei are superposed motions of *translation* and *rotation* of the system as a whole. The translational motion can be completely separated from the others by the introduction of the center of mass, which moves like a free particle whose mass equals the total mass of the system. In what follows we shall suppose this separation made, and consider only the vibrational and rotational motion of the nuclei.

Denote by m the mass of the electron, by M a mass of the order of magnitude of the nuclear masses, and by a the average mutual distance of the nuclei in the molecule. One has, roughly [1]:

$$a \simeq 10^{-8} \text{ cm} \qquad m/M \simeq 10^{-3}\text{–}10^{-5}. \tag{XVIII.33}$$

The linear dimensions of the molecule being of the order of a, this quantity gives the order of magnitude of the amplitude of the electronic motion. According to the uncertainty relations, the momentum of the electrons is of the order of $\hbar/a$, which corresponds to a kinetic energy of about $\hbar^2/ma^2$. This last mentioned quantity gives an order of magnitude for the binding energy of the ground state of the electrons and also an order of magnitude for the separation of the electronic levels:

$$\varepsilon_{\text{el}} \simeq \hbar^2/ma^2 \tag{XVIII.34}$$

(compare this evaluation with that of § XI.3).

As for the motion of the nuclei, we first consider their rotations. The moment of inertia of the system is of the order of Ma^2. Since the square of the angular momentum varies by quanta of the order of $\hbar^2$, the rotational energy will vary by quanta of the order of

$$\varepsilon_{\text{rot}} \simeq \hbar^2/Ma^2. \tag{XVIII.35}$$

In a first approximation, the vibrations can be considered as harmonic vibrations of quanta $\varepsilon_{\text{vib}} = \hbar\omega$. We take the zero of potential energy at the position of stable equilibrium of the nuclei. If one of the nuclei is removed a distance a, the system acquires a potential energy $\frac{1}{2}M\omega^2a^2$; but, since this amounts to completely separating one

[1] The mass ratio is the least favourable for the hydrogen molecule. In that case:

$$(m/M)_{H_2} \simeq 0.5 \times 10^{-3}.$$

of the atoms from the rest, and therefore to an increase in the energy of the order of ε_{el}, we have

$$M\omega^2 a^2 \simeq \hbar^2/ma^2$$

and therefore,

$$\varepsilon_{vib} \simeq \hbar^2/(mM)^{\frac{1}{2}}\, a^2. \tag{XVIII.36}$$

Comparison of (XVIII.34–36) leads to:

$$\varepsilon_{rot} \ll \varepsilon_{vib} \ll \varepsilon_{el}.$$

In a more precise way, if we follow Born and Oppenheimer [1]) and introduce the parameter

$$\varkappa = (m/M)^{\frac{1}{2}}, \tag{XVIII.37}$$

we find

$$\varepsilon_{rot} \simeq \varkappa^2\, \varepsilon_{vib} \simeq \varkappa^4\, \varepsilon_{el}. \tag{XVIII.38}$$

Now a level separation ε is associated with a classical frequency $\varepsilon/\hbar$. We can therefore conclude that the motion of the electrons is more rapid than the vibrational motion of the nuclei, itself more rapid than the rotational motion of the ensemble, in agreement with the remarks at the beginning of this paragraph. The ratio of the frequencies of these different motions is of the order of $\varkappa^2$, i.e. about 0.01: during a complete revolution of the whole molecule the nuclei make about 100 oscillations about their equilibrium positions and the electrons effect about 10 000 revolutions.

13. Motion of the Electrons in the Presence of Fixed Nuclei

Keeping in mind the semi-classical discussion of the preceding paragraph, we now consider the problem of finding the stationary states of the molecule. Since our purpose is to give more an outline than a detailed account, we shall not hesitate to make simplifying assumptions when they do not affect the principle of the method. In particular we shall treat the electrons and the nuclei as spinless particles.

We label the degrees of freedom of the electrons and of the nuclei with the indices i and j respectively. We denote by x_i and X_j the coordinates of the electrons and of the nuclei respectively and by

---

[1]) M. Born and J. R. Oppenheimer, Ann. der Phys. 84 (1927) 457; with regard to the original theory of Born and Oppenheimer, see § 16.

M_j the mass of the nucleus of coordinate X_j. We denote by T_e the kinetic energy of the electrons, by T_N that of the nuclei, and by V the interaction potential between the various particles in the molecule. The molecular Hamiltonian H is the sum of these three terms:

$$H = T_N + T_e + V. \tag{XVIII.39}$$

We have:

$$T_N = -\sum_j \frac{\hbar^2}{2M_j}\frac{\partial^2}{\partial X_j^2}, \qquad T_e = -\sum_i \frac{\hbar^2}{2m}\frac{\partial^2}{\partial x_i^2} \tag{XVIII.39'}$$

$V \equiv V(x, X)$ is a certain function of the electron and nuclear coordinates – it is the sum of the Coulomb interactions of each pair of particles in the system.

Consider the simpler Hamiltonian

$$H^{(0)} = T_e + V \tag{XVIII.40}$$

obtained by neglecting the kinetic energy of the nuclei in (XVIII.39). $H^{(0)}$ is the Hamiltonian of the system in the limit when $M_j \to \infty$, and its stationary states are those of a system of electrons in the presence of fixed nuclei. Indeed, since $H^{(0)}$ contains no derivatives with respect to the X_j,

$$[X_j, H^{(0)}] = 0,$$

and therefore $H^{(0)}$ and the X_j can simultaneously be diagonalized. In other words, to solve the eigenvalue problem for $H^{(0)}$, we can assign definite values X_j' to the nuclear position coordinates, and look for the eigenvectors of $H^{(0)}$ among the eigenvectors corresponding to these particular values of the X_j. Denote a given set of the X_j' by the symbol X'. To each set X' there corresponds a set of eigenvalues $W_n(X')$ of $H^{(0)}$ labelled by the quantum number n. For a given n and X' there will be one or several linear independent eigenvectors, which in the latter case we may distinguish one from another by an additional index, s say. Thus the Schrödinger equation for $H^{(0)}$ takes the form

$$H^{(0)}|nsX'\rangle = W_n(X')|nsX'\rangle. \tag{XVIII.41}$$

In the $\{xX\}$ representation, the eigenvector $|nsX'\rangle$ is represented by the wave function

$$\varphi_{ns}(x, X')\,\delta(X - X') \tag{XVIII.42}$$

and the function φ is the solution of the Schrödinger equation

$$[T_e + V(x, X')]\,\varphi_{ns}(x, X') = W_n(X')\,\varphi_{ns}(x, X'). \quad \text{(XVIII.43)}$$

In (XVIII.43), X' plays the role of a simple parameter, and we have the Schrödinger equation for the electrons of the molecule when the positions of the nuclei are fixed at X'. To each solution of this equation there corresponds an eigenfunction of $H^{(0)}$ of the form (XVIII.42), and all the solutions of (XVIII.43) for all possible values of X' constitute a complete set of eigenfunctions of $H^{(0)}$. Normalization being important in what follows, we shall always take orthonormal eigenfunctions, which will automatically be the case if $\varphi_{ns}(x, X')$, considered as a function of the x alone, has a norm equal to 1.

The eigenvalue problem (XVIII.43) is analogous to the problem of finding the stationary states of an atom and can be solved by analogous methods, that of self-consistent fields for example. In the present case however there are several centers of force, and the symmetry of atomic problems is thereby largely, if not totally, destroyed.

Let us briefly discuss this question of symmetry [1]). The potential $V(x, X)$ is invariant under translation, rotation and reflection of the whole system, and also under time reversal and permutation of the identical particles. Suppose that the positions of the nuclei are now fixed at X'. The symmetries of $V(x, X')$ considered as a function of the x alone will be those of the ones given above that leave the nuclear configuration X' unchanged. Thus in a diatomic molecule like ClH, $V(x, X')$ is invariant under rotation about the axis joining the hydrogen and chlorine nuclei, and under reflection in planes passing through that axis. With the invariance under time reversal these are the only remaining symmetries. In sufficiently complex molecules, only the invariance under time reversal remains. The symmetries of $H^{(0)}$, considered as a function of the x alone, are obviously the same as those of $V(x, X')$. The degeneracy of the level $W_n(X')$ is associated with these symmetry properties (cf. Chapter XV).

To avoid complications due to degeneracy, we shall assume that $V(x, X')$ has only the time-reversal invariance. Then, since we are

---

[1]) For a general treatment of the symmetry properties of electronic wave functions in molecules cf. L. D. Landau and E. M. Lifshitz, *Quantum Mechanics* (Pergamon Press, London, 1958).

dealing with a molecule (and not with a free radical) the number of electrons is even, and are in the first case discussed in § XV.21. Assuming further that there is no accidental degeneracy, we conclude that $W_n(X')$ is non-degenerate and we choose the phase so as to have $\varphi_n(x, X')$ real.

We conclude this paragraph with the following two remarks.

Considered as an operator acting on the variables x alone, $H^{(0)}$ is a continuous function of the parameters X', as are its eigenvalues, and also its eigenfunctions once their phases have been fixed, i.e. for a given n, $W_n(X')$ and $\varphi_n(x, X')$ are continuous functions of X'.

Considered as an operator acting on all of the variables [cf. eq. (XVIII.41)], $H^{(0)}$ is invariant with respect to the various transformations enumerated above in connection with $V(x, X)$. Therefore, any vector obtained from $|nX'\rangle$ by one of these transformations, is an eigenvector of $H^{(0)}$ belonging to the same eigenvalue $W_n(X')$. In other words, $W_n(X')$ is unchanged by the application of these transformations to the X'. In particular, $W_n(X')$ depends only on the mutual distances of the nuclei, i.e. on the geometrical figure that they form, and does not change when that figure is displaced (translation and rotation) or replaced by its image (reflection).

14. The Adiabatic Approximation

In the preceding paragraph we have considered the stationary states of the electrons when the nuclei are fixed. Suppose now that the nuclei are slowly moving according to a certain law $X'(t)$. If this motion is sufficiently slow, the dynamical state of the electrons will adiabatically follow the resulting modification of the potential in which they move: if at time t_0 they are in the state $(n, X'(t_0))$ corresponding to the level $W_n(X'(t_0))$ – i.e. the state represented by the wave function $\varphi_n(x, X'(t_0))$ – then at time t they will be in the state $(n, X'(t))$ obtained from $(n, X'(t_0))$ by continuity, n remaining fixed.

The conditions for the validity of this approximation were discussed in § XVII.13 [cf. criterion (XVII.114)]. The probability η_n of finding the electrons in a state different to (n, X') is given by

$$\eta_n \simeq \left|\frac{\alpha_n}{\omega_n}\right|^2, \qquad \text{(XVIII.44)}$$

where α_n is the "angular velocity" of the vector $\varphi_n(x, X')$ and ω_n

the minimum Bohr frequency associated with the level $W_n(X')$.

To evaluate this quantity we make use of the semi-classical analysis of § 12. The separation of the electronic levels being given by (XVIII.34), we have

$$\omega_n \simeq \varepsilon_{el}/\hbar \simeq \hbar/ma^2. \tag{XVIII.45}$$

To evaluate α_n, we first make an estimate of the norm of $\partial\varphi_n/\partial X_j'$, φ_n being by hypothesis a real function of norm 1. We know that to separate an atom from the rest of the molecule its nucleus must be removed to a distance of the order of a from its position of equilibrium; i.e. an increment of the order of $\Delta X_j' \simeq a$ is necessary to transform the function $\varphi_n(x, X')$ into a function that is orthogonal to it. It follows that $\partial\varphi_n/\partial X_j$ is roughly equal to this last-mentioned function divided by a, and its norm is of the order of $1/a^2$:

$$\int \left(\frac{\partial\varphi_n}{\partial X_j}\right)^2 \mathrm{d}x \simeq \frac{1}{a^2}. \tag{XVIII.46}$$

Now $\alpha_n{}^2$ is by definition the norm of the (real) function $\mathrm{d}\varphi_n/\mathrm{d}t$. If $v_j = \mathrm{d}X_j'/\mathrm{d}t$ is the velocity of X_j', $\alpha_n{}^2$ is the norm of the function $\sum_j v_j\,(\partial\varphi_n/\partial X_j)$, giving roughly

$$\alpha_n{}^2 \simeq \frac{1}{a^2}\sum_j v_j{}^2.$$

If we denote the kinetic energy of the nuclei by t_N, we have

$$\alpha_n{}^2 \simeq \frac{t_N}{Ma^2}. \tag{XVIII.47}$$

Substituting (XVIII.45) and (XVIII.47) into (XVIII.44) we find

$$\eta_n \simeq \varkappa^4 \frac{t_N}{\varepsilon_{el}}. \tag{XVIII.48}$$

This semi-classical analysis can be used as a guide for finding the stationary states of the molecule. Consider then the eigenvalue problem for the Hamiltonian H. From (XVIII.39) and (XVIII.40),

$$H = H^{(0)} + T_N. \tag{XVIII.49}$$

If the term for the kinetic energy of the nuclei, T_N, could be neglected, the molecular Hamiltonian would reduce to $H^{(0)}$, and each stationary

state $|nX'\rangle$ would correspond to a definite electronic quantum number n, and a definite arrangement of the nuclei X'. The nuclei would remain fixed at X' and the motion of the electrons would be represented by the function $\varphi_n(x, X')$. T_N couples the eigenvectors of $H^{(0)}$ corresponding to neighbouring X'.

In the adiabatic approximation, one neglects the coupling between vectors with different electronic quantum numbers, and one assumes n to be a good quantum number:

$$n \simeq Cst. \tag{XVIII.50}$$

The eigenvectors of H are then linear combinations of $|nX'\rangle$ corresponding to a definite value of n, hence of the form

$$\int |nX'\rangle \psi(X')\, dX', \tag{XVIII.51}$$

where $\psi(X')$ is an arbitrary function of the X'. In the $\{x, X\}$ representation, such vectors are represented by wave functions of the form

$$\Phi_n(x, X) \equiv \varphi_n(x, X)\, \psi(X). \tag{XVIII.52}$$

We can obtain the eigenfunctions of H in this approximation by *application of the variational method with $\Phi_n(x, X)$ as the trial function.* We have here a case where the trial function varies in a subspace of state-vector space, namely the space $\mathscr{E}_n$ of vectors of the form (XVIII.51). We know (§ 3) that the method then leads to an eigenvalue equation in this subspace [eq. (XVIII.9)]. In the present case, one obtains a Schrödinger equation for the unknown function $\psi(X)$. The Hamiltonian of this equation, which we denote by H_n, is the projection of H onto $\mathscr{E}_n$; it acts only on the dynamical variables of the nuclei. The eigenvalues of H_n are the molecular energy levels relative to the electronic quantum number n.

The error being made can be evaluated on the basis of the discussion of the beginning of this paragraph. Let Φ_n be a particular one of our approximate solutions, and Ψ_n the corresponding exact solution. They are both supposed normalized to unity. The difference

$$\delta\Psi \equiv \Psi_n - \Phi_n$$

represents the deviation of the exact solution from the adiabatic limit. It is a vector essentially outside the subspace $\mathscr{E}_n$, whose norm

is equal to the quantity η_n defined above. Thus, with the aid of (XVIII.48),

$$\iint |\delta\Psi|^2 \, dx \, dX \simeq \varkappa^4 \left\langle \frac{T_N}{\varepsilon_{el}} \right\rangle \simeq \varkappa^4 \frac{\varepsilon_{vib}}{\varepsilon_{el}} \simeq \varkappa^6.$$

In other words, we have roughly

$$\delta\Psi \simeq \varkappa^3 \Phi', \tag{XVIII.53}$$

where Φ' is a function orthogonal to the space $\mathscr{E}_n$ and of norm 1. Since the functional

$$E[\Psi] \equiv \frac{\langle \Psi | H | \Psi \rangle}{\langle \Psi | \Psi \rangle}$$

is stationary about the solution Ψ_n, the calculation of the energy with the function Φ_n involves an error of the second order with respect to the variation $\delta\Psi$, i.e.,

$$\begin{aligned} \delta E &\simeq \langle \delta\Psi | H | \delta\Psi \rangle = \varkappa^6 \langle \Phi' | H | \Phi' \rangle \\ &\simeq \varkappa^6 \, \varepsilon_{el}. \end{aligned}$$

Comparing with (XVIII.38),

$$\delta E / \varepsilon_{rot} \simeq \varkappa^2 \qquad (\simeq 10^{-2} \ll 1). \tag{XVIII.54}$$

Thus, the adiabatic approximation introduces an error of order $\varkappa^3$ in the determination of the molecular wave function [eq. (XVIII.53)] and an error in the determination of the energy that is $\varkappa^2$ times smaller than the separation of the rotational levels [eq. (XVIII.54)].

15. Hamiltonian for the Nuclei in the Adiabatic Approximation

By applying the variational method, we shall now determine the "Schrödinger equation" for the function $\psi(X)$.

We recall that $\varphi_n(x, X)$ is the real solution of equation (XVIII.43) whose norm with respect to integration over the x is equal to 1:

$$\varphi_n^*(x, X) = \varphi_n(x, X) \tag{XVIII.55}$$

$$\int \varphi_n^2(x, X) \, dx = 1. \tag{XVIII.56}$$

φ_n, being a continuous function of the X, is thus defined up to a sign, which may be arbitrarily fixed. The following equations, obtained

by differentiating (XVIII.56), will be of use in what follows:

$$\int \varphi_n \frac{\partial \varphi_n}{\partial X_j} \, dx = 0. \tag{XVIII.56'}$$

$$\int \varphi_n \frac{\partial^2 \varphi_n}{\partial X_j^2} \, dx = - \int \left(\frac{\partial \varphi_n}{\partial X_j} \right)^2 dx. \tag{XVIII.56''}$$

The function $\psi(X)$ is an arbitrary, square-integrable function of the coordinates X.

The domain of variation of the trial function being thus specified, we now propose to put the functional $E[\Phi_n]$ in the form of a functional of $\psi(X)$. From the properties of φ_n one easily obtains:

$$\langle \Phi_n | \Phi_n \rangle \equiv \iint |\Phi_n|^2 \, dx \, dX = \int |\psi|^2 \, dX. \tag{XVIII.57}$$

One may also write

$$\langle \Phi_n | H | \Phi_n \rangle \equiv \iint \Phi_n^* (H \Phi_n) \, dx \, dX = \int \psi^* (H_n \psi) \, dX \tag{XVIII.58}$$

if one agrees to put

$$H_n \psi \equiv \int \varphi_n(x, X) \, [H \varphi_n(x, X) \, \psi(X)] \, dx. \tag{XVIII.59}$$

Consequently,

$$E[\Phi_n] = \frac{\int \psi^* (H_n \psi) \, dX}{\int |\psi|^2 \, dX}. \tag{XVIII.60}$$

Identity (XVIII.59) is the defining relation for a certain operator H_n of $\psi(X)$ space. It is clearly a linear operator and also can be shown to be Hermitean (as will be verified below). In the adiabatic approximation, the eigenfunctions of the molecular Schrödinger equation are those for which $E[\Phi_n]$ is stationary with respect to variations of $\psi(X)$. According to (XVIII.60), these are the solutions of the eigenvalue equation

$$H_n \psi = E \psi. \tag{XVIII.61}$$

This is the required "Schrödinger equation".

The molecular energy levels corresponding to the electronic quantum number n are the eigenvalues of H_n. The corresponding eigenfunctions are the Φ_n obtained by substituting for $\psi(X)$ in (XVIII.52) the corresponding solution (or solutions) of the Schrödinger equation (XVIII.61).

Making use of the properties of the φ_n we can carry out many of the integrations in (XVIII.59) and obtain a more workable expression for H_n.

From (XVIII.43)

$$H^{(0)}\,\varphi_n(x, X)\,\psi(X) = W_n(X)\,\varphi_n(x, X)\,\psi(X);$$

from which, with the aid of (XVIII.56),

$$\int \varphi_n[H^{(0)}\,\varphi_n\,\psi]\,\mathrm{d}x = W_n(X)\,\psi(X). \qquad \text{(XVIII.62)}$$

From expression (XVIII.39′) for the kinetic energy of the nuclei one has

$$T_N\,\varphi_n\,\psi = \sum_j \left(\frac{-\hbar^2}{2M_j}\right)\left[\varphi_n \frac{\partial^2\psi}{\partial X_j^2} + 2\frac{\partial\varphi_n}{\partial X_j}\frac{\partial\psi}{\partial X_j} + \frac{\partial^2\varphi_n}{\partial X_j^2}\psi\right],$$

from which one may calculate

$$\int \varphi_n(T_N\,\varphi_n\,\psi)\,\mathrm{d}x$$

by multiplying both sides by $\varphi_n(x, X)$ and integrating over x. We separately consider the contributions from the three terms in the bracket on the right-hand side. Using (XVIII.56) we immediately have that the contribution of the first term is simply.

$$T_N\,\psi \equiv \sum_j \left(\frac{-\hbar^2}{2M_j}\right)\frac{\partial^2\psi}{\partial X_j^2}. \qquad \text{(XVIII.63)}$$

The contribution of the second term vanishes because of (XVIII.56′). That of the third term is the product of ψ by a certain function $W_n'(X)$ that with the aid of (XVIII.56″) can be written in the form

$$W_n'(X) = \sum_j \frac{\hbar^2}{2M_j}\int \left(\frac{\partial\varphi_n}{\partial X_j}\right)^2 \mathrm{d}x. \qquad \text{(XVIII.64)}$$

Thus, finally

$$\int \varphi_n(T_N\,\varphi_n\,\psi)\,\mathrm{d}x = (T_N + W_n')\,\psi. \qquad \text{(XVIII.65)}$$

From the definition (XVIII.59) and equations (XVIII.49), (XVIII.62) and (XVIII.65) we have

$$H_n = T_N + W_n' + W_n. \qquad \text{(XVIII.66)}$$

The three terms in formula (XVIII.66) are easily interpreted.

The potential energy $W_n(X)$ is the average of $H^{(0)}$, that is, it is the

sum of the mutual interaction energy of the nuclei and the average value of the energy of the electrons in the quantum state n corresponding to a determined position X of the nuclei. According to the discussion of § 12, $W_n(X)$ has an absolute minimum for a certain point X_0, the point of stable equilibrium for the figure formed by the nuclei.

The other two terms come from the kinetic energy of the nucleons averaged over the dynamical state of the electrons, which gives, in addition to the kinetic energy proper, T_N, a potential energy term $W_n'(X)$. The latter [1]) constitutes a small correction to the potential W_n; substituting the estimate (XVIII.46) into expression (XVIII.64), we find

$$W_n'(X_0) \simeq \frac{\hbar^2}{Ma^2}.$$

W_n' is therefore a positive quantity of the order of magnitude of the rotational quanta.

16. The Born-Oppenheimer Method

The original treatment of molecules, due to Born and Oppenheimer [2]) differs somewhat from the variational method described above. It consists in expanding the Hamiltonian H into a power series in $\varkappa$ and solving the corresponding eigenvalue problem by the usual methods of perturbation theory.

In the preceding paragraph we denoted the position of the minimum of $W_n(X)$ by X_0. This point of equilibrium is, in fact, defined only up to a rotation since W_n is invariant in an overall rotation of the nuclei (note in passing that in general W_n' is not). More precisely, let us put $X=(\omega, \xi)$, where ω denotes the three angular variables (two for diatomic molecules) fixing the overall orientation of the nuclei and ξ the set of radial variables fixing the positions of the nuclei relative to one another; W_n depends only on the ξ, and the equilibrium position corresponds to a certain set of values for the radial variables.

Following Born and Oppenheimer, one introduces the new radial variables u defined by

$$\xi = \xi_0 + \varkappa u.$$

The u represent, in convenient units, the distance of the nuclei from their equilibrium positions. Since $\varkappa$ is roughly equal to the

[1]) The appearance of this term is analogous to that of the centrifugal energy term in the equations of motion of a classical system when one adopts a rotating system of reference.

[2]) *Loc. cit.*, note, p. 783.

ratio of the amplitude of vibration of the nuclei to the amplitude of the motion of the electrons, the domain of variation of the u is of the same order of magnitude as that of the x, i.e. of the order of a.

Having made this change of variable, one obtains the expansion of H in powers of $\varkappa$ by expanding the potential in H in powers of u. The term T_N is of order $\varkappa^2$. The expansion must be pushed to the order of $\varkappa^4$ to accurately obtain the rotational levels. If it is stopped at order $\varkappa^5$ one obtains, to within infinitesimals of higher order, the results given by the adiabatic approximation. Differences appear only when we proceed to order $\varkappa^6$ and beyond, in agreement with the order-of-magnitude analysis of § 12.

17. Notions on Diatomic Molecules

We shall not proceed any further with the general study of molecules in the adiabatic approximation. In particular, we shall not discuss the separation of vibrations and rotations; in order to perform this separation, it is convenient to somewhat modify the application of the variational method by introducing three sets of variables in place of two, namely, the angular variables fixing the orientation of the whole, the variables fixing the positions of the nuclei with respect to one another and the variables fixing the positions of the electrons with respect to the nuclei. In concluding we shall limit ourselves to a few summary indications concerning diatomic molecules.

Once the separation of the center of mass has been effected, the dynamical variables of the two nuclei describe the relative motion of the one with respect to the other. Thus in this case the set of coordinates X reduces to the coordinates of the vector $\boldsymbol{R} = \boldsymbol{R}_1 - \boldsymbol{R}_2$ giving the position of the first nucleus with respect to the second. The coordinates of the electrons with respect to the center of mass of the nuclei will still be collectively denoted by x. Neglecting, for simplicity, the spins [1]), we denote the (orbital) angular momentum of the electrons by $\boldsymbol{L}$, that of the nuclei by $\boldsymbol{G}(\boldsymbol{G} = \boldsymbol{R} \times \boldsymbol{P})$ and the total angular momentum of the system by $\boldsymbol{K}$

$$\boldsymbol{K} = \boldsymbol{G} + \boldsymbol{L}. \qquad \text{(XVIII.67)}$$

[1]) If the spin-orbit coupling is small as compared with the rotational quanta, the presence of the spins will not sensibly modify the results and will be reflected simply in the appearance of fine structure as in the case of LS coupling for atoms. In the opposite case, the general form of the spectrum is rather considerably modified. Cf. L. D. Landau and E. M. Lifshitz, *loc. cit.*, note, p. 785.

Finally, we denote the unit vector along the molecular axis by $\boldsymbol{u}(\boldsymbol{u}=\boldsymbol{R}/R)$, and label vector components along this axis with the index u. In passing, we note the operator identity

$$\boldsymbol{G}\cdot\boldsymbol{R} = 0. \tag{XVIII.68}$$

$G_u(\equiv \boldsymbol{G}\cdot\boldsymbol{u})$ is therefore identically null and we have

$$K_u = L_u.$$

For the rest we retain the notation of the preceding paragraphs.

Consider first the motion of the electrons. The symmetries of $H^{(0)}$, considered as an operator acting on the electron variables alone, are the same as those of the Stark effect: $H^{(0)}$ is invariant under rotations about $\boldsymbol{u}$ and under reflections in planes passing through $\boldsymbol{u}$. If we have two identical nuclei, $H^{(0)}$ is also invariant under reflection proper $(x \to -x)$; we leave aside this particular case for the moment.

To classify the states, we need only to refer to the discussion of the Stark-Effect in § XV.14. Each level of $H^{(0)}$ corresponds to a given eigenvalue of $L_u{}^2$. The eigenvalues of $L_u{}^2$ may be written in the form $\hbar^2\Lambda^2$, where Λ is a quantum number taking all non-negative integral values. Following spectroscopic tradition, the terms corresponding to the first three values $\Lambda = 0, 1, 2$, are denoted respectively by the greek capitals Σ, Π, Δ. If $\Lambda \neq 0$, the level is doubly degenerate and we can find two corresponding eigenvectors corresponding respectively to the values $\hbar\Lambda$ and $-\hbar\Lambda$ for the component of angular momentum along the molecular axis. The Σ levels $(\Lambda = 0)$ are non-degenerate; they can be put into two categories, Σ^+ and Σ^-, according as their eigenvector is invariant or changes sign under reflection in a plane passing through the axis. In the great majority of diatomic molecules, the ground state is a Σ^+ state.

We now consider the total Hamiltonian. The appearance of degeneracy in the electronic levels is a complicating factor only in appearance. It can be shown, in fact, that the coupling introduced by T_N between the states $\hbar\Lambda$ and $-\hbar\Lambda$ is smaller than the separation between rotational levels; it is therefore simply neglected. In other words in addition to approximation (XVIII.50) we make the second approximation $L_u = Cst.$

On the other hand, although the mode of application of the varia-

tional method employed in § 14 and § 15 is well suited to a discussion of a general nature it does not permit the separation of rotations and vibrations to be easily exhibited. The difficulty stems from the correction $W_n'(\mathbf{R})$ being generally not invariant under rotation. It is preferable to take the rotational invariance of H and $H^{(0)}$ into account from the outset and to look for the eigenvectors of H having a given total angular momentum. The eigenvectors of $H^{(0)}$ can be characterized by:

(*i*) the quantum numbers of total angular momentum K and m_K (m_K, component of $\mathbf{K}$ along a *fixed* axis in space, not to be confused with K_u);

(*ii*) the quantum numbers specifying the state of the electrons, namely n, and when $\Lambda \neq 0$, s ($s \equiv$ sign of L_u);

(*iii*) the internuclear distance R.

The Schrödinger equation (XVIII.41) then reads

$$H^{(0)}|Km_K n s R\rangle = W_n(R)|Km_K n s R\rangle. \qquad \text{(XVIII.69)}$$

The adiabatic approximation consists in looking for the eigenvectors of H in the subspace of vectors corresponding to well-defined values of K, m_K, n and s, i.e. among the vectors of the form

$$\int_0^\infty |K\, m_K\, n\, s\, R\rangle\, y(R)\, \mathrm{d}R \qquad \text{(XVIII.70)}$$

[to be compared with the form (XVIII.51)]. The variational method is now applied as in § 14 and § 15 and leads to a Schrödinger equation for the radial function $y(R)$ whose eigenvalues are the molecular energy-levels corresponding to quantum numbers K and n (these levels are independent of m_K and s).

The Hamiltonian h of this radial equation is obtained by the method of § 15. From $H^{(0)}$ we have the contribution $W_n(R)$. To calculate the contribution from T_N, the kinetic energy of the nuclei, it is convenient to write this operator in the form

$$T_N = \frac{P_R^2}{2M} + \frac{G^2}{2MR^2} \qquad \text{(XVIII.71)}$$

(M reduced mass of the nuclei; P_R radial momentum) and to separately

treat the two terms on the right-hand side. It is easily shown that the first term gives

$$-\frac{\hbar^2}{2M}\frac{\mathrm{d}^2}{\mathrm{d}R^2}+w_n'(R), \tag{XVIII.72}$$

where $w_n'(R)$ is a small correction independent of K, m_K and s. The contribution from the second term, representing the kinetic energy of rotation, requires a more careful examination. It is equal to the product of $\hbar^2/2MR^2$ by the average value of the operator $\boldsymbol{G}^2$ in the subspace of the vectors given by (XVIII.70). With the aid of (XVIII.67) we can write

$$\begin{aligned}\langle \boldsymbol{G}^2\rangle &= \langle(\boldsymbol{K}-\boldsymbol{L})^2\rangle \\ &= \hbar^2K(K+1)-2\langle \boldsymbol{K}\cdot\boldsymbol{L}\rangle+\langle \boldsymbol{L}^2\rangle.\end{aligned}$$

Since $|Km_KnsR\rangle$ is an eigenstate of L_u, the average value of the component of $\boldsymbol{L}$ along an axis perpendicular to $\boldsymbol{u}$ is null; hence, taking (XVIII.68) into account,

$$\langle \boldsymbol{K}\cdot\boldsymbol{L}\rangle = \langle K_u\, L_u\rangle = \langle {L_u}^2\rangle = \hbar^2\Lambda^2.$$

$\langle \boldsymbol{L}^2\rangle$, on the other hand, is a certain positive quantity depending only on n. If we put $w_n''(R) \equiv (\langle \boldsymbol{L}^2\rangle - 2\hbar^2\Lambda^2)/2MR^2$, we obtain for the contribution of the kinetic energy of rotation the expression

$$\frac{\hbar^2K(K+1)}{2MR^2}+w_n''(R). \tag{XVIII.73}$$

The small corrections $w_n'(R)$ and $w_n''(R)$ may be incorporated in the potential $W_n(R)$. The radial Hamiltonian then becomes

$$h \equiv -\frac{\hbar^2}{2M}\frac{\mathrm{d}^2}{\mathrm{d}R^2}+\frac{\hbar^2K(K+1)}{2MR^2}+W_n(R). \tag{XVIII.74}$$

This is the Hamiltonian for a particle of mass M and angular momentum K in the potential $W_n(R)$.

Note that the possible values for the integer K depend on n. By fixing n we fix Λ, and since $K \geqslant |K_u|$,

$$K \geqslant \Lambda. \tag{XVIII.75}$$

The energy levels E are the eigenvalues of h and are obtained by solving the radial equation

$$h\, y(R) = E\, y(R).$$

For given K and n, these are the levels of a particle of mass M and angular momentum K in the potential $W_n(R)$; to distinguish them one from another we shall use another quantum number which we denote v. n is the electronic quantum number, v the vibrational quantum number, K the rotational quantum number.

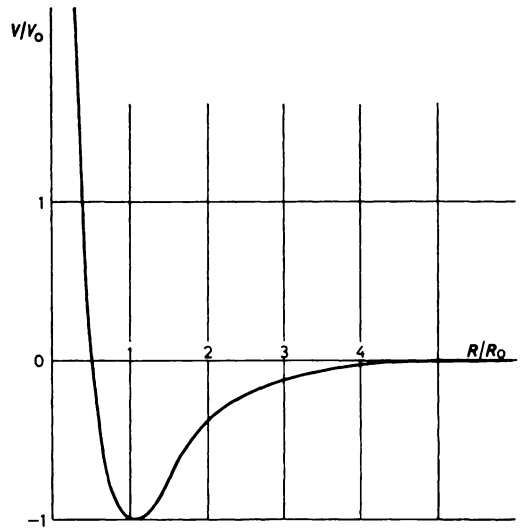


Fig. XVIII.2. The Morse potential: $V = V_0\,(e^{-2(R-R_0)/b} - 2e^{-(R-R_0)/b})$. R_0 is the equilibrium internuclear distance, V_0 the potential depth at the equilibrium position. The sharpness of the minimum varies as the smallness of the ratio b/R_0. The value adopted in the figure, $b/R_0 = 0.68$, corresponds to the hydrogen molecule.

The spectrum obtained owes its essential characteristics to the general form of $W_n(R)$; this function has a sharp minimum for a certain value R_0 of the internuclear distance corresponding to the position of stable equilibrium for the molecule. Experiment shows that $W_n(R)$ can most often be represented to a good approximation by a Morse potential about its position of equilibrium (cf. Fig. XVIII.2 and Problem XVIII.5). For the lowest levels the wave function $y(R)$ remains concentrated in a small domain about R_0 (approximate width $\varkappa R_0$) and represents vibrations of the nuclei about the point of equilibrium.

In a first (the harmonic) approximation, one can replace the R in the rotational energy term by R_0, and substitute for $W_n(R)$ the first two terms of its expansion in powers of $q \equiv R - R_0$. With the notations

$$B_n = \hbar^2/2MR_0^2, \quad W_n(R_0) = E_n, \quad W_n''(R_0) = M\omega_n^2,$$

this gives the radial equation

$$\left[-\frac{\hbar^2}{2M}\frac{\mathrm{d}^2}{\mathrm{d}q^2} + E_n + \tfrac{1}{2}M\omega_n^2 q^2 + B_n K(K+1) \right] y = E_{nvK}\, y.$$

This is the Schrödinger equation for a harmonic oscillator, so that we have

$$\begin{gathered} E_{nvK} = E_n + (v+\tfrac{1}{2})\hbar\omega_n + K(K+1)B_n \\ (v = 0, 1, 2, \ldots; \quad K = \Lambda, \Lambda+1, \Lambda+2, \ldots). \end{gathered} \tag{XVIII.76}$$

The energy is thus the sum of three terms: the energy of the electrons E_n, the vibrational energy $(v+\frac{1}{2})\hbar\omega_n$ and the rotational energy $K(K+1)B_n$. The order of magnitude found for the quanta $\hbar\omega_n$ and B_n confirms the analysis of § 12 and the subsequent deductions concerning molecular spectra [1]).

We shall not dwell here upon the necessary modifications of the theory due to the existence of spin.

We only mention one particularly striking effect due to spin and nuclear statistics. Since the coupling of the nuclear spins with the rest of the molecule is entirely negligible, the main effect is the existence of a spin degeneracy or multiplicity for each of the levels. The molecular eigenfunctions are of the form

$$\Psi = \Phi(\mathbf{R}, x)\, \chi(\mu_1, \mu_2)$$

and the order of degeneracy is equal to the number of linearly independent functions $\chi(\mu_1, \mu_2)$ that can be formed. If j_1 and j_2 are respectively

---

[1]) For the electronic ground state of the hydrogen atom (Σ^+ state), the first levels are accurately represented by this formula with $E_0 = -4.72$ eV, $\hbar\omega_0 = 0.54$ eV and $B_0 = 0.0074$ eV. These figures correspond to an internuclear distance: $R_0 = 0.74 \times 10^{-8}$ cm. A better representation of the vibration and rotation spectra is obtained by taking a Morse potential for $W_0(R)$ with $V_0 = 4.72$ eV, $R_0 = 0.74 \times 10^{-8}$ cm and $b = 0.68\, R_0$ (cf. Fig. XVIII.2 and Problem XVIII.5); the dissociation energy of the molecule is $V_0 - \frac{1}{2}\hbar\omega_0$.

the spins of the two nuclei, there exist in all $(2j_1+1)(2j_2+1)$ of these functions. If the two nuclei are different, there is no restriction on Ψ and the multiplicity of each level is just $(2j_1+1)(2j_2+1)$. If we have two identical nuclei of spin j ($j_1=h_2=j$), Ψ must be symmetrical or antisymmetrical with respect to their exchange

$$(\mathbf{R} \leftrightarrow -\mathbf{R},\ \mu_1 \leftrightarrow \mu_2)$$

according as they are bosons or fermions, i.e. according as j is integral or half-integral. One can form $(j+1)(2j+1)$ symmetrical functions and $j(2j+1)$ antisymmetrical functions in the $(2j+1)^2$-dimensional space of the χ [Problem (XIII.13)]. The multiplicity g therefore depends on whether $\Psi(\mathbf{R}, x)$ is symmetrical or antisymmetrical, in other words it depends on the parity of Φ in a reflection of the nuclei alone. Denoting this parity by ϖ ($\varpi=\pm 1$) we have

$$g = \begin{cases} (j+1)(2j+1) & j(2j+1) \quad \text{if} \quad \varpi = +1 \\ j(2j+1) & (j+1)(2j+1) \quad \text{if} \quad \varpi = -1. \end{cases}$$

(j integral) (j half-integral)

The interesting effect is due to the fact that ϖ depends on the parity of the total orbital angular momentum. It can be shown that

$$\varpi = (-)^K \varpi_e,$$

where ϖ_e ($=\pm 1$) depends on the behavior under reflection of the electronic wave function corresponding to Φ. If $\Lambda \neq 0$ each level has two corresponding wave functions of opposite ϖ_e. For the Σ levels ($\Lambda=0$) we have on the contrary only one electronic wave function, ϖ_e has a well-defined value, $+1$ or -1, and the multiplicity g varies in a characteristic manner from one rotational level to another [1]).

These features of the energy spectrum of the Σ levels of homonuclear diatomic molecules show up clearly in the "band spectra" of these molecules. Since the probabilities for optical transitions between states with different spin functions are very small, the transitions observed in practice all conserve the parity of K. In addition, under the usual conditions of observation the relative intensity of rays corresponding to even values of K and rays corre-

[1]) In particular, if $j = 0$ the levels with $\varpi = -1$ are absent and K can take only values of well-defined parity.

sponding to odd values of K is directly related to the ratio of the multiplicities calculated above, either $(j+1)/j$ or $j/(j+1)$ according to the sign of ϖ_e and according as the nuclei are bosons or fermions. This provides a particularly direct method for measuring nuclear spins.

EXERCISES AND PROBLEMS

1. E_0 and Ψ_0 are the ground-state energy and wave function respectively of a given quantum system. A variational calculation gives the energy E_{var} and the wave function Ψ_{var}. Ψ_0 and Ψ_{var} being supposed normalized to unity, we put $\varepsilon = 1-|\langle\Psi_0|\Psi_{\text{var}}\rangle|^2$. ε is the norm of the projection of Ψ_{var} onto the subspace orthogonal to Ψ_0, and is a measure of the square of the distance from the state Ψ_{var} to the state Ψ_0. Prove that

$$E_{\text{var}} - E_0 \geqslant \varepsilon(E_1 - E_0)$$

(E_1 energy of the first excited level). Verify that it is satisfied by the results of § 4. Discuss.

2. Supposing the ground-state wave function for the hydrogen atom known, calculate the level of the first excited state by taking as trial function that part of $u = \varrho\, e^{-b\varrho}$ which is orthogonal to the ground-state wave function (notation of § 4). Compare with the exact eigenvalue and eigenfunction. Discuss.

3. Prove inequalities (XVIII.11) and (XVIII.12).

4. Taking the trial function $u = \varrho^{l+1} e^{-b\varrho}$ (notation of § 4) calculate the lowest hydrogen atom level corresponding to angular momentum l. Compare with the exact results and comment.

5. Consider a particle in one dimension of mass M in the potential

$$V(q) = V_0[e^{-2q/b} - 2e^{-q/b}]$$

(*the Morse potential*). Determine the corresponding wave functions and energy levels.

N.B. The wave equation is taken over into a Laplace equation by the following changes of function and variable

$$\xi = 2K_0\, b\, e^{-q/b}, \qquad w(\xi) = e^{\frac{1}{2}\xi}\, \xi^{-\varkappa b}\, \psi(q),$$

where

$$K_0 = \sqrt{2MV_0}/\hbar \qquad \varkappa = \sqrt{-2M\overline{E}}/\hbar.$$

We find a finite number of discrete levels depending on the index n in accordance with the formula:

$$E_n = -V_0\left(1 - \frac{n+\frac{1}{2}}{K_0 b}\right)^2 \qquad (0 \leqslant n \leqslant K_0 b - \tfrac{1}{2}).$$

CHAPTER XIX

COLLISION THEORY

1. Introduction

Up to the present, we have considered only the simplest collision problems, namely the scattering of an elementary particle by a potential (and notably by a central potential) and the mutual scattering of two elementary particles, the second problem reducing to the first when the relative and center-of-mass motions are separated. The treatment of this type of problem was described in Chapter X, and was complemented by a treatment of Coulomb scattering (Ch. XI, Section II) and of the mutual scattering of two identical particles (§ XIV.9–10). In addition, in § XVII.5 we obtained a particularly simple expression for the scattering of a particle by a potential $V(\boldsymbol{r})$ by treating $V(\boldsymbol{r})$ as a perturbation and considering only the first-order effect [eq. (XVII.54)]. However, the work of § XVII.5 is based on a somewhat questionable definition of the cross section and it is worth giving a rigorous derivation of this formula.

The object of this chapter is, on the one hand, to set up a formalism in which we can treat collisions between complex particles, and on the other, to extend the perturbation and variational methods of the preceding chapters to the calculation of collision cross sections.

These problems may be approached in two different ways. The first consists in giving a rigorous justification of the definition of § XVII.5 equating cross sections with probabilities per unit time and per unit incident flux, and then relating these quantities to the matrix elements of the evolution operator $U(t, t')$ in the limit when $t \to +\infty$ and $t' \to -\infty$. The second is a simple generalization of the treatment of Chapter X (Section I) in which the cross sections are directly related to the asymptotic behavior of the stationary solutions of the Schrödinger equation. It is the second of these clearly equivalent procedures that we adopt in the present chapter [1]).

---

[1]) For an account of the first approach, see B. Lippman and J. Schwinger, Phys. Rev. **79** (1949) 469 or M. Gell-Mann and M. L. Goldberger, Phys. Rev. **91** (1953) 398. The definitions of cross sections in both of these accounts are somewhat questionable: a discussion of this point and the possible remedies is given by S. Sunakawa, Prog. Theor. Phys. **14** (1955) 175.

This chapter has in all five sections. The first two are devoted to establishing the formalism and describing the perturbation method in the simple case of a particle in a potential $V(\boldsymbol{r})$. This is extended to complex collisions in Section III. Section IV is devoted to variational methods. In Section V we give a certain number of properties of the collision amplitudes that follow directly from general properties of the Hamiltonian such as Hermiticity, time-reversal invariance and other invariance properties.

I. FREE WAVE GREEN'S FUNCTION AND THE BORN APPROXIMATION

2. Integral Representations of the Scattering Amplitude

In this and the following sections we shall discuss the scattering of a particle of mass m by a potential $V(\boldsymbol{r})$. We denote the kinetic energy operator by H_0 and the total Hamiltonian by H:

$$H_0 \equiv -\frac{\hbar^2}{2m}\Delta \tag{XIX.1}$$

$$H \equiv -\frac{\hbar^2}{2m}\Delta + V(\boldsymbol{r}). \tag{XIX.2}$$

We shall suppose that the potential V tends asymptotically to 0 more rapidly than $1/r$. Potentials with the $1/r$ behavior will briefly be examined in § 15.

In what follows we shall introduce various types of waves which we wish to distinguish with suitable notations. The letter φ will be reserved for plane waves, the letter ψ for the stationary solutions of the Hamiltonian H. For a given propagation vector $\boldsymbol{k}$ we can define

(*i*) the plane wave $\varphi_{\boldsymbol{k}} \equiv e^{i\boldsymbol{k}\cdot\boldsymbol{r}}$

(*ii*) the stationary waves $\psi_{\boldsymbol{k}}^{(+)}$ and $\psi_{\boldsymbol{k}}^{(-)}$ respectively characterized by the asymptotic behavior $e^{i\boldsymbol{k}\cdot\boldsymbol{r}}+$ outgoing wave and $e^{i\boldsymbol{k}\cdot\boldsymbol{r}}+$ incoming wave.

In particular, if $\boldsymbol{k}_a$ is the propagation vector of the incident particles and E their energy $[k_a = k = (2mE)^{\frac{1}{2}}]$, then the complete stationary scattering wave function (definition of § X.3) is $\psi_{\boldsymbol{k}_a}^{(+)}$. This wave and

the corresponding scattering amplitude $f^{(+)}_{k_a}$ are determined by the conditions:

$$H\psi^{(+)}_{k_a} = E\psi^{(+)}_{k_a} \qquad \psi^{(+)}_{k_a} \underset{r\to\infty}{\sim} e^{ik_a\cdot r} + f^{(+)}_{k_a}(\Omega)\frac{e^{ikr}}{r}. \tag{XIX.3}$$

We denote by $d\sigma_{a\to b}/d\Omega$ the scattering cross section in the direction Ω_b. In Chapter X we showed that

$$d\sigma_{a\to b}/d\Omega = |f^{(+)}_{k_a}(\Omega_b)|^2. \tag{XIX.4}$$

Let k_b be a vector of length k pointing in the direction Ω_b. Throughout this section the only stationary waves to appear will be those defined by the vectors k_a and k_b so we shall simplify the writing by systematically replacing the indices k_a and k_b by a and b respectively.

We begin by establishing a very general property of the scattering amplitudes. Consider two potentials $U(r)$ and $\hat{U}(r)$ and denote by ξ and $\hat{\xi}$ the respective stationary solutions of the corresponding Hamiltonians. We are especially interested in the solutions $\hat{\xi}_b^{(-)}$ and $\xi_a^{(+)}$:

$$\left[-\frac{\hbar^2}{2m}\Delta + \hat{U}\right]\hat{\xi}_b^{(-)} = E\,\hat{\xi}_b^{(-)} \tag{XIX.5}$$

$$\left[-\frac{\hbar^2}{2m}\Delta + U\right]\xi_a^{(+)} = E\,\xi_a^{(+)} \tag{XIX.6}$$

$$\hat{\xi}_b^{(-)} \underset{r\to\infty}{\sim} e^{ik_b\cdot r} + \hat{f}_b^{(-)}(\Omega)\frac{e^{-ikr}}{r} \tag{XIX.7}$$

$$\xi_a^{(+)} \underset{r\to\infty}{\sim} e^{ik_a\cdot r} + f_a^{(+)}(\Omega)\frac{e^{ikr}}{r}. \tag{XIX.8}$$

It will now be shown that

$$\begin{aligned}\langle\hat{\xi}_b^{(-)}|(U-\hat{U})|\xi_a^{(+)}\rangle &\equiv \int \hat{\xi}_b^{(-)*}(r)\,[U(r)-\hat{U}(r)]\,\xi_a^{(+)}(r)\,dr \\ &= -\frac{2\pi\hbar^2}{m}\,[f_a^{(+)}(\Omega_b) - \hat{f}_b^{(-)*}(-\Omega_a)]\end{aligned} \tag{XIX.9}$$

(by convention, $-\Omega_a$ denotes the direction opposite to Ω_a).

The proof is analogous to the one used to obtain integral representations of the phase shifts in § X.17. There being no possibility of confusion, we shall sometimes omit the indices b and a and the

symbol $\hat{}$. Multiply (XIX.6) by $\xi^{(-)*}$ and subtract the equation obtained by multiplying the complex conjugate of (XIX.5) by $\xi^{(+)}$. Since $\hat{U}$ is real, this gives

$$-\frac{\hbar^2}{2m}\left[\xi^{(-)*}(\Delta\xi^{(+)}) - (\Delta\xi^{(-)*})\,\xi^{(+)}\right] + \xi^{(-)*}(U-\hat{U})\,\xi^{(+)} = 0.$$

Integrating over the volume of a sphere of radius R centered at the origin, we get

$$\langle\xi^{(-)}|(U-\hat{U})|\xi^{(+)}\rangle = \frac{\hbar^2}{2m}\lim_{R\to\infty}\int_{r<R}\left[\xi^{(-)*}(\Delta\xi^{(+)}) - (\Delta\xi^{(-)*})\,\xi^{(+)}\right]\mathrm{d}\boldsymbol{r}. \quad \text{(XIX.10)}$$

In spite of the Hermitean character of Δ, the integral on the right-hand side does not necessarily vanish in the limit since the ξ are not square-integrable functions. To calculate this integral, we use Green's theorem to transform it into a surface integral over the sphere of radius R.

If ξ_1 and ξ_2 are functions of $\boldsymbol{r}$, we shall denote the integral of the "radial Wronskian"

$$W[\xi_1, \xi_2] \equiv \xi_1\frac{\partial\xi_2}{\partial r} - \xi_2\frac{\partial\xi_1}{\partial r},$$

over the sphere by the notation $\{\xi_1, \xi_2\}$:

$$\begin{aligned}\{\xi_1, \xi_2\} &\equiv \int W[\xi_1, \xi_2]|_{r=R}\, R^2\,\mathrm{d}\Omega \\ &\equiv \int W[r\xi_1, r\xi_2]|_{r=R}\,\mathrm{d}\Omega.\end{aligned} \quad \text{(XIX.11)}$$

Using Green's theorem,

$$\{\xi_1, \xi_2\} = \int_{r<R}\left[\xi_1(\Delta\xi_2) - (\Delta\xi_1)\xi_2\right]\mathrm{d}\boldsymbol{r}.$$

Equation (XIX.10) may then be written:

$$\langle\xi^{(-)}|(U-\hat{U})|\xi^{(+)}\rangle = \frac{\hbar^2}{2m}\lim_{R\to\infty}\{\xi^{(-)*}, \xi^{(+)}\}. \quad \text{(XIX.12)}$$

The surface integral $\{\xi^{(-)*}, \xi^{(+)}\}$ tends asymptotically to a constant which may easily be determined by substituting for the ξ and their radial derivatives the first term in their asymptotic expansions

in powers of $1/r$. Substituting from (XIX.7) and (XIX.8), we find:

$$\lim_{R\to\infty}\{\xi^{(-)*},\xi^{(+)}\}=\lim_{R\to\infty}\{e^{-ik_b\cdot r},e^{ik_a\cdot r}\}+\lim_{R\to\infty}\left\{e^{-ik_b\cdot r},f^{(+)}\frac{e^{ikr}}{r}\right\}$$
$$+\lim_{R\to\infty}\left\{f^{(-)*}\frac{e^{ikr}}{r},e^{ik_a\cdot r}\right\}+\lim_{R\to\infty}\left\{f^{(-)*}\frac{e^{ikr}}{r},f^{(+)}\frac{e^{ikr}}{r}\right\}. \quad \text{(XIX.13)}$$

The asymptotic form of a plane wave is given by (Problem XIX.1):

$$e^{ik\cdot r}\underset{r\to\infty}{\sim}\frac{2\pi}{ikr}[\delta(\Omega_r-\Omega_k)\,e^{ikr}-\delta(\Omega_r+\Omega_k)\,e^{-ikr}]+O\left(\frac{1}{r^2}\right). \quad \text{(XIX.14)}$$

In the bracket on the right-hand side, Ω_r and Ω_k denote the directions of the vectors $\boldsymbol{r}$ and $\boldsymbol{k}$ respectively. The δ functions are defined by the property

$$\int\delta(\Omega'-\Omega)\,\varphi(\Omega')\,d\Omega'=\varphi(\Omega)$$

for any function φ of the angles Ω. If we replace the plane waves in the right-hand side of (XIX.13) by their asymptotic expressions, the angular integrations are easily performed. As $\boldsymbol{k}_b$ and $\boldsymbol{k}_a$ have the same length, the first term vanishes (in agreement with the result of Problem XIX.2). In the second, the only non-vanishing contribution comes from the *incoming part* of the plane wave, and is equal to $-4\pi f^{(+)}(\Omega_b)$. The same is true for the third term, whose contribution is $4\pi f^{(-)*}(-\Omega_a)$. The fourth term is null. We have, therefore:

$$\lim_{R\to\infty}\{\xi^{(-)*},\xi^{(+)}\}=-4\pi[f^{(+)}(\Omega_b)-f^{(-)*}(-\Omega_a)].$$

Substituting this result in (XIX.12), we obtain relation (XIX.9). Q.E.D.

Property (XIX.9) is interesting chiefly because it is independent of the particular form of the potentials U and $\hat{U}$. It is required only that the latter be real and that they fall off asymptotically more rapidly than $1/r$.

As a first application, we take

$$U=V \qquad \hat{U}=0.$$

Relation (XIX.9) gives:

$$\langle\varphi_b|V|\psi_a^{(+)}\rangle=-\frac{2\pi\hbar^2}{m}f_a^{(+)}(\Omega_b). \quad \text{(XIX.15)}$$

This integral representation of the scattering amplitude will be used in the next two paragraphs as a starting point for the Born approximation.

We next take

$$U = \hat{U} = V.$$

Since the left-hand side of (XIX.9) vanishes, the two amplitudes in the right-hand side are equal:

$$f_a^{(+)}(\Omega_b) = f_b^{(-)*}(-\Omega_a). \tag{XIX.16}$$

Finally, let us take

$$U = 0, \qquad \hat{U} = V.$$

Relation (XIX.9) gives:

$$\langle \psi_b^{(-)} | V | \varphi_a \rangle = -\frac{2\pi\hbar^2}{m} f_b^{(-)*}(-\Omega_a). \tag{XIX.17}$$

From (XIX.15–17) we obtain the important equality:

$$\langle \psi_b^{(-)} | V | \varphi_a \rangle = \langle \varphi_b | V | \psi_a^{(+)} \rangle. \tag{XIX.18}$$

3. Cross Sections and the T Matrix. Microreversibility

If the scattering amplitude in (XIX.4) is replaced by its integral representation (XIX.15), we get

$$\frac{d\sigma_{a\to b}}{d\Omega} = \frac{m^2}{4\pi^2\hbar^4} |\langle \varphi_b | V | \psi_a^{(+)} \rangle|^2.$$

In order to have an expression resembling (XVII.54), we introduce the density of states at energy E

$$\varrho(E) = (2\pi\hbar)^{-3} p^2 (dp/dE) = m\hbar k/(2\pi\hbar)^3$$

in accordance with the definition of § XVII.5 [eq. (XVII.52)], and denote the initial velocity by $v = \hbar k/m$. The preceding equation then becomes [1]:

$$\frac{d\sigma_{a\to b}}{d\Omega} = \frac{2\pi}{\hbar v} |\langle \varphi_b | V | \psi_a^{(+)} \rangle|^2 \varrho(E). \tag{XIX.19}$$

Since $|\psi_a^{(+)}\rangle$ and $|\varphi_b\rangle$ cannot be basis vectors in the same representa-

[1] In the other formulation of collision theory (cf. reference, note, p. 801), this form follows in a natural way from the very definition of cross section.

tion, $\langle \varphi_b | V | \psi_a^{(+)} \rangle$ is not really a matrix element of the operator V. It is convenient to introduce a matrix T such that:

$$T_{a \to b} \equiv \langle \varphi_b | T | \varphi_a \rangle = \langle \varphi_b | V | \psi_a^{(+)} \rangle. \tag{XIX.20}$$

T will be called the *transition matrix*, and the element $T_{a \to b}$ the *amplitude for the transition* $a \to b$. It is to be observed that these matrix elements are taken between plane waves of the same energy. One can obviously conceive of an operator T in Hilbert space satisfying condition (XIX.20); but it is not completely defined by this condition which merely determines certain of its matrix elements in a plane wave representation, namely the elements between waves corresponding to a specified energy. To completely define T we must also determine the matrix elements between waves corresponding to energies different to E. This will be done further on.

Formula (XIX.19) may now be written

$$\frac{\mathrm{d}\sigma_{a \to b}}{\mathrm{d}\Omega} = \frac{2\pi}{\hbar v} |T_{a \to b}|^2 \varrho(E). \tag{XIX.19'}$$

Before going on, we mention the property of microreversibility. We label the various waves corresponding to momenta $-\boldsymbol{k}_a$ and $-\boldsymbol{k}_b$ by the indices Ka and Kb respectively. We have the obvious relation between plane waves:

$$\varphi_{-\boldsymbol{k}} = \varphi_{\boldsymbol{k}}^*.$$

In addition, since the Hamiltonian H is real,

$$\psi_{-\boldsymbol{k}}^{(\mp)} = \psi_{\boldsymbol{k}}^{(\pm)*}.$$

In particular

$$\varphi_a = \varphi_{Ka}^* \qquad \psi_b^{(-)} = \psi_{Kb}^{(+)*}$$

and, since V is real, the "matrix elements" $\langle \psi_b^{(-)} | V | \varphi_a \rangle$ and $\langle \varphi_{Ka} | V | \psi_{Kb}^{(+)} \rangle$ represent the same integral. Thus (XIX.18) may also be written

$$\langle \varphi_{Ka} | V | \psi_{Kb}^{(+)} \rangle = \langle \varphi_b | V | \psi_a^{(+)} \rangle$$

or again, from definition (XIX.20),

$$T_{Kb \to Ka} = T_{a \to b}. \tag{XIX.21}$$

A fortiori, these two amplitudes have the same modulus. Returning

to (XIX.19'), we obtain *the microreversibility property for elastic scattering* [1]:

$$d\sigma_{Kb\to Ka}/d\Omega = d\sigma_{a\to b}/d\Omega. \tag{XIX.22}$$

4. The Born Approximation

Formula (XIX.19) is *exact*. Like formula (X.2) from which it was derived, it relates the cross section to the stationary scattering solution $\psi_a^{(+)}$. In the present case the latter no longer appears in its asymptotic form but as a factor in an integral. If we replace it by an approximate solution we get an approximate expression for the cross section.

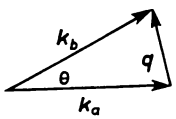


Fig. XIX.1.

In particular, for $V(r)$ sufficiently small, $\psi_a^{(+)}$ differs little from the incident plane wave φ_a, and may be replaced by the latter in the calculation of the transition amplitude. This gives the *Born approximation*:

$$T_{a\to b} \simeq T^{(B)}_{a\to b} = \langle \varphi_b | V | \varphi_a \rangle. \tag{XIX.23}$$

In this approximation (XIX.19) reduces to formula (XVII.54), justifying our conjecture that it gives an approximate expression for the cross section in the limit when $V(r)$ may be treated as a perturbation.

Put:

$$\boldsymbol{q} = \boldsymbol{k}_b - \boldsymbol{k}_a,$$

$\hbar\boldsymbol{q}$ is the momentum transferred to the particle in the course of the collision (Fig. XIX.1). The length of this vector depends on the scattering angle θ between the two vectors $\boldsymbol{k}_a$ and $\boldsymbol{k}_b$ as follows

$$q = 2k \sin \tfrac{1}{2}\theta \qquad (k = k_a = k_b). \tag{XIX.24}$$

From (XIX.23),

$$T^{(B)}_{a\to b} = \int e^{-i\boldsymbol{q}\cdot\boldsymbol{r}} V(\boldsymbol{r})\, d\boldsymbol{r} \equiv \mathscr{V}(\boldsymbol{q}). \tag{XIX.25}$$

Thus in the Born approximation the scattering cross section takes the particularly simple form

$$\frac{d\sigma^{(B)}_{a\to b}}{d\Omega} = \frac{m^2}{4\pi^2\hbar^4} |\mathscr{V}(\boldsymbol{q})|^2. \tag{XIX.26}$$

It is proportional to the squared modulus of the Fourier component of

[1] For a central potential this property can also be deduced from rotational invariance. Here we see that it holds for non-central potentials as well.

$V(\mathbf{r})$ corresponding to the momentum transferred in the collision. Note that the dependence on energy and angle is entirely contained in the vector $\boldsymbol{q}$.

With a central potential the situation is simpler still. After integration over the angles the right-hand side of (XIX.25) becomes

$$\mathscr{V}(q) = \frac{1}{q}\int_0^\infty \sin qr\, V(r)\, r\, \mathrm{d}r. \tag{XIX.27}$$

The differential cross section now depends only on the length $\hbar q$ of the momentum transfer. As for the total cross section σ_{tot}, if we take q as the variable of integration instead of θ, then

$$\sigma_{\text{tot}}^{(\text{B})} = \frac{m^2}{2\pi\hbar^4 k^2}\int_0^{2k} \mathscr{V}^2(q)\, q\, \mathrm{d}q. \tag{XIX.28}$$

From these formulas, one can draw a certain number of general conclusions regarding the behavior of the cross sections at high energies. Denote the range of the potential $V(\mathbf{r})$ by a. Since the real function $V(\mathbf{r})$ takes appreciable values only in a region with linear dimensions a, its Fourier transform takes appreciable values only in a region of linear dimensions $1/a$ about the origin [1]).

The cross section is therefore relatively important only in the region $q \lesssim 1/a$. According to (XIX.24) this region corresponds to scattering angles such that

$$\sin \tfrac{1}{2}\theta \lesssim \frac{1}{2ka}.$$

At high energies ($ka \gg 1$), the scattering is *essentially concentrated in the forward direction*, in a cone with θ less than $1/ka$. These conclusions, based on the Born approximation, may be compared with the results of § X.13 for the high-energy scattering by a hard sphere.

The total cross section tends to zero like $1/E$. This result is evident from (XIX.28), whose asymptotic form is

$$\sigma_{\text{tot}}^{(\text{B})} \underset{k\to\infty}{\sim} \frac{m^2}{2\pi\hbar^4 k^2}\int_0^\infty \mathscr{V}^2(q)\, q\, \mathrm{d}q. \tag{XIX.29}$$

It may be extended without difficulty to non-central potentials.

[1]) In the present discussion, a may be taken as a length of the order of the mean-square deviation of the distribution

$$\frac{V^2(\mathbf{r})}{\int V^2(\mathbf{r}')\, \mathrm{d}\mathbf{r}'}.$$

5. Integral Equation for Scattering

Up to the present, $\psi_a^{(+)}$ has been defined as a solution of the Schrödinger equation satisfying certain boundary conditions. It will now be shown that it is also the solution of a certain integral equation. With this new definition we shall be able to write $\psi_a^{(+)}$ as an expansion in powers of V and evaluate the corrections to the Born approximation.

To this end, we make use of the identity [1])

$$(\Delta + k^2)\frac{e^{ikr}}{r} = -4\pi\delta(\mathbf{r}).$$

It follows that the function

$$\mathcal{G}(\mathbf{r}, \mathbf{r}') = -\frac{m}{2\pi\hbar^2}\frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} \tag{XIX.30}$$

has the property

$$\frac{\hbar^2}{2m}(\Delta + k^2)\,\mathcal{G}(\mathbf{r}, \mathbf{r}') = \delta(\mathbf{r}-\mathbf{r}'). \tag{XIX.31}$$

$\mathcal{G}(\mathbf{r}, \mathbf{r}')$ is a free-particle Green's function for the energy $E = \hbar^2k^2/2m$. The complex conjugate function has the the same property.

The Schrödinger equation for $\psi_a^{(+)}$ may be written

$$\frac{\hbar^2}{2m}(\Delta + k^2)\,\psi = V\psi. \tag{XIX.32}$$

Suppose for the moment that the function $F(\mathbf{r}) \equiv V(\mathbf{r})\,\psi(\mathbf{r})$ is known. Equation (XIX.32) will then be an inhomogeneous partial differential equation for ψ. It follows from (XIX.31) that the function

$$\hat{\psi}(\mathbf{r}) \equiv \int \mathcal{G}(\mathbf{r}, \mathbf{r}')\,F(\mathbf{r}')\,d\mathbf{r}'$$

is a solution of this equation. The general solution is obtained by adding to $\hat{\psi}(\mathbf{r})$ the general solution of the homogeneous equation. In other words, if ψ obeys eq. (XIX.32), $\psi - \hat{\psi}$ is a free wave of the same energy:

$$\frac{\hbar^2}{2m}(\Delta + k^2)(\psi - \hat{\psi}) = 0 \tag{XIX.33}$$

and conversely.

[1]) Considered as a distribution (in the sense of Schwartz) e^{ikr}/r is differentiable in all orders. To obtain its Laplacian, we use the identity

$$\Delta(1/r) = -4\pi\delta(\mathbf{r})$$

and the property

$$\Delta(fg) = (\Delta f)g + 2(\nabla f)\cdot(\nabla g) + f\Delta g.$$

To completely determine $\psi - \hat{\psi}$ we need only to find its asymptotic form. We first show that $\hat{\psi}$ tends asymptotically to an outgoing wave. Substituting in (XIX.30) the asymptotic expression

$$|\boldsymbol{r}-\boldsymbol{r}'| \equiv r\left[1-2\frac{\boldsymbol{r}\cdot\boldsymbol{r}'}{r^2}+\frac{r'^2}{r^2}\right]^{\frac{1}{2}} \underset{r\to\infty}{\sim} r-\frac{\boldsymbol{r}}{r}\cdot\boldsymbol{r}'+O\left(\frac{r'}{r}\right),$$

we have, for $r \gg r'$,

$$\frac{e^{ik|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|} \underset{r\to\infty}{\sim} \frac{e^{ikr}}{r}\, e^{-i\boldsymbol{k}\cdot\boldsymbol{r}'}+O\left(\frac{r'}{r}\right), \qquad (\boldsymbol{k}=k\boldsymbol{r}/r) \tag{XIX.34}$$

from which

$$\hat{\psi}(\boldsymbol{r}) \underset{r\to\infty}{\sim} -\frac{m}{2\pi\hbar^2}\frac{e^{ikr}}{r}\int e^{-i\boldsymbol{k}\cdot\boldsymbol{r}'}\, V(\boldsymbol{r}')\,\psi(\boldsymbol{r}')\,d\boldsymbol{r}'. \tag{XIX.35}$$

This expression is valid when r is very much larger than the range of the potential ($r \gg a$). Consequently, if the asymptotic form of ψ is given by (XIX.3), then $\psi - \hat{\psi}$ will also have the asymptotic form ($e^{i\boldsymbol{k}_a\cdot\boldsymbol{r}}$ + outgoing wave). Since it also satisfies eq. (XIX.33) it must be the plane wave $e^{i\boldsymbol{k}_a\cdot\boldsymbol{r}}$. Thus

$$\psi_a^{(+)} = e^{i\boldsymbol{k}_a\cdot\boldsymbol{r}} - \frac{m}{2\pi\hbar^2}\int \frac{e^{ik|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|}\, V(\boldsymbol{r}')\,\psi_a^{(+)}(\boldsymbol{r}')\,d\boldsymbol{r}'. \tag{XIX.36}$$

Eq. (XIX.36) is equivalent to (XIX.3). It is called the *scattering integral equation.*

With the aid of (XIX.34) the asymptotic behavior of $\psi_a^{(+)}$ may be read directly from the right-hand side of (XIX.36); property (XIX.15) can be proved in this way.

The preceding analysis may be repeated using in place of $\mathscr{G}(\boldsymbol{r}, \boldsymbol{r}')$ its complex conjugate $\mathscr{G}^*(\boldsymbol{r}, \boldsymbol{r}')$. The latter is also a free-particle Green's function, but corresponds to an incoming asymptotic behavior. This procedure leads to the integral equation:

$$\psi_a^{(-)} = e^{i\boldsymbol{k}_a\cdot\boldsymbol{r}} - \frac{m}{2\pi\hbar^2}\int \frac{e^{-ik|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|}\, V(\boldsymbol{r}')\,\psi_a^{(-)}(\boldsymbol{r}')\,d\boldsymbol{r}'. \tag{XIX.37}$$

6. The Born Expansion

Equations (XIX.36) and (XIX.37) can be solved by iteration. The method resembles the one used to solve eq. (XVII.13) of the time-dependent perturbation theory.

Consider eq. (XIX.36). Substituting the plane wave $e^{ik_a \cdot r}$ for $\psi_a^{(+)}$ in the integral on the right-hand side, we obtain

$$\psi_1^{(+)} = e^{ik_a \cdot r} + \int \mathcal{G}(r, r')\, e^{ik_a \cdot r'}\, dr'.$$

Subsequently substituting $\psi_1^{(+)}$, we obtain

$$\psi_2^{(+)} = e^{ik_a \cdot r} + \int \mathcal{G}(r, r')\, \psi_1^{(+)}(r')\, dr'.$$

Substituting $\psi_2^{(+)}$, we obtain a new wave $\psi_3^{(+)}$, and so on. Under suitable conditions, the series given by this procedure converges to the exact solution $\psi_a^{(+)}$, which is then represented by an expansion in powers of V, namely

$$\psi_a^{(+)} = e^{ik_a \cdot r} + \sum_{n=1}^{\infty} \int K_n(r, r')\, e^{ik_a \cdot r'}\, dr', \qquad \text{(XIX.38)}$$

where

$$K_n(r, r') = \int K_1(r, r'')\, K_{n-1}(r'', r')\, dr'' \qquad (n > 1)$$
$$K_1(r, r') = \mathcal{G}(r, r')\, V(r').$$

This is the *Born expansion* of the stationary scattering wave.

We can obtain an analogous expansion for $T_{a \to b}$ by substituting the above expansion of $\psi_a^{(+)}$ into the integral $\langle \varphi_b | V | \psi_a^{(+)} \rangle$. It is an expansion in powers of V and converges for V sufficiently small [1]). The Born approximation consists in keeping only the first term.

7. Validity Criterion for the Born Approximation

For the Born approximation to be justified the error made in calculating $T_{a \to b}$ by replacing $\psi_a^{(+)}$ by the plane wave φ_a must be negligible. We must therefore have the exact stationary wave differing little from the plane wave in the domain of the potential V. Putting

$$\psi'(r) \equiv \psi_a^{(+)}(r) - e^{ik_a \cdot r},$$

---

[1]) For the study of the convergence of the Born approximation see R. Jost and A. Pais, Phys. Rev. 82 (1951) 840; W. Kohn, Rev. Mod. Phys. 26 (1954) 292. The integral equation that occurs in the time-dependent perturbation method (Ch. XVII) is of the Volterra type; if $U^{(0)}$ and V were functions, and not operators, the expansion of U obtained from it by iteration would always converge; the only possible causes of divergence are due to $U^{(0)}$ and V being operators of Hilbert space. On the other hand, the integral equation with which we are concerned here is of the Fredholm type, and the Born expansion does not always converge.

we obtain the condition

$$|\psi'(\boldsymbol{r})| \ll 1 \tag{XIX.39}$$

(at any point $\boldsymbol{r}$ where V is relatively large).

We can estimate ψ' by taking the first term in its Born expansion. Taking the z axis along $\hat{\boldsymbol{k}}_a$ we thereby obtain after an obvious change of variable

$$\psi'(\boldsymbol{r}) \simeq -\frac{m}{2\pi\hbar^2} e^{ikz} I(\boldsymbol{r})$$

$$I(\boldsymbol{r}) \equiv \int e^{ik(R+Z)} V(\boldsymbol{R}+\boldsymbol{r}) \, d\boldsymbol{R}/R.$$

Due to the presence of k in the exponential, $|I|$ depends on the energy. However, we can obtain an energy-independent upper limit by replacing the quantity under the integral sign by its modulus:

$$|I| < \int |V(\boldsymbol{R}+\boldsymbol{r})| \, d\boldsymbol{R}/R. \tag{XIX.40}$$

Denote the range of the potential by a [1]) and its average depth by V_0; in the relevant domain of the variable $\boldsymbol{r}$ this upper limit is approximately $2\pi V_0 a^2$, so that the Born approximation is justified at all energies if

$$V_0 \ll \hbar^2/ma^2. \tag{XIX.41}$$

This condition is obviously too restrictive. The integral $|I|$ attains the limit defined by (XIX.40) only for energy sufficiently low that the factor $e^{ik(R+Z)}$ is practically constant in the region of the potential [2]) i.e. when $ka \ll 1$. If instead we have $ka \gg 1$, this factor oscillates rapidly over the region of the potential, and the preceding evaluation is much too pessimistic. When $k \to \infty$ these oscillations become increasingly rapid and the integral $|I|$ tends to zero. We may therefore expect the Born approximation to be valid for sufficiently high energies.

The asymptotic behavior of I at high energies may be obtained by the method of stationary phases. The main contribution comes from the region surrounding the points where the oscillating phase factor is stationary, i.e. the region surrounding the half-line

---

[1]) Cf. note, p. 809.

[2]) If the sign of V is constant. To be exact, we have

$$I(\boldsymbol{r}) \underset{k\to 0}{\sim} \int V(\boldsymbol{R}+\boldsymbol{r}) \, d\boldsymbol{R}/R.$$

$X = Y = 0, Z < 0$. The details of the calculation will not be given here [1]). One obtains:

$$I \underset{k\to\infty}{\sim} \frac{2\pi i}{k} \int_{-\infty}^{z} V(x, y, z')\, dz' + O(k^{-2}). \qquad \text{(XIX.42)}$$

Thus in the domain of r in which we are interested, $|I|$ is of order $2\pi V_0 a/k$ in this limit. Criterion (XIX.39) is therefore satisfied if we have

$$ka \gg 1 \qquad V_0 \ll \hbar^2 k/ma. \qquad \text{(XIX.43)}$$

In order to obtain a validity criterion for the Born approximation at intermediate energies when (XIX.41) is not satisfied, we must evaluate the integral I for each individual case. In practice, it is usually evaluated for a single suitably chosen point in the potential region, $r = 0$ say. With a central potential we find, after integration over the angles,

$$I(0) = \frac{2\pi i}{k} \int_0^{\infty} (1 - e^{2ikr})\, V(r)\, dr,$$

in good agreement with properties (XIX.40) and (XIX.42); whence the criterion

$$\left| \int_0^{\infty} (1 - e^{2ikr})\, V(r)\, dr \right| \ll \hbar^2 k/m. \qquad \text{(XIX.44)}$$

In a collision, a characteristic magnitude relating to the cross section is the *geometrical cross section* $4\pi a^2$ corresponding to the range a of the forces between the particles. When the Born approximation is valid, the total cross section is small compared to this quantity:

$$\sigma_{tot} \ll \text{geometric cross section.} \qquad \text{(XIX.45)}$$

Let us demonstrate this inequality for the two limiting cases $ka \ll 1$ and $ka \gg 1$. To this effect we make use of expression (XIX.28) for the total cross section. According to its definition [eq. (XIX.25)], the function $\mathscr{V}(q)$ is of order $(2\pi)^{3/2} V_0 a$ in a domain of linear dimensions $1/a$ surrounding the origin, and practically null elsewhere. Substituting these results in the right-hand side of (XIX.28), we obtain

$$\sigma_{tot}^{(B)}/4\pi a^2 \approx 2\pi (V_0 ma^2/\hbar^2)^2 \qquad \text{if } ka \ll 1$$
$$\sigma_{tot}^{(B)}/4\pi a^2 \approx \tfrac{1}{2}\pi (V_0 ma/\hbar^2 k)^2 \qquad \text{if } ka \gg 1.$$

---

[1]) See, for example: L. I. Schiff, Phys. Rev. **103** (1956) 443.

The conditions for the validity of the Born approximation are inequalities (XIX.41) in the first case, (XIX.43) in the second. In both cases: $\sigma_{tot}^{(B)}/4\pi a^2 \ll 1$.

Inequality (XIX.45) is a necessary, but not a sufficient condition for the validity of the Born approximation. It sometimes happens that due to resonance effects – such as the Ramsauer–Townsend effect in the scattering of slow electrons by atoms – the cross section is much smaller than the geometric cross section without the Born expansion having a rapid convergence. In actual practice, resonant effects can easily be recognized by their extreme sensitivity to variations of the incident energy. With these qualifications, inequality (XIX.45) is a most useful criterion for judging the validity of the Born approximation.

8. Elastic Scattering of Electrons by an Atom

As an application, let us consider the elastic scattering of a charged particle by an atom. In the simplified treatment to be given here the atom is treated as a distribution of electric charges, and we calculate in the Born approximation the cross section for the scattering of the charged particle in the potential created by this distribution. This method actually gives the correct cross section if we are in the region where the Born approximation is valid (see § 21). To be definite we shall suppose that the incident particle is an electron. We should then normally have to consider the exchange effect between the incident electron and the electrons of the atom, but since at energies where the Born approximation is valid this gives only a small correction, such effects will be ignored.

The atom, by hypothesis, is neutral. We denote the atomic number by Z and the density of the electrons by $\varrho(\boldsymbol{r})$:

$$\int \varrho \, \mathrm{d}\boldsymbol{r} = Z.$$

The electric potential satisfies the Poisson equation

$$\Delta\varphi = -4\pi e[Z\,\delta(\boldsymbol{r}) - \varrho(r)]. \qquad \text{(XIX.46)}$$

The electron is subject to the potential

$$V(r) = -e\varphi(r),$$

which statisfies an analogous equation whose Fourier transform gives the relation

$$q^2\mathscr{V}(q) = 4\pi e^2[Z - F(q)] \qquad \text{(XIX.47)}$$

between the function $\mathscr{V}(q)$ [definition (XIX.25) or (XIX.27)] and the function

$$F(q) \equiv \int e^{-i\mathbf{q}\cdot\mathbf{r}}\,\varrho(r)\,d\mathbf{r} = q^{-1}\int_0^\infty \sin qr\,\varrho(r)\,r dr. \qquad \text{(XIX.48)}$$

With the aid of (XIX.47), the cross section in the Born approximation [eq. (XIX.26)] is

$$\frac{d\sigma}{d\Omega} = \frac{4[Z - F(q)]^2\,m^2\,e^4}{\hbar^4\,q^4}. \qquad \text{(XIX.49)}$$

$F(q)$ is called the *form factor* for the electronic density. Its general form may easily be deduced from that of $\varrho(r)$. In particular, we have

$$F(0) = Z.$$

If a is the radius of the atom, i.e. the average distance of the electrons from the nucleus, the factor F has appreciable values only in the region $q \lesssim 1/a$.

If, for example, the electronic density is represented by the function

$$\varrho(r) = \frac{Ze^{-r/a}}{2a^3},$$

the form factor is given by

$$F(q) = Z[1 + q^2a^2]^{-2}.$$

When $qa \gg 1$, i.e. for large scattering angles:

$$\sin\tfrac{1}{2}\theta \gg \frac{1}{2ka},$$

the form factor is practically null [$F(q) \ll Z$] and formula (XIX.49) reduces essentially to the Rutherford formula giving the cross section for scattering by the nucleus alone: the effect of the electrons is negligible.

On the other hand, their effect becomes appreciable once $qa \lesssim 1$: the screening effect is therefore appreciable for small angles, i.e. for

$$\sin\tfrac{1}{2}\theta \lesssim \frac{1}{2ka}. \qquad \text{(XIX.50)}$$

This is in agreement with the classical picture in which small scattering angles correspond to large impact parameters.

To conclude, let us examine the validity of the Born approximation.

The range of the potential is of the order of a. An estimate based on the Thomas–Fermi model gives

$$a \simeq \frac{Z^{-\frac{1}{3}} \hbar^2}{me^2}.$$

For the average value of the potential we may take the value at $r = a$ of the nuclear Coulomb potential:

$$V_0 \simeq \frac{Ze^2}{a}.$$

In what follows we use the familiar notations

$$v = \frac{\hbar k}{m}, \qquad \gamma = \frac{Ze^2}{\hbar v};$$

whence

$$ka = Z^{\frac{1}{3}} \gamma^{-1}.$$

Criterion (XIX.41) requires that $Z^{\frac{1}{3}} \ll 1$ and is therefore never realized. Hence the Born approximation is justified only for sufficiently high energies. In the high-energy domain ($ka \gg 1$), that is for $\gamma \ll Z^{\frac{1}{3}}$, criterion (XIX.43) requires $\gamma \ll 1$.

In summary, the Born approximation is justified only at energies sufficiently high that

$$\gamma \ll 1.$$

In this case the screening effect is appreciable only at very small angles, and becomes practically negligible outside the region defined by (XIX.50), i.e. outside the region

$$\theta \lesssim \gamma.$$

The form factor can be experimentally determined by making precise measurements of the angular distribution at small angles and high energies, and using formula (XIX.49) to interpret the results.

9. Central Potential. Calculation of Phase Shifts

When we have a central potential the amplitude $T_{a\to b}$ depends only on the scattering angle θ ($\cos\theta = \mathbf{k}_a \cdot \mathbf{k}_b / k^2$) and may be expanded as follows:

$$T_{a\to b} = 16\pi^2 \sum_{lm} T_l \, Y_l^{m*}(\hat{\mathbf{k}}_b) \, Y_l^m(\hat{\mathbf{k}}_a) \qquad \text{(XIX.51)}$$
$$= 4\pi \sum_l (2l+1) \, T_l \, P_l(\cos\theta).$$

The coefficients T_l are easily related to the phase shifts by comparing this expansion with (X.31):

$$T_l = -\hbar^2 \, e^{i\delta_l} \sin \delta_l / 2mk. \tag{XIX.52}$$

On the other hand, expansion (XIX.51) can be directly obtained by expanding the waves φ_b and $\psi_a^{(+)}$ in the integral $\langle \varphi_b | V | \psi_a^{(+)} \rangle$ into spherical harmonics. The angular integrations are easily carried out with the aid of the orthonormality relations for the latter; the coefficients T_l appear as radial integrals. Putting

$$\psi_a^{(+)} = 4\pi \sum_{lm} i^l \psi_l(r) \, Y_l^m(\hat{\mathbf{k}}_a) \, Y_l^{m*}(\hat{\mathbf{r}}), \tag{XIX.53}$$

we find:

$$T_l = \int_0^\infty j_l(kr) \, V(r) \, \psi_l(r) \, r^2 \, dr. \tag{XIX.54}$$

It is easily seen (Problem XIX.4) that:

$$\psi_l \underset{r \to \infty}{\sim} e^{i\delta_l} \sin (kr - \tfrac{1}{2} l\pi + \delta_l)/kr.$$

Comparing relations (XIX.52) and (XIX.54), we easily obtain the integral representation of $\sin \delta_l$ given in § X.17 [eq. (X.73)]; this reduces to the Born formula for the phase shifts [eq. (X.75)] in the limit when the l-wave is sufficiently close to the free wave.

From the scattering integral equation we obtain in the same way an integral equation for ψ_l. For this it is convenient to use the following expansion of the Green's function (Problem XIX.4):

$$\frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} = 4\pi k \sum_{lm} j_l(kr_<) \, h_l^{(+)}(kr_>) \, Y_l^{m*}(\hat{\mathbf{r}}') \, Y_l^m(\hat{\mathbf{r}}). \tag{XIX.55}$$

In this formula $r_<$ and $r_>$ denote respectively the smaller and the larger of the two lengths r and r'. Substituting (XIX.53), (XIX.55) and the expansion of the plane wave into the integral equation (XIX.36), we obtain an integral equation for the outgoing partial wave:

$$\psi_l(r) = j_l(kr) - (2mk/\hbar^2) \int_0^\infty j_l(kr_<) \, h_l^{(+)}(kr_>) \, V(r') \, \psi_l(r') \, r'^2 \, dr'. \tag{XIX.56}$$

Iteration of this equation gives ψ_l as an expansion in powers of V.

10. Green's Function as an Operator. Relation to the Resolvent of H_0

In the preceding paragraphs the Green's function has appeared as a kernel in an integral equation. It may also be thought of as the matrix representative of a certain operator $G_0^{(+)}$ according to the definition:

$$\mathcal{G}(\mathbf{r}, \mathbf{r}') \equiv \langle \mathbf{r}|G_0^{(+)}|\mathbf{r}'\rangle. \tag{XIX.57}$$

Thus, consider a vector $|u\rangle$ and denote by $u(\mathbf{r})$ the corresponding wave function; $u(\mathbf{r}) \equiv \langle \mathbf{r}|u\rangle$; the vector $G_0^{(+)}|u\rangle$ obtained by application of $G_0^{(+)}$ is represented by the function

$$\langle \mathbf{r}|G_0^{(+)}|u\rangle = \int \mathcal{G}(\mathbf{r}, \mathbf{r}')\, u(\mathbf{r}')\, d\mathbf{r}'. \tag{XIX.58}$$

We similarly define the operator $G_0^{(-)}$ by

$$\mathcal{G}^*(\mathbf{r}, \mathbf{r}') \equiv \langle \mathbf{r}|G_0^{(-)}|\mathbf{r}'\rangle. \tag{XIX.59}$$

Since $\mathcal{G}(\mathbf{r}, \mathbf{r}')$ is symmetrical in $\mathbf{r}$ and $\mathbf{r}'$, one has

$$G_0^{(-)} = G_0^{(+)\dagger}. \tag{XIX.60}$$

The index 0 indicates that these operators are related to the free-particle Hamiltonian H_0. The $(+)$ and $(-)$ indicate the respective asymptotic behavior, which from the asymptotic form of the Green's function [eq. (XIX.34)] is seen to be

$$\langle \mathbf{r}|G_0^{(\pm)}|u\rangle \underset{r\to\infty}{\sim} -\frac{m}{2\pi\hbar^2}\frac{e^{\pm ikr}}{r}\int e^{\mp i\mathbf{k}\cdot\mathbf{r}'}\, u(\mathbf{r}')\, d\mathbf{r}'. \tag{XIX.61}$$
$$(\mathbf{k} = k\mathbf{r}/r)$$

This result is valid for any vector $|u\rangle$ of finite norm (the latter condition guarantees the convergence of the integral on the right-hand side).

Note the extreme singularity of these operators. Unless the Fourier components of $u(\mathbf{r})$ relative to propagation vectors of length k all vanish, the fall-off of $\langle \mathbf{r}|G_0^{(\pm)}|u\rangle$ at infinity is not sufficiently rapid for it to be a square-integrable function. In other words, the application of $G_0^{(+)}$ or $G_0^{(-)}$ to a vector of Hilbert space gives a vector in general outside this space. Therefore, $G_0^{(+)}$ and $G_0^{(-)}$ are not, properly speaking, operators in Hilbert space.

They can, however, be defined as limits of operators in Hilbert space as follows:

$$G_0^{(\pm)} = \lim_{\varepsilon\to 0^+} \frac{1}{E - H_0 \pm i\varepsilon} \qquad (\varepsilon > 0). \tag{XIX.62}$$

On the right-hand side will be recognized the resolvent $(z-H_0)^{-1}$ (cf. Ch. XVI, Sec. III); it is a (bounded) operator of Hilbert space for all values of the complex variable z except for the eigenvalues of H_0, i.e. except for points on the positive real axis. The behavior of the resolvent in the neighborhood of this line is given by (XIX.62). If z tends to E from above the real axis ($\mathrm{Im}\, z>0$), $(z-H_0)^{-1}$ tends toward $G_0^{(+)}$; if it tends to E from below the real axis ($\mathrm{Im}\, z<0$), $(z-H_0)^{-1}$ tends toward $G_0^{(-)}$.

The demonstration of (XIX.62) proceeds as follows. Starting from the (diagonal) matrix of $(z-H_0)^{-1}$ in the $\{\boldsymbol{p}\}$ representation, we apply the well-known unitary transformation to obtain its matrix in the $\{\boldsymbol{r}\}$ representation:

$$\langle \boldsymbol{r}|(z-H_0)^{-1}|\boldsymbol{r}'\rangle = \frac{1}{(2\pi)^3}\int \mathrm{e}^{\mathrm{i}\boldsymbol{k}'\cdot(\boldsymbol{r}'-\boldsymbol{r})}\left(z-\frac{\hbar^2 k'^2}{2m}\right)^{-1}\mathrm{d}\boldsymbol{k}'.$$

The angular integrations are easily carried out. Introducing the quantity ζ defined by

$$z=\hbar^2\zeta^2/2m \qquad \mathrm{Re}\,\zeta>0$$

and writing

$$\boldsymbol{R}=\boldsymbol{r}-\boldsymbol{r}',$$

we find:

$$\langle \boldsymbol{r}|(z-H_0)^{-1}|\boldsymbol{r}'\rangle = \frac{m}{\pi^2\hbar^2 R}\int_0^\infty \frac{k'\sin k'R}{\zeta^2-k'^2}\,\mathrm{d}k'$$

$$= \frac{m}{\pi\hbar^2 R}\frac{1}{2\pi\mathrm{i}}\int_{-\infty}^{+\infty}\frac{k'\,\mathrm{e}^{\mathrm{i}k'R}}{\zeta^2-k'^2}\,\mathrm{d}k'.$$

The value of this last integral is not modified if we close the path of integration with a semi-circular contour at infinity in the upper half plane. The contour integral thus obtained is equal to $2\pi\mathrm{i}$ times the sum of the residues of the poles in the upper half-plane; in the present case there is just one pole, whose position depends on the sign of $\mathrm{Im}\,\zeta$, and therefore on the sign of $\mathrm{Im}\, z$ (Fig. XIX.2):

(*a*) $\mathrm{Im}\, z>0$. Pole for $k'=\zeta$ with residue $-\frac{1}{2}\mathrm{e}^{\mathrm{i}\zeta R}$;

(*b*) $\mathrm{Im}\, z<0$. Pole for $k'=-\zeta$ with residue $-\frac{1}{2}\mathrm{e}^{-\mathrm{i}\zeta R}$.

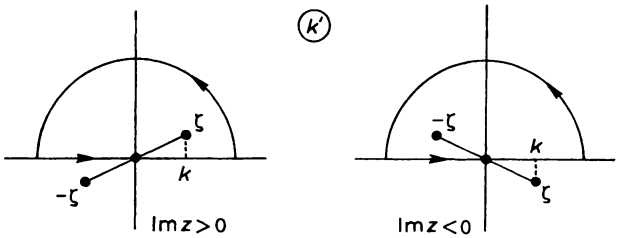


Fig. XIX.2.

Consequently:

$$\langle \boldsymbol{r}|(z-H)^{-1}|\boldsymbol{r}'\rangle = -\frac{m}{2\pi\hbar^2} \times \begin{cases} e^{i\zeta R}/R & \text{if} \quad \text{Im } z > 0 \\ e^{-i\zeta R}/R & \text{if} \quad \text{Im } z < 0. \end{cases}$$

These expressions are to be compared with those for $\mathscr{G}(\boldsymbol{r}, \boldsymbol{r}')$ [eq. (XIX.30)]. In the limit when z tends to E, ζ tends to k and the matrix element calculated above tends to $\mathscr{G}(\boldsymbol{r}, \boldsymbol{r}')$ or to $\mathscr{G}^*(\boldsymbol{r}, \boldsymbol{r}')$ according as the imaginary part of z is positive or negative in the passage to the limit. Q.E.D.

Operators are often more easily manipulated algebraically than their representative matrices. Thus there is frequently advantage in using the operators $G_0^{(+)}$ and $G_0^{(-)}$ rather than the Green's functions themselves. Nevertheless, these are singular operators, and their manipulation requires caution, and a certain flair [1]). In all rigor, they should be replaced by the regular operators $(E - H_0 \pm i\varepsilon)^{-1}$, and the passage to the limit $\varepsilon \to 0$ examined with care. These details of rigor will not be dwelt upon in what follows. When the $G_0^{(\pm)}$ are replaced by the operators $(E - H_0 \pm i\varepsilon)^{-1}$ it is always to be understood that

[1]) In particular, care must be taken that G acts only on vectors of *finite norm*. The kind of pitfall to which one is exposed if one forgets the singular character of these operators is illustrated by the following example. From relation (XIX.31) and the symmetry of the Green's function in $\boldsymbol{r}$ and $\boldsymbol{r}'$ we have

$$(E - H_0)\, G_0^{(\pm)} = 1, \qquad G_0^{(\pm)}(E - H_0) = 1.$$

These relations mean that the action of the left-hand side of each of these equations on a vector of finite norm gives the same vector, and nothing more. However, taking them literaly, one would be tempted to conclude that $E - H_0$ has an inverse, which is obviously false, and that $G_0^{(+)}$ and $G_0^{(-)}$ are both equal to that inverse and therefore equal to each other, which is also false.

ε is a real positive quantity, and that the formulas are to be considered in the limit $\varepsilon \to 0^+$.

We may now write the integral equations (XIX.36) and (XIX.37) in the form

$$\psi_a^{(\pm)} = \varphi_a + G_0^{(\pm)} V \psi_a^{(\pm)} \tag{XIX.63}$$

or

$$\psi_a^{(\pm)} = \varphi_a + \frac{1}{E - H_0 \pm i\varepsilon} V \psi_a^{(\pm)}. \tag{XIX.63'}$$

By iteration of these equations, we obtain the Born expansions:

$$\psi_a^{(\pm)} = \left[1 + \sum_{n=1}^{\infty} (G_0^{(\pm)} V)^n\right] \varphi_a. \tag{XIX.64}$$

In these expressions the symbols ψ and φ denote either the stationary waves themselves or the kets they represent (no confusion being possible, we drop the symbols $|\rangle$).

From these we obtain the conjugate equations by replacing the kets on both sides by their respective conjugate bras. Since H_0 is Hermitean, we have

$$G_0^{(+)\dagger} = G_0^{(-)} \qquad G_0^{(-)\dagger} = G_0^{(+)}. \tag{XIX.65}$$

From (XIX.63),

$$\langle \psi_a^{(\pm)}| = \langle \varphi_a| + \langle \psi_a^{(\pm)}| V G_0^{(\mp)} \tag{XIX.66}$$

and from (XIX.64) we have the Born expansions:

$$\langle \psi_a^{(\pm)}| = \langle \varphi_a| \left[1 + \sum_{n=1}^{\infty} (V G_0^{(\mp)})^n\right]. \tag{XIX.67}$$

II. GENERALIZATION TO DISTORTED WAVES

11. Generalized Born Approximation

It may happen that the potential $V(r)$ is too large for the Born approximation to be used but that we can exactly solve the scattering problem for a neighboring potential V. It is then useful to treat the difference as a perturbation, which may easily be done by a simple generalization of the treatment of Section I.

Let us put

$$\begin{aligned} V(r) &= U_1(r) + W_1(r) \\ H_1 &= H_0 + U_1(r). \end{aligned} \tag{XIX.68}$$

We therefore have:

$$H = H_1 + W_1. \tag{XIX.69}$$

H_1 is the unperturbed Hamiltonian, W_1 the perturbation. We shall suppose that U_1 and V tend to 0 more rapidly than $1/r$; the generalization to $1/r$ potentials will be made in § 15. The stationary states of H_1 are assumed to be known. They will be denoted by the letter χ and the corresponding amplitudes by g; for the rest we follow the notation of § 2. $\chi_a^{(+)}$ and $g_a^{(+)}$ are defined by

$$H_1 \chi_a^{(+)} = E \chi_a^{(+)} \qquad \chi_a^{(+)} \underset{r\to\infty}{\sim} e^{ik_a\cdot r} + g_a^{(+)}(\Omega)\, e^{ikr}/r.$$

All of the results of Section I relating to the Hamiltonian H are obviously true for H_1 as well. In particular we can define a transition matrix for collisions at energy E governed by the Hamiltonian H_1, which we shall denote by $T^{(1)}$

$$T^{(1)}_{a\to b} \equiv \langle \varphi_b | T^{(1)} | \varphi_a \rangle = \langle \varphi_b | U_1 | \chi_a^{(+)} \rangle = \langle \chi_b^{(-)} | U_1 | \varphi_a \rangle. \tag{XIX.70}$$

Among the results of § 2, recall that:

$$g_b^{(-)*}(-\Omega_a) = g_a^{(+)}(\Omega_b) = -\frac{m}{2\pi\hbar^2} T^{(1)}_{a\to b}.$$

From (XIX.9) we can obtain an integral representation of the difference $T_{a\to b} - T^{(1)}_{a\to b}$. Taking

$$\hat{U} = U_1, \qquad U = V,$$

we obtain:

$$\begin{aligned} \langle \chi_b^{(-)} | W_1 | \psi_a^{(+)} \rangle &= -\frac{2\pi\hbar^2}{m} [f_a^{(+)}(\Omega_b) - g_b^{(-)*}(-\Omega_a)] \\ &= -\frac{2\pi\hbar^2}{m} [f_a^{(+)}(\Omega_b) - g_a^{(+)}(\Omega_b)] \\ &= T_{a\to b} - T^{(1)}_{a\to b}. \end{aligned} \tag{XIX.71}$$

In the same way, by taking $\hat{U} = V$ and $U = U_1$, we obtain:

$$\langle \psi_b^{(-)} | W_1 | \chi_a^{(+)} \rangle = T_{a\to b} - T^{(1)}_{a\to b}. \tag{XIX.72}$$

Note in passing the relation

$$\langle \psi_b^{(-)} | W_1 | \chi_a^{(+)} \rangle = \langle \chi_b^{(-)} | W_1 | \psi_a^{(+)} \rangle, \tag{XIX.73}$$

an obvious consequence of (XIX.71) and (XIX.72) and a generalization of relation (XIX.18).

Either of these two integral representations can be used as a starting-point for the perturbation treatment. Let us take (XIX.71), which gives:

$$\boxed{T_{a\to b} = T^{(1)}_{a\to b} + \langle \chi_b^{(-)} | W_1 | \psi_a^{(+)} \rangle} \qquad \text{(XIX.74)}$$

This expression for the transition amplitude is exact. The first term is what it would be in the absence of W_1, the second the correction due to the latter. For W_1 sufficiently small, the exact stationary wave, $\psi_a^{(+)}$, differs little from $\chi_a^{(+)}$, and may be replaced by the latter in the correction term. We thereby obtain the transition amplitude to the first order in the perturbation W_1:

$$T_{a\to b} \simeq T^{(1)}_{a\to b} + \langle \chi_b^{(-)} | W_1 | \chi_a^{(+)} \rangle. \qquad \text{(XIX.75)}$$

This is the generalization of the Born approximation [cf. eq. (XIX.23)]. Note the simultaneous presence, and respective position, of the stationary solutions corresponding to outgoing and incoming waves in the perturbation integral:

$$\langle \chi_b^{(-)} | W_1 | \chi_a^{(+)} \rangle \equiv \int \chi_b^{(-)*}(\boldsymbol{r})\, W_1(\boldsymbol{r})\, \chi_a^{(+)}(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r}. \qquad \text{(XIX.76)}$$

If we take $U_1 = 0$ we obtain the Born result. In the general case, when $U_1 \neq 0$, it is obviously more difficult to calculate this integral than the one appearing in the Born formula. The situation becomes somewhat simpler when U_1 and V are both spherically symmetrical. We may then use the expansions (Problem XIX.4):

$$\chi_{\boldsymbol{k}}^{(\pm)} = \frac{4\pi}{kr} \sum_{lm} \mathrm{i}^l\, \mathrm{e}^{\pm \mathrm{i}\eta_l}\, F_l(k; r)\, Y_l^{m*}(\hat{\boldsymbol{k}})\, Y_l^m(\hat{\boldsymbol{r}}), \qquad \text{(XIX.77)}$$

where $F_l(k; r)$ denotes the regular solution of the radial equation

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}r^2} - \frac{l(l+1)}{r^2} - \frac{2m}{\hbar^2} U_1 + k^2\right] y = 0, \qquad \text{(XIX.78)}$$

whose asymptotic behavior is given by

$$F_l(k; r) \underset{r\to\infty}{\sim} \sin\left(kr - \tfrac{1}{2}l\pi + \eta_l\right). \qquad \text{(XIX.79)}$$

Here η_l is the phase shift of the l-wave in the potential U_1. Substituting

the expansions of $\chi_a^{(+)}$ and $\chi_b^{(-)}$ into the integral (XIX.76) and integrating over the angles, we find

$$\langle \chi_b^{(-)} | W_1 | \chi_a^{(+)} \rangle = \frac{4\pi}{k^2} \sum_l (2l+1)\, P_l(\cos\theta)\, e^{2i\eta_l} \int_0^\infty F_l^2(k; r)\, W_1(r)\, dr.$$

The convergence of this expansion is obviously better the smaller the range of W_1

12. Generalization of the Born Expansion

To find the higher order corrections we must expand $\psi_a^{(+)}$ into powers of W_1. In analogy with the method of section I we do this by finding an integral equation for $\psi_a^{(+)}$ that can be iterated to give the desired expansion.

Denote by $G_1^{(+)}(\boldsymbol{r}, \boldsymbol{r}')$ the Green's function for the Hamiltonian H_1 corresponding to the energy $E = \hbar^2 k^2/2m$ and having the outgoing-wave asymptotic behavior. By definition, it is a symmetrical function of $\boldsymbol{r}$ and $\boldsymbol{r}'$ obeying the partial differential equation

$$(E - H_1)\, G_1^{(+)} \equiv \left[\frac{\hbar^2}{2m}(\Delta + k^2) - U_1\right] G_1^{(+)}(\boldsymbol{r}, \boldsymbol{r}') = \delta(\boldsymbol{r}' - \boldsymbol{r}) \quad \text{(XIX.80)}$$

and behaving asymptotically like e^{ikr}/r when $r \to \infty$, r' being kept fixed. It is the representative matrix of a certain operator $G_1^{(+)}$ according to the definition:

$$\langle \boldsymbol{r} | G_1^{(+)} | \boldsymbol{r}' \rangle \equiv G_1^{(+)}(\boldsymbol{r}, \boldsymbol{r}').$$

It will be shown in the next paragraph that there always exists one, and only one function $G_1^{(+)}(\boldsymbol{r}, \boldsymbol{r}')$ having these properties.

We obtain the integral equation by the method of § 5 using this function in place of the free-wave Green's function. We write the Schrödinger equation for $\psi_a^{(+)}$ in the form:

$$(E - H_1)\, \psi_a^{(+)} \equiv \left[\frac{\hbar^2}{2m}(\Delta + k^2) - U_1\right] \psi_a^{(+)} = W_1\, \psi_a^{(+)}$$

and note that in view of eq. (XIX.80) the function

$$\psi_a^{(+)} - \int G_1^{(+)}(\boldsymbol{r}, \boldsymbol{r}')\, W_1(\boldsymbol{r}')\, \psi_a^{(+)}(\boldsymbol{r}')\, d\boldsymbol{r}'$$

satisfies the "homogeneous equation"

$$(E - H_1)\chi = 0.$$

Since this function has the same asymptotic form as $\chi_a^{(+)}$, it is necessarily equal to $\chi_a^{(+)}$. $\psi_a^{(+)}$ therefore satisfies the integral equation

$$\psi_a^{(+)} = \chi_a^{(+)}(\boldsymbol{r}) + \int G_1^{(+)}(\boldsymbol{r}, \boldsymbol{r}')\, W_1(\boldsymbol{r}')\, \psi_a^{(+)}(\boldsymbol{r}')\, \mathrm{d}\boldsymbol{r}'. \quad \text{(XIX.81)}$$

The expansion of $\psi_a^{(+)}$ in powers of W_1 is easily obtained from eq. (XIX.81). Substituting this expansion into the "matrix element" in the right-hand side of (XIX.74) we obtain the expansion of $T_{a\to b}$ in powers of W_1. If we keep only the first two terms of this expansion we obtain the approximate expression (XIX.75).

The complex conjugate of $G_1^{(+)}(\boldsymbol{r}, \boldsymbol{r}')$ is also a Green's function of H_1 for the energy E: it is the function $G_1^{(-)}(\boldsymbol{r}, \boldsymbol{r}')$ which asymptotically has the incoming-wave behavior. From these properties we deduce by a similar method the integral equation

$$\psi_a^{(-)} = \chi_a^{(-)} + \int G_1^{(-)}(\boldsymbol{r}, \boldsymbol{r}')\, W_1(\boldsymbol{r}')\, \psi_a^{(-)}(\boldsymbol{r}')\, \mathrm{d}\boldsymbol{r}'. \quad \text{(XIX.82)}$$

We could also obtain the expansion of $T_{a\to b}$ in powers of W_1 by starting from eq. (XIX.82), and using (XIX.73); the result would obviously be the same.

13. Green's Functions for Distorted Waves

The existence of the Green's functions $G_1^{(+)}$ and $G_1^{(-)}$ remains to be demonstrated. It will be shown that by a natural generalization of eq. (XIX.62) they can be defined as limits of the resolvent of H_1, *viz.*

$$G_1^{(\pm)} = \lim_{\varepsilon\to 0+} \frac{1}{E - H_1 \pm \mathrm{i}\varepsilon}. \quad \text{(XIX.83)}$$

Similarly, for each value of E we associate with the Hamiltonian H the functions:

$$G^{(\pm)} = \lim_{\varepsilon\to 0+} \frac{1}{E - H \pm \mathrm{i}\varepsilon}. \quad \text{(XIX.84)}$$

Let us first consider the operators $G^{(+)}$ and $G^{(-)}$ and show that

$$(E - H)G^{(\pm)} = 1, \qquad G^{(\pm)}(E - H) = 1. \quad \text{(XIX.85)}$$

$|u\rangle$ being any vector of finite norm, we obviously have

$$(E - H)\frac{1}{E - H + \mathrm{i}\varepsilon}|u\rangle = |u\rangle - \frac{\mathrm{i}\varepsilon}{E - H + \mathrm{i}\varepsilon}|u\rangle. \quad \text{(XIX.86)}$$

The norm of the second term on the right-hand side is

$$N(\varepsilon) = \varepsilon^2 \langle u| \frac{1}{(E-H)^2+\varepsilon^2} |u\rangle$$

$$= \varepsilon \int |u(E')|^2 \frac{\varepsilon}{(E-E')^2+\varepsilon^2} \, \mathrm{d}E'.$$

The second line is obtained by adopting a representation in which H is diagonal; $|u(E')|^2 \mathrm{d}E'$ is the norm of the component of $|u\rangle$ of energy E'. By a known property of the δ function [eq. (A.15c)], the integral tends to $\pi|u(E)|^2$ when $\varepsilon \to 0$. Thus the norm $N(\varepsilon)$ vanishes in this limit. Therefore, when $\varepsilon \to 0$ equation (XIX.86) gives:

$$(E - H) G^{(+)} |u\rangle = |u\rangle.$$

This being true for any $|u\rangle$ [*of finite norm* [1])], we have:

$$(E - H) G^{(+)} = 1.$$

The other relations of (XIX.85) may be demonstrated in the same way.

In similar fashion

$$(E - H_1) G_1^{(\pm)} = 1, \qquad G_1^{(\pm)} (E - H_1) = 1. \qquad \text{(XIX.87)}$$

Equations (XIX.85) and (XIX.87) show that $G^{(\pm)}$ and $G_1^{(\pm)}$ are respectively the Green's functions for H and H_1 for the energy E. In the $\{\boldsymbol{r}\}$ representation they give the partial differential equations characteristic of the Green functions: thus the first of eqs. (XIX.87) gives eq. (XIX.80).

In the preceding demonstration, we examined with some care the passage to the limit $\varepsilon \to 0^+$ since it plays an essential role. In what follows, we leave aside these points of rigor and, with the conventions of page 822, shall often replace $G^{(\pm)}$ by $[E-H \pm \mathrm{i}\varepsilon]^{-1}$ and $G_1^{(\pm)}$ by $[E-H_1 \pm \mathrm{i}\varepsilon]^{-1}$, where ε is a real, positive, infinitesimal quantity.

Proceeding as at the beginning of § XVI.16 we obtain the fundamental identities

$$\frac{1}{E-H \pm \mathrm{i}\varepsilon} - \frac{1}{E-H_0 \pm \mathrm{i}\varepsilon} = \frac{1}{E-H \pm \mathrm{i}\varepsilon} V \frac{1}{E-H_0 \pm \mathrm{i}\varepsilon} \qquad \text{(XIX.88}a\text{)}$$

$$= \frac{1}{E-H_0 \pm \mathrm{i}\varepsilon} V \frac{1}{E-H \pm \mathrm{i}\varepsilon}. \qquad \text{(XIX.88}b\text{)}$$

[1]) The remarks of note p. 821 fully apply here.

There are analogous identities with H_1 and U_1 in place of H and V respectively, and also with H_1 and W_1 in place of H_0 and V respectively.

As a first application of identities (XIX.88) we can deduce the relations [1])

$$(1 + G^{(\pm)}V)(1 - G_0^{(\pm)}V) = 1 \qquad \text{(XIX.89}a\text{)}$$

$$(1 - G_0^{(\pm)}V)(1 + G^{(\pm)}V) = 1. \qquad \text{(XIX.89}b\text{)}$$

Now eq. (XIX.63) may be written in the form

$$[1 - G_0^{(\pm)}V]\,\psi_a^{(\pm)} = \varphi_a.$$

Applying the operator $(1 + G^{(\pm)}V)$ to both sides of this equation and using (XIX.89a) to simplify the left-hand side, we obtain the important formula:

$$\psi_a^{(\pm)} = \left(1 + \frac{1}{E - H \pm i\varepsilon}V\right)\varphi_a. \qquad \text{(XIX.90)}$$

Since H is Hermitean, we obviously have

$$G^{(\pm)\dagger} = G^{(\mp)}. \qquad \text{(XIX.91)}$$

The conjugate equation to eq. (XIX.90) is therefore

$$\langle\psi_a^{(\pm)}| = \langle\varphi_a|\left(1 + V\frac{1}{E - H \mp i\varepsilon}\right). \qquad \text{(XIX.92)}$$

We are now in a position to obtain the asymptotic behavior of $G^{(+)}$ and $G^{(-)}$ from that of $G_0^{(+)}$ and $G_0^{(-)}$ respectively. If $|u\rangle$ is a vector with a sufficiently rapid fall-off at infinity, we have from (XIX.88b),

$$G^{(\pm)}|u\rangle = G_0^{(\pm)}|\mathscr{V}\rangle,$$

where

$$|\mathscr{V}\rangle = (1 + VG^{(\pm)})|u\rangle.$$

If the vector $|\mathscr{V}\rangle$ also falls off sufficiently rapidly at infinity [2]), we can apply (XIX.61), which gives

$$\langle\boldsymbol{r}|G^{(\pm)}|u\rangle \underset{r\to\infty}{\sim} -\frac{m}{2\pi\hbar^2}\frac{e^{\pm ikr}}{r}\langle\varphi_{\pm\boldsymbol{k}}|\mathscr{V}\rangle. \qquad \text{(XIX.93)}$$

$$(\boldsymbol{k} = k\boldsymbol{r}/r).$$

---

[1]) From these identities we cannot conclude that $[1 + G^{(\pm)}V]$ and $[1 - G_0^{(\pm)}V]$ are inverses of one another, nor even that each of them has an inverse. Cf. note, p. 821.

[2]) The following deductions are not valid if V behaves like $1/r$ at infinity.

$G^{(+)}$ acting on a vector of Hilbert space will therefore generally give a vector which asymptotically behaves as a purely outgoing wave, $G^{(-)}$ gives a vector which asymptotically behaves like a purely incoming wave. Replacing $|\mathscr{V}\rangle$ on the right-hand side of (XIX.93) by its definition, and using the fact [eq. (XIX.91)] that

$$\langle \varphi_{\pm \boldsymbol{k}}|(1 + VG^{(\pm)}) = \langle \psi^{(\mp)}_{\pm \boldsymbol{k}}|,$$

we find

$$\langle \boldsymbol{r}|G^{(\pm)}|u\rangle \underset{r\to\infty}{\sim} -\frac{m}{2\pi\hbar^2}\frac{\mathrm{e}^{\pm ikr}}{r}\langle \psi^{(\mp)}_{\pm \boldsymbol{k}}|u\rangle. \tag{XIX.94}$$

It is to be observed that it is the wave $\psi^{(-)}_{\boldsymbol{k}}$ that appears in the asymptotic expression for $G^{(+)}$, and $\psi^{(+)}_{-\boldsymbol{k}}$ in that of $G^{(-)}$.

$G_1^{(+)}$ and $G_1^{(-)}$ are found to have analogous properties, and notably the asymptotic behavior

$$\langle \boldsymbol{r}|G_1^{(\pm)}|u\rangle \underset{r\to\infty}{\sim} -\frac{m}{2\pi\hbar^2}\frac{\mathrm{e}^{\pm ikr}}{r}\langle \chi^{(\mp)}_{\pm \boldsymbol{k}}|u\rangle. \tag{XIX.95}$$

$$(\boldsymbol{k} = k\boldsymbol{r}/r).$$

When U_1 is spherically symmetrical, the Green functions $G_1^{(\pm)}$ are easily expressed in terms of the solutions of the radial equation (XIX.78). One finds (Problem XIX.4):

$$\langle \boldsymbol{r}|G_1^{(\pm)}|\boldsymbol{r}'\rangle = -\frac{2m}{\hbar^2 k}\frac{1}{rr'}\sum_{lm}\mathrm{e}^{\pm i\eta_l}F_l(k; r_<)\,u_l^{(\pm)}(k; r_>)\,Y_l^{m*}(\hat{\boldsymbol{r}}')\,Y_l^m(\hat{\boldsymbol{r}}). \tag{XIX.96}$$

Here, $r_<$ and $r_>$ are respectively the smaller and the larger of the two lengths r and r'; η_l and F_l were defined in § 11; $u_l^{(\pm)}$ are the irregular solutions of eq. (XIX.78) having the asymptotic behavior

$$u_l^{(\pm)} \underset{r\to\infty}{\sim} \mathrm{e}^{\pm i(kr - \frac{1}{2}l\pi)}.$$

The explicit forms (XIX.96) are readily shown to have the characteristic properties of $G_1^{(+)}$ and $G_1^{(-)}$.

14. Applications. Definition and Formal Properties of T

By using the formal properties of the Green's functions given in § 13 we can obtain most of our previous results by a few simple algebraic manipulations. The advantage of such a procedure is that on the one hand it is formally very simple, and on the other it lends itself easily to generalization.

We obtain in particular the integral equations (XIX.81) and (XIX.82), which may be put in the form:

$$\psi_a^{(\pm)} = \chi_a^{(\pm)} + \frac{1}{E - H_1 \pm i\varepsilon} W_1 \psi_a^{(\pm)}. \qquad \text{(XIX.97)}$$

By comparing the asymptotic behavior of both sides of (XIX.97) we obtain relation (XIX.74).

It can also be shown that

$$\chi_a^{(\pm)} = \left[1 + \frac{1}{E - H_1 \pm i\varepsilon} U_1\right] \varphi_a, \qquad \text{(XIX.98)}$$

$$\psi_a^{(\pm)} = \left[1 + \frac{1}{E - H \pm i\varepsilon} W_1\right] \chi_a^{(\pm)}, \qquad \text{(XIX.99)}$$

the proof of which is left to the reader.

We can also give a formal definition of T. Relation (XIX.20) gives its matrix elements only between plane waves of the same energy E. Taking equation (XIX.90) into account, this relation can be put in the form

$$\langle \varphi_b | T | \varphi_a \rangle = \langle \varphi_b | \left[V + V \frac{1}{E - H + i\varepsilon} V \right] | \varphi_a \rangle. \qquad \text{(XIX.100)}$$

We generalize this relation by putting:

$$T = V + V \frac{1}{E - H + i\varepsilon} V. \qquad \text{(XIX.101)}$$

In this definition, E plays the role of a simple parameter, T is the value taken by the function

$$T(z) = V + V \frac{1}{z - H} V$$

of the complex variable z in the limit when z tends towards the real value E while remaining in the upper half-plane ($\mathrm{Im}\, z > 0$).

Transforming the right-hand side of eq. (XIX.101) with the aid of identities (XIX.88a) and (XIX.88b), we easily obtain the following integral equations for the operator T:

$$T = V + T \frac{1}{E - H_0 + i\varepsilon} V \qquad \text{(XIX.102}a\text{)}$$

$$T = V + V \frac{1}{E - H_0 + i\varepsilon} T. \qquad \text{(XIX.102}b\text{)}$$

In practice, what usually appears are the matrix elements of T between free waves of energy E, and in particular between plane waves of energy E.

As an illustration we shall demonstrate the microreversibility relation (XIX.21) using the formal properties of T alone. Denote by K the (antilinear) time-reversal operator. H_0 is obviously invariant under time reversal, and furthermore:

$$\varphi_{Ka} = K\varphi_a, \qquad \varphi_{Kb} = K\varphi_b. \tag{XIX.103}$$

Let us suppose that H has this invariance property as well:

$$KHK^\dagger = H.$$

i.e.

$$KVK^\dagger = V.$$

From definitions (XIX.84) and (XIX.101) we successively obtain

$$KG^{(+)}K^\dagger = G^{(-)} = G^{(+)\dagger}$$

and

$$KTK^\dagger = T^\dagger. \tag{XIX.104}$$

From the above law for the transformation of T under time-reversal we deduce:

$$\begin{aligned} T_{Kb\to Ka} &= (\langle\varphi_a|K^\dagger)T(K|\varphi_b\rangle) \\ &= \langle\varphi_a|(K^\dagger TK)|\varphi_b\rangle^* \\ &= \langle\varphi_b|(K^\dagger T^\dagger K)|\varphi_a\rangle \\ &= \langle\varphi_b|T|\varphi_a\rangle = T_{a\to b}. \end{aligned}$$

This is the microreversibility relation.

15. Note on the $1/r$ Potentials

In the preceding sections, we have everywhere supposed that the potentials $V(r)$, $U_1(r)$ and $W_1(r)$ fall off at infinity more rapidly than $1/r$. This is a necessary condition for the various waves and Green's functions used to have the indicated asymptotic behavior. However, with a few modifications, the same treatment can be used when both $V(r)$ and $U_1(r)$ have a $1/r$ behavior provided that their difference $W_1(r)$ tends to 0 more rapidly than $1/r$. The asymptotic forms differ only through the presence of an additional term proportional to $\ln 2kr$ in the argument of the exponential; the arguments of the asymp-

totic forms of u_l and F_l undergo the same modification. Apart from these differences the method and the results are essentially the same. In particular, definitions (XIX.83) and (XIX.84) of the Green's functions are still valid, integral equation (XIX.97) is still satisfied as is the fundamental equation (XIX.74) and also formula (XIX.75) in the limit when W_1 can be treated as a small perturbation.

III. COMPLEX COLLISIONS AND THE BORN APPROXIMATION

16. Generalities. Cross Sections

In this section [1]), we extend the considerations of the two preceding sections to collisions between any two particles. We use the word "particle" in the general sense of the term where the particle may have an internal structure. The internal variables reduce to the spin variables alone in the case of the elementary particles such as electrons, protons and neutrons; but we shall also consider complex particles, ions, atoms, atomic nuclei, etc., made up of several elementary particles.

Suppose then that we have two particles, A and X, elementary or complex, entering into collision. In general, three types of effects may be observed:

(a) *elastic scattering*, where the particles are simply deviated without any change in their respective internal quantum states

$$A + X \rightarrow A + X$$

(b) *inelastic scattering*, where in addition to deviation, there is a change in the internal states of the particles [2]); they emerge, for example, in states A' and X' respectively:

$$A + X \rightarrow A' + X'$$

(c) a *rearrangement collision* (or, properly speaking, a *reaction*), where A and X exchange a certain number of their constituent elemen-

[1]) For more details concerning the terminology introduced in this section see J. Blatt and V. Weisskopf *Theoretical Nuclear Physics* Ch. VIII (Wiley, New York, 1952).

[2]) This could be a simple reversal of spin. The definition adopted here differs slightly from the usual one according to which a scattering process is inelastic or elastic according as there is a change in the internal energy of the particles or not.

tary particles in the course of the collision, or more generally, when the system $(A+X)$ splits into two or more different particles; for example when particles B and Y (different to A and X) emerge from the collision:

$$A+X \to B+Y.$$

A chemical reaction is a rearrangement collision between molecules, a nuclear reaction a rearrangement collision between atomic nuclei; several examples will be given in what follows.

The characteristic magnitude in the collision of two particles is the cross section. This notion has already been defined in Chapter X (§ 2) for elastic collisions. More generally, *the cross section for a certain type of event in a given collision is equal to the number of events of this type per unit time per target divided by the flux of incident particles relative to the target.*

To unambiguously define a cross section relating to the collision of A and X the initial conditions of the collision, that is, the quantum states and the respective velocities of particles A and X must be precisely specified. In the usual experimental conditions one of the velocities is null, that of X say: a target of particles of type X practically at rest is bombarded with a mono-ergic beam of particles of type A. However, since the motion of the center of mass can be completely separated from the motion of the relative coordinates, simple kinematic considerations relate cross sections corresponding to initial conditions differing only by a change of referential; in particular, by an evident generalization of the method of § X.7 (cf. Problem X.2), we can relate the cross sections in the laboratory system (velocity of X null) to the cross section in the center-of-mass system (velocity of the center of mass null). All the cross sections to be considered in what follows are with respect to the center-of-mass system.

In what follows E will denote the total incident energy in the center-of-mass system and H the Hamiltonian for the relative variables (one obtains the total Hamiltonian by adding the energy of the center of mass).

17. Channels

A concept that is very useful in the treatment of complex collisions is the concept of a *channel*. A channel is any possible mode of frag-

mentation of the system in the course of the collision. One of these modes is the two incident particles themselves: it is called the *entrance channel*. In an elastic collision, the two particles remain in the entrance channel; in the other two types of reaction, the out-going channel is different to the entrance channel: it is an inelastic collision or a rearrangement collision according as the particles which make up the channel are the same or are different to those of the entrance channel. We shall henceforth denote each channel by a particular Greek letter.

With each channel may be associated a certain number of parameters and magnitudes. For simplicity, we shall always suppose that the reaction channels are made up of two particles.

Consider then a channel γ made up of two particles, C and X. The vector $\mathbf{r}_\gamma$ fixes the position of C with respect to X, in other words, we have $\mathbf{r}_\gamma = \mathbf{r}_C - \mathbf{r}_Z$, where $\mathbf{r}_C$ and $\mathbf{r}_Z$ denote the respective positions of the centers of mass of C and X. We likewise define the relative momentum $\mathbf{p}_\gamma$ and the reduced mass M_γ:

$$M_\gamma = \frac{M_C M_Z}{M_C + M_Z};$$

The kinetic energy of channel γ is $\mathbf{p}_\gamma^2/2M_\gamma$.

The wave function φ_γ (of norm 1), describing the internal quantum state of the particles of the channel is the product of the wave functions φ_C and φ_Z of the particles C and Z respectively. If h_C and h_Z are the Hamiltonians of these particles, we have

$$h_C \varphi_C = e_C \varphi_C, \qquad h_Z \varphi_Z = e_Z \varphi_Z$$

$$\varphi_\gamma = \varphi_C \varphi_Z, \qquad e_\gamma = e_C + e_Z$$

e_γ is the total internal energy of the particles of channel γ.

Finally, let V_γ be the interaction between particles C and Z, that is, the sum of the interaction potentials between each elementary particle of C and each elementary particle of Z. V_γ tends to zero in the limit when $r_\gamma \to \infty$. In what follows, we shall always suppose that this decrease is more rapid than $1/r_\gamma$; the generalization to potentials in $1/r_\gamma$ is easily accomplished as indicated in § 15.

The Hamiltonian of the system is

$$H = H_\gamma + V_\gamma \tag{XIX.105}$$

with

$$H_\gamma = h_C + h_Z + \mathbf{p}_\gamma^2/2M_\gamma. \tag{XIX.106}$$

18. Calculation of Cross Sections. T Matrices

In the collision considered above, the entrance channel is the channel $(A+X)$ which we denote by the symbol α. To completely specify the initial conditions, we need to know the relative momentum $\hbar\boldsymbol{k}_a$ of the particles in the initial state. We shall represent the initial conditions by the letter or index $a \equiv (\alpha, \boldsymbol{k}_a)$. One has:

$$E = e_\alpha + \hbar^2 k_a^2/2M_\alpha. \tag{XIX.107}$$

Let us suppose that the reaction $A+X \to B+Y$ is possible. In this case one says that the channel $(B+Y)$, or the channel β, is an *open channel*. The particles of β emerge with a well-defined relative momentum of length $\hbar k_b$ given by the conservation of energy law

$$E = e_\beta + \hbar^2 k_b^2/2M_\beta. \tag{XIX.108}$$

For channel β to be open, it is obviously necessary that $E - e_\beta$ be positive.

One can associate a plane wave,

$$\Phi_a = \varphi_\alpha \, e^{i\boldsymbol{k}_a \cdot \boldsymbol{r}_\alpha}, \tag{XIX.109}$$

and two stationary waves $\Psi_a^{(+)}$ and $\Psi_a^{(-)}$ with each set of initial conditions a. $\Psi_a^{(+)}$ is the stationary solution of H corresponding to energy E having an ($e^{i\boldsymbol{k}_a \cdot \boldsymbol{r}_\alpha}$ + outgoing wave) asymptotic behavior in the entrance channel, and a purely outgoing-wave behavior in all other open channels. $\Psi_a^{(-)}$ is similarly defined for incoming waves. In accordance with these definitions we have

$$H_\alpha \Phi_a = E\Phi_a \tag{XIX.110}$$

$$H\Psi_a^{(\pm)} = E\Psi_a^{(\pm)} \tag{XIX.111}$$

and, for any open channel γ (we shall suppose that none of them has any more than two particles):

$$\begin{aligned} \Psi_a^{(\pm)} &\underset{r_\alpha \to \infty}{\sim} \varphi_\alpha \left[e^{i\boldsymbol{k}_a \cdot \boldsymbol{r}_\alpha} + f_{a\alpha}^{(\pm)}(\Omega_\alpha) \, e^{\pm i k_a r_\alpha}/r_\alpha \right] \\ &\underset{r_\gamma \to \infty}{\sim} \varphi_\gamma \, f_{a\gamma}^{(\pm)}(\Omega_\gamma) \, e^{\pm i k_c r_\gamma} \qquad (\gamma \neq \alpha). \end{aligned} \tag{XIX.112}$$

We assume that the stationary solutions $\Psi_a^{(+)}$ and $\Psi_a^{(-)}$ exist.

The cross sections are directly related to the asymptotic behavior

of $\Psi_a^{(+)}$. Denote by $d\sigma_{a\to b}/d\Omega_\beta$ the cross section for emission of particle B in the direction Ω_b. One finds:

$$\frac{d\sigma_{a\to b}}{d\Omega_\beta} = \frac{v_b}{v_a} \mid f_{a\beta}^{(+)}(\Omega_b) \mid^2, \tag{XIX.113}$$

where $v_a = \hbar k_a/M_\alpha$ is the incident velocity and $v_b = \hbar k_b/M_\beta$ the relative velocity of the particle B emerging from the collision. The demonstration of this result is a simple generalization of that given in § 4–6 of Chapter X [1]). It is left to the reader.

Generalizing the considerations of § 3, we are led to describing the phenomenon of collision at energy E by a certain transition matrix T defined such that

$$T_{a\to b} \equiv \langle b|T|a\rangle = -\frac{2\pi\hbar^2}{M_\beta} f_{a\beta}^{(+)}(\Omega_b). \tag{XIX.114}$$

From (XIX.113) and (XIX.114),

$$\boxed{\frac{d\sigma_{a\to b}}{d\Omega_\beta} = \frac{2\pi}{\hbar v_a} |T_{a\to b}|^2 \varrho_b(E)} \tag{XIX.115}$$

where $\varrho_b(E)$ denotes the density of final states in accordance with the definition of § XVII.5. Formula (XIX.115) is the generalization of formula (XIX.19′). $T_{a\to b}$ is the amplitude for the transition $a \to b$.

19. Integral Representations of the Transition Amplitude

In order to find the integral representations of $T_{a\to b}$ we generalize property (XIX.9).

Let H and $\hat{H}$ be two possible Hamiltonians for the quantum system in question: H and $\hat{H}$ have the same kinetic energy, but their potential energies may be different; we shall assume that any such difference is made up of terms tending asymptotically to zero more rapidly than $1/r$. The stationary solutions of H and $\hat{H}$ at energy E will be denoted by Ψ and $\hat{\Psi}$ respectively. It is to be noted that certain channels may be open to collisions governed by H and closed to collisions governed by $\hat{H}$, and *vice versa*. We shall suppose that channel β is open to collisions governed by $\hat{H}$. Therefore there exists a solution

[1]) This result can also be obtained by generalizing the argument of § X.2: the cross section is equal to the ratio of the flux of B particles, $v_b|f_{a\beta}^{(+)}(\Omega_b)|^2$ to the incident flux v_a.

$\hat{\Psi}_b^{(-)}$ corresponding to the initial conditions $b \equiv (\beta, \mathbf{k}_b)$. We thus have:

$$\hat{H}\hat{\Psi}_b^{(-)} = E\hat{\Psi}_b^{(-)} \tag{XIX.116}$$

and, for any open channel δ (we shall suppose that no channel has any more than two particles),

$$\begin{aligned} \hat{\Psi}_b^{(-)} &\underset{r_\beta \to \infty}{\sim} \varphi_\beta [e^{i\mathbf{k}_b \cdot \mathbf{r}_\beta} + \hat{f}_{b\beta}^{(-)}(\Omega_\beta)\, e^{-ik_b r_\beta}/r_\beta] \\ &\underset{r_\delta \to \infty}{\sim} \varphi_\delta \hat{f}_{b\delta}^{(-)}(\Omega_\delta)\, e^{-ik_\alpha r_\delta}/r_\delta \qquad (\delta \neq \beta). \end{aligned} \tag{XIX.117}$$

One can show that

$$\langle \hat{\Psi}_b^{(-)} | (H - \hat{H}) | \Psi_a^{(+)} \rangle = -\frac{2\pi\hbar^2}{M_\beta} f_{a\beta}^{(+)}(\Omega_b) + \frac{2\pi\hbar^2}{M_\alpha} \hat{f}_{b\alpha}^{(-)*}(-\Omega_a). \tag{XIX.118}$$

The demonstration can be carried through the same way as that of property (XIX.9). Multiplying equation (XIX.111) by $\hat{\Psi}_b^{(-)*}$ and the complex conjugate of equation (XIX.116) by $\Psi_a^{(+)}$ and subtracting, we have

$$[\hat{\Psi}_b^{(-)*}(H\Psi_a^{(+)}) - (H\hat{\Psi}_b^{(-)})^*\Psi_a^{(+)}] + \hat{\Psi}_b^{(-)*}(H - \hat{H})\Psi_a^{(+)} = 0.$$

We obtain relation (XIX.118) by summing over the spin variables and integrating over the whole of configuration space. In spite of the Hermitean character of H, the contribution from the bracket does not necessarily vanish, since neither Ψ nor $\hat{\Psi}$ has a finite norm. To calculate this contribution, we calculate the integral over a finite volume of configuration space and then take the limit of that integral when this volume is extended to infinity. By use of Green's theorem the volume integral can be transformed into a surface integral [1]),

[1]) Recall that there are several ways to separate out the center of mass, but that the quantity

$$x \equiv (\textstyle\sum_j M_j \mathbf{r}_j^2)^{\frac{1}{2}}$$

is independent of the mode of reduction; $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_j, \ldots$ here denote the relative coordinates corresponding to this particular mode of reduction, and $M_1, M_2, \ldots, M_j, \ldots$, the corresponding masses (cf. § IX,13). The volume of integration involved here can be defined by the condition $x \leqslant X$; where X is a positive quantity, large with respect to the range of the interactions entering into play. With Green's theorem we can transform the volume integral into an integral over the hypersurface $x = X$, which tends to the expression on the right hand side of (XIX.119) in the limit when $X \to \infty$.

which takes the form of a sum of terms relative to the different open channels ν common to Ψ and $\hat{\Psi}$. One finds

$$\langle \hat{\Psi}_b^{(-)} | (H - \hat{H}) | \Psi_a^{(+)} \rangle = \sum_\nu \frac{\hbar^2}{2M_\nu} \lim_{R_\nu \to \infty} \{ \hat{\Psi}_a^{(-)*}, \Psi_b^{(+)} \}_\nu, \quad \text{(XIX.119)}$$

The symbol $\{...\}_\nu$ is a generalization of the notation introduced in § 2. By definition

$$\{F_1, F_2\}_\nu = \left\langle \int \left(F_1 \frac{\partial F_2}{\partial r_\nu} - F_2 \frac{\partial F_1}{\partial r_\nu} \right) \Big|_{r_\nu = R_\nu} R_\nu^2 \, d\Omega_\nu \right\rangle$$

the symbol $\langle \int ... \, d\Omega_\nu \rangle$ signifying summation or integration over all the variables except the relative distance r_ν. The different terms on the right-hand side of (XIX.119) may be evaluated by substituting for Ψ and $\hat{\Psi}$ and for their radial derivatives the respective asymptotic forms given by expressions (XIX.112) and (XIX.117). For any channel different to α and β, $\Psi_a^{(+)}$ is a purely outgoing wave, $\Psi_b^{(-)}$ a purely incoming wave, and therefore $\{...\}_\nu$ tends asymptotically to zero. For channels α and β, on the other hand the contribution from the bracket does not vanish, due to the presence of the plane-wave term in one of the two asymptotic forms. The calculation is identical to that of § 2 and gives the two terms on the right-hand side of relation (XIX.118). Q.E.D.

The fundamental relation (XIX.118) can be put in a more workable form. To this end we apply (XIX.118) in the special case where $\hat{H} = H$. The left-hand side being zero, we have

$$\frac{1}{M_\beta} f_{a\beta}^{(+)}(\Omega_b) = \frac{1}{M_\alpha} f_{b\alpha}^{(-)*}(-\Omega_a).$$

Comparing this with the relation defining T [eq. (XIX.114)], we obtain the equivalent definition

$$\langle b|T|a\rangle = -\frac{2\pi\hbar^2}{M_\alpha} f_{b\alpha}^{(-)*}(-\Omega_a). \quad \text{(XIX.114')}$$

These two definitions apply equally well to the transition matrix $\hat{T}$ associated with the Hamiltonian $\hat{H}$. Taking these into account, property (XIX.118) becomes

$$\langle b|T|a\rangle = \langle b|\hat{T}|a\rangle + \langle \hat{\Psi}_b^{(-)} | (H - \hat{H}) | \Psi_a^{(+)} \rangle. \quad \text{(XIX.120)}$$

It is in this form that we shall make use of it.

This relation remains true when channel α is closed to wave $\hat{\Psi}_b^{(-)}$ if we put $\langle b|\hat{T}|a\rangle = 0$. Similarly, it remains true when channel β is closed to wave $\Psi_a^{(+)}$ if we put $\langle b|T|a\rangle = 0$.

In particular, if $\hat{H} = H_\beta$, $\hat{\Psi}_\beta^{(-)}$ reduces to the plane wave Φ_β, and since $H - H_\beta = V_\beta$, relation (XIX.120) gives

$$\langle b|T|a\rangle = \langle \Phi_b| V_\beta |\Psi_a^{(+)}\rangle. \tag{XIX.121}$$

Similarly, if we replace H by H_α and $\hat{H}$ by H, (XIX.120) gives

$$\langle b|T|a\rangle = \langle \Psi_b^{(-)}| V_\alpha |\Phi_a\rangle. \tag{XIX.122}$$

All of these properties are valid whatever the type of collision involved, elastic collision ($\alpha = \beta$), inelastic collision ($\alpha \neq \beta$, $V_\alpha = V_\beta$) or rearrangement collision ($V_\alpha \neq V_\beta$).

20. The Born Approximation and its Generalizations

By substituting the plane wave Φ_a for $\Psi_a^{(+)}$ in the right-hand side of equation (XIX.121), one obtains the *amplitude for the transition* $a \to b$ *in the Born approximation* [cf. eq. (XIX.23)]:

$$T_{a\to b} \simeq T_{a\to b}^{(B)} = \langle \Phi_b| V_\beta |\Phi_a\rangle. \tag{XIX.123}$$

We obtain *the same* approximate value by replacing $\Psi_b^{(-)}$ by Φ_b in the right-hand side of equation (XIX.122). Indeed, if we replace H by H_α and $\hat{H}$ by H_β, (XIX.120) gives

$$\langle \Phi_b| V_\beta |\Phi_a\rangle = \langle \Phi_b| V_\alpha |\Phi_a\rangle. \tag{XIX.124}$$

Thus we can also write

$$T_{a\to b}^{(B)} = \langle \Phi_b| V_\alpha |\Phi_a\rangle. \tag{XIX.123'}$$

The passage from (XIX.121) to (XIX.123) is justified if $\Psi^{(+)}$ differs sufficiently little from Φ_a in the region covered by the potential V_β; that from (XIX.122) to (XIX.123'), if $\Psi_b^{(-)}$ differs sufficiently little from Φ_b in the region covered by the potential V_α. These two conditions are obviously equivalent, although they appear different. They both suppose that $\Psi_a^{(+)}$ and $\Psi_b^{(-)}$ may be respectively replaced by the plane waves Φ_a and Φ_b.

More exact expressions may be obtained by using wave functions nearer to the exact stationary solutions $\Psi_a^{(+)}$ and $\Psi_b^{(-)}$ in place of the plane waves.

Suppose, for example, that V_α can be put in the form

$$V_\alpha = U_\alpha + W_\alpha$$

and that the stationary solutions of the Hamiltonian $H_\alpha + U_\alpha$ corresponding to the energy E are known. These solutions will be denoted by X and the corresponding transition matrix by $T^{(X)}$. In particular, $X_a^{(+)}$ will denote the "outgoing wave" solution corresponding to initial conditions a and we have

$$\begin{aligned} T^{(X)}_{a\to b} &= \langle \Phi_b | (H_\alpha + U_\alpha - H_\beta) | X_a^{(+)} \rangle \\ &= \langle \Phi_b | (V_\beta - W_\alpha) | X_a^{(+)} \rangle. \end{aligned} \tag{XIX.125}$$

If W_α is sufficiently small, $X_a^{(+)}$ will differ little from $\Psi_a^{(+)}$.

Similarly, suppose that V_β can be put in the form

$$V_\beta = U_\beta + W_\beta$$

and that the stationary solutions of the Hamiltonian $H_\beta + U_\beta$ corresponding to the energy E are known. We shall denote these solutions by Ξ and the associated transition matrix by $T^{(\Xi)}$. In particular, $\Xi_b^{(-)}$ is the "incoming wave" solution corresponding to the initial conditions b and we have

$$T^{(\Xi)}_{a\to b} = \langle \Xi_b^{(-)} | (V_\alpha - W_\beta) | \Phi_a \rangle. \tag{XIX.126}$$

If W_β is sufficiently small, $\Xi_b^{(-)}$ will differ little from $\Psi_b^{(-)}$.

Now according to (XIX.120)

$$T_{a\to b} = T^{(X)}_{a\to b} + \langle \Psi_b^{(-)} | W_\alpha | X_a^{(+)} \rangle \tag{XIX.127}$$

$$= T^{(\Xi)}_{a\to b} + \langle \Xi_b^{(-)} | W_\beta | \Psi_a^{(+)} \rangle. \tag{XIX.128}$$

These two expressions for $T_{a\to b}$ are exact. We obtain approximate expressions by replacing $\Psi_b^{(-)}$ by $\Xi_b^{(-)}$ in the first one and $\Psi_a^{(+)}$ by $X_a^{(+)}$ in the second

$$T_{a\to b} \simeq T^{(X)}_{a\to b} + \langle \Xi_b^{(-)} | W_\alpha | X_a^{(+)} \rangle \tag{XIX.129}$$

$$\simeq T^{(\Xi)}_{a\to b} + \langle \Xi_b^{(-)} | W_\beta | X_a^{(+)} \rangle. \tag{XIX.130}$$

Although formally different if $W_\alpha \neq W_\beta$, these two expressions are always equal, as may easily be seen by applying (XIX.120) with $H_\alpha + U_\alpha$ in place of H, and $H_\beta + U_\beta$ in place of $\hat{H}$.

These formulas constitute a generalization of the Born approximation [cf. eq. (XIX.75)]. They are better than the one obtained by

simply replacing $\Psi_a^{(+)}$ by $X_a^{(+)}$ in expression (XIX.121). Indeed since one has [eq. (XIX.125)]

$$\langle \Phi_b | V_\beta | X_a^{(+)} \rangle = T^{(X)}_{a \to b} + \langle \Phi_b | W_a | X_a^{(+)} \rangle,$$

this last approximation amounts to replacing $\Psi_b^{(-)}$ in expression (XIX.127) by Φ_b and not by $\Xi_b^{(-)}$. For similar reasons, formulas (XIX.129) and (XIX.130) are better than the expression obtained by replacing $\Psi_b^{(-)}$ by $\Xi_b^{(-)}$ in the right-hand side of (XIX.122).

21. Scattering of Fast Electrons by an Atom

As an illustration of the Born approximation proper, consider the scattering of fast electrons by an atom.

We shall suppose that the nucleus is infinitely heavy and fixed at the origin. The Hamiltonian of the system is

$$H = \frac{\mathbf{p}^2}{2m} + h + V(\mathbf{r}, A).$$

The first term is the kinetic energy of the incident electron, the second the Hamiltonian of the atom and the third the interaction energy

$$V(\mathbf{r}, A) = -\frac{Ze^2}{r} + \sum_{i=1}^{Z} \frac{e^2}{|\mathbf{r} - \mathbf{r}_i|}, \qquad \text{(XIX.131)}$$

where $\mathbf{r}$ is the position vector of the incident electron and

$$\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_Z$$

the position vectors of the electrons of the atom. Finally, let

$$\varphi_0, \varphi_1, \ldots, \varphi_n, \ldots,$$

denote the eigenfunctions of h and

$$e_0, e_1, \ldots, e_n, \ldots,$$

the corresponding eigenvalues.

We consider an inelastic scattering process in which the momentum of the electron changes from $\hbar\mathbf{k}_0$ to $\hbar\mathbf{k}_n$ while the atom passes from the ground state φ_0 to the excited state φ_n. Energy being conserved in the collision we have

$$\hbar^2(k_0^2 - k_n^2) = 2m(e_n - e_0).$$

We denote the momentum transfer by $\hbar\boldsymbol{q}$,

$$\boldsymbol{q} = \boldsymbol{k}_n - \boldsymbol{k}_0.$$

Note that relation (XIX.24) does not apply here, but that

$$q^2 = (k_0 - k_n)^2 + 4k_0k_n \sin^2 \tfrac{1}{2}\theta. \qquad \text{(XIX.132)}$$

In addition to the direct process in which the incident electron is simply scattered with loss of a part of its kinetic energy, there may be an exchange of the incident electron with one of the electrons of the atom. This exchange effect can be important when the velocity of the scattered electron is of the same order of magnitude as the velocity of the electrons of the atom, i.e. when $k_n a \simeq 1$; it is small under conditions where the Born approximation is valid, and will be neglected here. In the same spirit, we shall treat the incident electron as a particle different from the electrons of the atom. Applying formula (XIX.123), we then obtain the transition amplitude in the Born approximation:

$$T_{0\to n} \simeq \int \mathrm{e}^{-\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}}\, \varphi_n^*(A)\, V(\boldsymbol{r}, A)\, \varphi_0(A)\, \mathrm{d}\boldsymbol{r}\, \mathrm{d}A. \qquad \text{(XIX.133)}$$

This is the Born approximation transition amplitude for a particle undergoing the momentum transfer $\hbar\boldsymbol{q}$ in the potential

$$V_n(\boldsymbol{r}) \equiv \int \varphi_n^*(A)\, V(\boldsymbol{r}, A)\, \varphi_0(A)\, \mathrm{d}A. \qquad \text{(XIX.134)}$$

This result is true whatever n and applies equally well in the elastic scattering case ($n = 0$).

In order to put it in a form analogous to that obtained in § 8 we introduce the "electron density"

$$\varrho_n(\boldsymbol{r}) = Z \int \varphi_n^*(\boldsymbol{r}, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_Z)\, \varphi_0(\boldsymbol{r}, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_Z)\, \mathrm{d}\boldsymbol{r}_2 \ldots \mathrm{d}\boldsymbol{r}_Z \qquad \text{(XIX.135)}$$

and the corresponding form factor

$$F_n(\boldsymbol{q}) = \int \mathrm{e}^{-\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}}\, \varrho_n(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r}.$$

In the elastic scattering case ($n = 0$), we simply have the ground state electron density and the corresponding form factor [cf. eq. (XIX.48)]:

$$\varrho_0(\boldsymbol{r}) = \varrho(r), \qquad F_0(\boldsymbol{q}) = F(q).$$

If we substitute the explicit form of $V(\boldsymbol{r}, A)$ given by equation

(XIX.131) in the integral defining $V_0(r)$, we see that $V_0(r)$ is the Coulomb interaction potential for an electron in the field created by the charge distribution

$$e[Z\delta(r) - \varrho(r)].$$

This is precisely the scattering potential adopted in § 8; thus the model adopted there is justified.

Consider now an inelastic collision ($n \neq 0$). We substitute the explicit form of $V(r, A)$ given by equation (XIX.131) in the integral defining $V_n(r)$. Due to the orthogonality of the functions φ_0 and φ_n, the contribution of the term Ze^2/r vanishes. The other terms give a contribution equal to the Coulomb potential felt by an electron in the field created by the charge distribution $-e\varrho_n(r)$. Repeating the manipulations of § 8, we find

$$T_{0\to n} \simeq -4\pi e^2 F_n(q) \qquad (n \neq 0)$$

whence the Born approximation result for the inelastic scattering cross section

$$\frac{d\sigma_{0\to n}}{d\Omega} = 4\frac{m^2e^4}{\hbar^4}\frac{k_n}{k_0}\frac{|F_n(q)|^2}{q^4}. \qquad \text{(XIX.136)}$$

Note the similarity of this formula and formula (XIX.49) for elastic scattering. The angular distribution depends on the factor $|F_n(q)|^2/q^4$. The form of $F_n(q)$ is easily deduced from that of $\varrho_n(r)$. In particular $F_n(0) = 0$. Moreover, if a is the radius of the atom, the factor $F_n(q)$ is appreciably different to zero only in the region $qa \lesssim 1$. When the Born approximation is valid ($\gamma \ll 1$), we have $k_0 a \gg 1$; this corresponds to an incident energy much greater than the distance between atomic levels, which is of the order of $\hbar^2/ma^2$; it follows that $(k_0 - k_n)a$ is of the order of $1/k_0 a$. Since q varies from $k_0 - k_n$ to $k_0 + k_n$ when θ increases from zero to 2π, qa varies from $1/k_0 a$ to $2k_0 a$ in the same interval [cf. eq. (XIX.132)]. According to the foregoing discussion, the electron is most probably scattered into angles for which $qa \lesssim 1$, that is, into the domain of small angles:

$$\theta \lesssim \frac{1}{k_0 a}.$$

In principle, inelastic scattering can be used to measure the position of the electron. By hypothesis, the transversal dimensions of the

incident wave packet are much greater than a, and the uncertainty in the transversal component of the momentum is therefore much smaller than $\hbar/a$. Immediately after the inelastic scattering, supposing that we can detect the change in quantum state of the atom, the position of the electron is known with an uncertainty of order a; however, the direction of its momentum being known only up to $1/k_0 a$, there is an uncertainty in the transversal component of the said momentum of order $\hbar/a$. This is in agreement with what was said in Chapter IV concerning position measurements (cf. in particular the discussion of measurements in the Wilson chamber, note, p. 142).

22. Coulomb Excitation of Nuclei

As an illustration of the generalized Born approximation, we take up again the problem treated in § XVII.3 with a view to avoiding the use of the classical approximation.

Unless otherwise indicated, we shall use the notation of § XVII.3. We shall suppose the incident energy sufficiently low that condition (XVII.28) is realized. On the other hand, conditions (XVII.29) and (XVII.30) upon which the classical treatment was founded will not necessarily be fulfilled. Since the incident energy is lower than the height Ze^2/R of the Coulomb barrier, the proton "penetrates" very little into the nucleus in the course of the collision. We therefore expect the effect of the specifically nuclear interactions to be small, and we neglect it for the moment. The collision is then governed by the Hamiltonian $H^{(0)} + V$.

We denote the initial and final conditions of the collision by a and b respectively and the corresponding eigensolutions of $H^{(0)}$ by $X_a^{(+)}$ and $X_b^{(-)}$. Since in $H^{(0)}$ the proton variables are completely separated from the nuclear variables, $X_a^{(+)}$ is the product of the nuclear wave-function in the initial state α by a Coulomb scattering wave $\xi_a^{(+)}(\boldsymbol{r})$ representing the stationary scattering state of a proton of energy E and initial momentum $\hbar\boldsymbol{k}_a$ by the potential Ze^2/r; similarly, $X_b^{(-)}$ is the product of the wave function of the nucleus in the state β by a Coulomb wave $\xi_b^{(-)}(\boldsymbol{r})$ of the "incoming wave" type [1]) representing the scattering of a proton of energy $(E - \Delta E)$ and of momentum

[1]) Its expansion into spherical harmonics is given by eq. (XIX.77), taking regular Coulomb solutions for F_l. Note that $\xi_b^{(-)} = \xi_{Kb}^{(+)*}$.

$\hbar \mathbf{k}_b$ by the potential Ze^2/r. Since $X_a^{(+)}$ gives no contribution to channel β the exact formula (XIX.127) gives

$$T_{a\to b} = \langle \Psi_b^{(-)} | V | X_a^{(+)} \rangle.$$

If V is treated as a small perturbation we can neglect its contribution to $\Psi_b^{(-)}$ and replace $\Psi_b^{(-)}$ by $X_b^{(-)}$, giving, in agreement with the generalized Born formula (XIX.129),

$$T_{a\to b} = \langle X_b^{(-)} | V | X_a^{(+)} \rangle.$$

Since the X-waves practically vanish for $r<R$, we may replace V by its expansion (XVII.34), giving

$$T_{a\to b} = \sum_{lm} (-)^m R_l^{-m} \langle \beta J_\beta M_\beta | Q_l^m | \alpha J_\alpha M_\alpha \rangle$$
$$R_l^m = \frac{4\pi Ze}{2l+1} \int \xi_b^{(-)*}(\mathbf{r}) \frac{Y_l^m(\Omega)}{r^{l+1}} \xi_a^{(+)}(\mathbf{r})\, d\mathbf{r}. \qquad \text{(XIX.137)}$$

By applying relation (XIX.115) one obtains the Coulomb excitation cross section. The expression obtained is very similar to the one given by the semi-classical theory. The latter is obtained by replacing R_l^m by $A(d\sigma_R/d\Omega)^{\frac{1}{2}} S_l^m$, where A is the positive constant defined by relation (XIX.20): $A^2 = 4\pi^2\hbar^4 v_a / M^2 v_b$. Indeed in the limit when conditions (XVII.29–30) are fulfilled, the chief contribution to the integral R_l^m comes from the neighborhood of the classical trajectory and the semi-classical result can be obtained by applying the method of stationary phases.

The discussion of selection rules given at the end of § XVII.3 applies here without change. In particular, for experiments in which the target nuclei are not oriented and in which the polarization of the excited nuclei is not measured, we have

$$\frac{d\sigma_{a\to b}}{d\Omega} = \frac{1}{(2J_\alpha+1)(2l_0+1)} \frac{M^2 v_b}{4\pi^2\hbar^4 v_a} |\langle \beta \| Q^{(l_0)} \| \alpha \rangle|^2 \left(\sum_m |R_{l_0}^m|^2\right).$$

The above treatment can be extended to include the specifically nuclear interactions. Denote by V_{pA} the nuclear interaction potential between the proton and the nucleons of nucleus A. V_{pA} is very strong for $r<R$ and practically vanishes for $r>R$. We shall henceforth include V_{pA} in $H^{(0)}$. It is convenient to separate the external and the internal parts of V: $V = V_{\text{int}} + V_{\text{ext}}$, in accordance with the definition

$$V_{\text{ext}} = \begin{cases} V & \text{if} \quad r > R \\ 0 & \text{if} \quad r < R \end{cases}$$

and also to include V_{int} in $H^{(0)}$. With these modifications the separation of variables effected above is no longer possible and $X_a^{(+)}$ and $X_b^{(-)}$ are no longer pure elastic scattering waves. The generalized Born approximation now gives

$$T_{a\to b} \simeq T^{(X)}_{a\to b} + \langle X_b^{(-)} | V_{\text{ext}} | X_a^{(+)} \rangle .$$

The first term, $T^{(X)}_{a\to b}$ is the nuclear excitation term. It is the transition amplitude obtained when we replace the interaction between proton and nucleus in the external region ($r > R$) by a point Coulomb interaction; the excitation $\alpha \to \beta$ can then occur only if the proton penetrates into the nucleus. The second term is the Coulomb excitation term proper.

Due to the existence of the Coulomb barrier and the nature of nuclear forces (short range and strong), the $X_a^{(+)}$ and $X_b^{(-)}$ waves have narrow resonances at low energies analogous to the scattering resonances studied in Chapter X (Section IV). Knowing the parameters characteristic of these resonances – resonant energy and resonance width in the different open channels – one can construct these waves in the external region [1]), and calculate the two terms of the transition amplitude. Outside the resonances, the proton practically does not penetrate into the nucleus, $T^{(X)}_{a\to b}$ is negligible, and the Coulomb excitation amplitude is given to a very good approximation by formula (XIX.137), in which the X are pure Coulomb scattering waves. When passing through a resonance, the contribution from $T^{(X)}_{a\to b}$ ceases to be negligible and its energy-dependence has the typical resonant shape [cf. eq. (X.64)]; at the same time the Coulomb excitation amplitude departs significantly from formula (XIX.137); the two amplitudes contribute coherently to the total cross section.

23. Green's Functions and Integral Equations for Stationary Scattering Waves

The considerations of § 13 can easily be extended to include collisions of complex systems. To each of the Hamiltonians H, H_α, etc., we can associate the Green's function $[E - H \pm i\varepsilon]^{-1}$, $[E - H_\alpha \pm i\varepsilon]^{-1}$, etc. Except for a few changes in notation, the results of § 13 all remain valid here, and may be deduced in an analogous way, in particular

[1]) A detailed treatment of this question is given in the references of note 1, p. 832.

property (XIX.85), identities (XIX.88) and (XIX.89) and the asymptotic forms of these functions. One has, for example, for each open channel γ [cf. eq. (XIX.94)]

$$\langle \boldsymbol{r}_\gamma | \frac{1}{E-H \pm i\varepsilon} | u \rangle \underset{r_\gamma \to \infty}{\sim} -\frac{M_\gamma}{2\pi\hbar^2} \varphi_\gamma \frac{e^{\pm i k_c r_\gamma}}{r_\gamma} \langle \Psi^{(\mp)}_{(\gamma, \pm \boldsymbol{k}_c)} | u \rangle$$

$$(\boldsymbol{k}_c = k_c \boldsymbol{r}_\gamma / r_\gamma).$$

Using these various properties of the Green's functions, one can deduce the integral equations

$$\Psi_a^{(\pm)} = \Phi_a + \frac{1}{E-H_\alpha \pm i\varepsilon} V_\alpha \Psi_a^{(\pm)} \qquad \text{(XIX.138)}$$

$$\Psi_a^{(\pm)} = X_a^{(\pm)} + \frac{1}{E-H_\alpha - U_\alpha \pm i\varepsilon} W_\alpha \Psi_a^{(\pm)} \qquad \text{(XIX.139)}$$

etc., and also the formulas

$$\Psi_a^{(\pm)} = \left(1 + \frac{1}{E-H \pm i\varepsilon} V_\alpha\right) \Phi_a \qquad \text{(XIX.140)}$$

$$= \left(1 + \frac{1}{E-H \pm i\varepsilon} W_\alpha\right) X_a^{(\pm)}. \qquad \text{(XIX.141)}$$

From these integral equations or from these formulas, it is easy to deduce the Born expansion for the transition amplitudes. Substituting formula (XIX.140) into (XIX.121) and expanding the Green's function, we obtain

$$\langle b|T|a\rangle = \langle \Phi_b | \left(V_\beta + V_\beta \frac{1}{E-H+i\varepsilon} V_\alpha\right) | \Phi_a \rangle \qquad \text{(XIX.142)}$$

$$= \langle \Phi_b | \left(V_\beta + V_\beta \frac{1}{E-H_\alpha+i\varepsilon} V_\alpha + \ldots\right) | \Phi_a \rangle. \qquad \text{(XIX.143)}$$

In the particular case of sections I and II, the transition amplitudes can all be defined as matrix elements of a certain operator T defined by equation (XIX.101), or else by one or the other of integral equations (XIX.102). When dealing with complex collisions, the definition of a unique operator is no longer possible. Nevertheless it will be seen from equation (XIX.142) that any transition amplitude from channel

α to channel β at energy E may be regarded as a matrix element of a certain operator $T^{\beta\alpha}$ defined by

$$T^{\beta\alpha} = V_\beta + V_\beta \frac{1}{E - H + i\varepsilon} V_\alpha. \tag{XIX.144}$$

Note that in the case of rearrangement collisions ($V_\beta \neq V_\alpha$), these matrix elements do not answer to the usual definition of the representation of operators by matrices, since the vectors $|\Phi_a\rangle$ and $\langle\Phi_b|$ occurring in the formula

$$T_{a\to b} = \langle\Phi_b|T^{\beta\alpha}|\Phi_a\rangle \tag{XIX.145}$$

are not orthogonal.

24. Scattering of a Particle by Two Scattering Centers

The main interest of the formal considerations of the preceding paragraph is their generality and the fact that they provide a framework for the treatment of complex problems. To gain a certain familiarity with this formalism we shall apply it to a simple problem and prove in this way a certain number of familiar results.

We consider the scattering of a particle by two scattering centers, for example the scattering of an electron by two atoms. In what follows we neglect exchange effects but make no particular hypothesis about the wavelength of the incident particle.

We suppose the nuclei of the scattering atoms to be infinitely heavy and treat them as fixed centers. We take nucleus 1 as the origin of coordinates and denote by $\boldsymbol{R}$ the vector joining nucleus 1 to nucleus 2 (cf. Fig. XIX.3). The distance R is supposed large with respect to atomic dimensions: $a \ll R$. The collision is governed by the Hamiltonian

$$H = H_0 + V, \tag{XIX.146}$$

where

$$H_0 = \frac{\boldsymbol{p}^2}{2m} + h_1 + h_2 \tag{XIX.146'}$$

$$V = V_1 + V_2. \tag{XIX.146''}$$

h_1 and h_2 are the Hamiltonians of atoms 1 and 2 respectively and V_1 and V_2 their potentials of interaction with the incident particle. Finally, we denote the free-wave Green's function by G_0:

$$G_0 \equiv (E - H_0 + i\varepsilon)^{-1}.$$

The transition matrix T associated with the (elastic or inelastic) scattering of the particle by the two atoms is represented by the Born expansion

$$T = V + VG_0V + VG_0VG_0V + \ldots \qquad \text{(XIX.147)}$$

Replacing V by $V_1 + V_2$, we obtain the expansion of T in powers of V_1 and V_2

$$T = V_1 + V_2 + V_1G_0V_1 + V_1G_0V_2 + V_2G_0V_1 + V_2G_0V_2 + V_1G_0V_1G_0V_1 + V_1G_0V_1G_0V_2 + \ldots$$

In what follows, we make no particular hypothesis about the strength of V_1 or V_2. This expansion is therefore not necessarily a rapidly converging one and thus cannot be taken, as it stands, as a starting point for an approximation method. However, it is possible to regroup its terms so as to form an expansion that is rapidly converging.

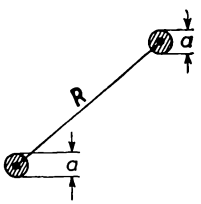


Fig. XIX.3.

The regroupment in question is based on the following remark. We consider the calculation of a matrix element of the second-order term $V_1G_0V_2$ in the representation where $\boldsymbol{r}$ is diagonal. Any matrix element of G_0, the free-wave Green's function, contains the factor $e^{ik|\boldsymbol{r}-\boldsymbol{r}'|}/|\boldsymbol{r}-\boldsymbol{r}'|$. Since V_1 is essentially localized in a small domain about the origin and V_2 in a small domain about the point $\boldsymbol{R}$, this factor is of order $1/R$. The same remark applies to the term $V_2G_0V_1$. Roughly speaking, the terms $V_1G_0V_2$ and $V_2G_0V_1$ are a/R times smaller than the terms $V_1G_0V_1$ and $V_2G_0V_2$. This remark applies equally well to higher-order terms. It suggests that we classify the various terms according to the number of times that G_0 appears between V_1 and V_2. Henceforth we shall say that a term is of the first order if G_0 never appears between V_1 and V_2, of the second order if this occurs once, ..., of the nth order if this occurs $(n-1)$ times, etc. In this terminology, $V_1G_0V_1$ and $V_2G_0V_2$ are of the first order, $V_1G_0V_2$ and $V_2G_0V_1$ of the second order.

We introduce the individual transition matrices T_1, T_2 defined by

$$\begin{aligned} T_i &= V_i + V_iG_0T_i \qquad (i = 1, 2) \qquad \text{(XIX.148)} \\ &= V_i + V_iG_0V_i + V_iG_0V_iG_0V_i + \ldots \end{aligned}$$

T_1 is the transition matrix associated with scattering from atom 1

when we neglect the interaction of the incident particle with atom 2, T_2 that associated with scattering from atom 2 when we neglect the interaction with atom 1. The contributions of successive orders are simply expressed in terms of the operators T_1, T_2 and G_0. Such expressions are easily determined by simple inspection. The first order gives T_1+T_2, the second $T_1G_0T_2+T_2G_0T_1$, etc. We therefore have

$$T = (T_1 + T_2) + (T_1G_0T_2 + T_2G_0T_1) + (T_1G_0T_2G_0T_1 + T_2G_0T_1G_0T_2) + \ldots \quad \text{(XIX.149)}$$

Expansion (XIX.149) is the starting point of our treatment. The various terms are easy to interpret. The first-order terms represent simple scattering of the particle either by atom 1 (term T_1), or by atom 2 (term T_2). The second-order terms represent double scattering: $T_1G_0T_2$ represents the scattering of the incident particle by atom 2 (operator T_2), then propagation of the scattered particle from 2 to 1 (operator G_0) and finally the scattering of the particle by atom 1 (operator T_1). Similarily each nth order term represents n successive scatterings of the particle alternatively by atoms 1 and 2.

The mechanism thus associated with each term can be schematically represented by diagrams of the type used in the time dependent perturbation theory (Fig. XVII.1); two such diagrams are represented in Fig. XIX.4.

Expansion (XIX.149) relates the transition amplitudes of the two-scatterer problem to the amplitude of the single-scatterer problem.

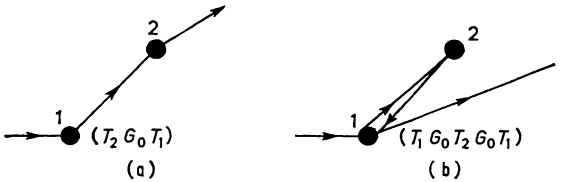


Fig. XIX.4. Diagrams representing two terms of expansion (XIX.149).

For simplicity, we shall suppose the two atoms identical and denote by $t(e)$ the transition matrix relative to energy e for the single-scatterer problem, that is, for the scattering of the particle by *one* of these atoms *placed at the origin.*

We denote the eigenvector of H_0 corresponding to the product of

the plane wave $e^{i\mathbf{k}\cdot\mathbf{r}}$ representing the incident particle with momentum $\hbar\mathbf{k}$ by the wave function φ_m of atom 1 and the wave function φ_n of atom 2 by $|\mathbf{k}mn\rangle$:

$$H_0|\mathbf{k}mn\rangle = \left(\frac{\hbar^2k^2}{2m} + e_m + e_n\right)|\mathbf{k}mn\rangle.$$

With the normalization adopted here,

$$\langle\mathbf{k}'m'n'|\mathbf{k}mn\rangle = (2\pi)^3\delta(\mathbf{k}-\mathbf{k}')\delta_{mm'}\delta_{nn'}.$$

Similarly $|\mathbf{k}m\rangle$ denotes the product of the plane wave $e^{i\mathbf{k}\cdot\mathbf{r}}$ by the wave function φ_m of the atom in the single-scatterer problem.

From the definition (XIX.148) of T_1, one easily deduces that

$$\langle\mathbf{k}'m'n'|T_1(E)|\mathbf{k}mn\rangle = \delta_{nn'}\langle\mathbf{k}'m'|t(E-e_n)|\mathbf{k}m\rangle. \quad \text{(XIX.150)}$$

Similarily. one can relate the elements of the matrix T_2 to the elements of the transition matrix associated with the scattering by a single atom *placed at the point* $\mathbf{R}$; the latter is obtained from $t(e)$ by an overall translation of $\mathbf{R}$; it follows (Problem XIX.5) that

$$\langle\mathbf{k}'m'n'|T_2(E)|\mathbf{k}mn\rangle = \delta_{mm'}e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{R}}\langle\mathbf{k}'n'|t(E-e_m)|\mathbf{k}n\rangle. \quad \text{(XIX.151)}$$

Formulas (XIX.150) and (XIX.151) apply to any transition whatsoever; they are valid even if the states $|\mathbf{k}mn\rangle$ and $|\mathbf{k}'m'n'\rangle$ correspond to an energy different to E.

25. Simple Scattering. Interference

As a first application of expansion (XIX.149), we calculate the cross section for the elastic scattering of the particle by the two atoms.

Suppose that the atoms are, and remain, in their ground state φ_0, and denote by $\mathbf{k}_0$ and $\mathbf{k}$ the wave vectors of the incident and scattered waves respectively, i.e. we consider the transition $(\mathbf{k}_0 00) \to (\mathbf{k}00)$ and we have

$$k = k_0, \qquad E = \hbar^2k^2/2m + 2e_0.$$

To calculate the transition amplitude we shall take only the first-order terms in expansion (XIX.149):

$$T \simeq T_1 + T_2. \quad \text{(XIX.152)}$$

Taking (XIX.150) and (XIX.151) into account and putting

$$\boldsymbol{q} \equiv \boldsymbol{k} - \boldsymbol{k}_0$$
$$\langle \boldsymbol{k}0|t(E - e_0)|\boldsymbol{k}_0 0\rangle \equiv \langle \boldsymbol{k}|t|\boldsymbol{k}_0\rangle,$$

we find

$$\langle \boldsymbol{k}\,00|T|\boldsymbol{k}_0\,00\rangle \simeq [1 + \mathrm{e}^{-i\boldsymbol{q}\cdot\boldsymbol{R}}]\,\langle \boldsymbol{k}|t|\boldsymbol{k}_0\rangle. \qquad \text{(XIX.153)}$$

This will be recognized as the relation giving the scattering amplitude in the elementary theory of interference. With the aid of (XIX.153) one can relate the cross section $\mathrm{d}\Sigma/\mathrm{d}\Omega$ for elastic scattering by the two atoms to the cross section $\mathrm{d}\sigma/\mathrm{d}\Omega$ for the transition $(\boldsymbol{k}_0 0) \to (\boldsymbol{k}0)$, that is, to the cross section for the same elastic scattering by a single atom. The first is obtained by multiplying the square of the modulus of $\langle \boldsymbol{k}00|T|\boldsymbol{k}_0 00\rangle$ by a suitable factor, the second by multiplying the square of the modulus of $\langle \boldsymbol{k}|t|\boldsymbol{k}_0\rangle$ by the same factor. Relation (XIX.153) thus leads to

$$\mathrm{d}\Sigma/\mathrm{d}\Omega \simeq 2I(\Omega)\,\mathrm{d}\sigma/\mathrm{d}\Omega \qquad \text{(XIX.154)}$$
$$I(\Omega) = 1 + \cos(\boldsymbol{q}\cdot\boldsymbol{R}). \qquad \text{(XIX.155)}$$

The presence of the factor $I(\Omega)$ in formula (XIX.154) is due to interference between the waves scattered by each atom. If this interference phenomenon did not occur, we would have $I(\Omega) = 1$ and the cross section $\mathrm{d}\Sigma/\mathrm{d}\Omega$ would be simply the sum of the scattering cross sections of each of the atoms 1 and 2.

The usual results of interference theory are obtained when we examine the behavior of $I(\Omega)$ as a function of the scattering angle. We discuss only the two limiting cases when the wavelength $\lambda = 2\pi/k$ is very large or very small beside R.

If $\lambda \gg R$, then $\boldsymbol{q}\cdot\boldsymbol{R} \ll 1$ whatever the scattering angle, and therefore $I(\Omega) = 2$. The cross section for scattering by the two atoms is therefore four times the individual cross section, that is, twice what it would be if the phenomenon of interference did not exist.

If $\lambda \ll R$, $I(\Omega)$ oscillates rapidly between the values 0 and 2 when the scattering angle is varied. Let the angles made by the vectors $\boldsymbol{k}_0$ and $\boldsymbol{k}$ with $\boldsymbol{R}$ be α_0 and α respectively:

$$\boldsymbol{q}\cdot\boldsymbol{R} = \frac{2\pi R\,(\cos\alpha_0 - \cos\alpha)}{\lambda}.$$

$I(\Omega)$ therefore vanishes each time that $(\cos\alpha - \cos\alpha_0)$ is a half-integral

multiple of λ/R, and is equal to 2 each time that $(\cos\alpha - \cos\alpha_0)$ is an integral multiple of λ/R. The corresponding values of α give respectively the directions of minimum and maximum interference. The width of the interference fringes is of order λ/R. The possibility of observing these interferences depends on the angular resolving power of the detecting apparatus. The latter counts the particles admitted into a certain solid angle of *finite* dimensions $\delta\Omega \equiv (\delta\alpha)^2$. The counting rate is given by the integral of $d\Sigma/d\Omega$ over this finite domain. The angular dependence of $d\Sigma/d\Omega$ in this domain is essentially given by the factor $I(\Omega)$. If $\delta\alpha \gg \lambda/R$, $I(\Omega)$ has many oscillations in the domain in question and may be replaced by its average value: the results will be the same as if $I(\Omega) = 1$ and as if the scattering from the two atoms was incoherent [1]). If, on the other hand, $\delta\alpha \ll \lambda/R$, $I(\Omega)$ is practically constant in this domain and the resolving power of the detector is sufficient for the observation of the interference phenomenon.

Let us now consider inelastic collisions in which one of the two atoms passes from its ground state to its nth excited state. Denote by $\boldsymbol{k}_n$ the wave vector for the scattered wave

$$E = \frac{\hbar^2 k_n^2}{2m} + e_0 + e_n.$$

The cross section $d\Sigma_n/d\Omega$ is the sum of the cross sections $d\Sigma_{n0}/d\Omega$ and $d\Sigma_{0n}/d\Omega$ corresponding to the transitions $(\boldsymbol{k}_0 00) \to (\boldsymbol{k}_n n0)$ and $(\boldsymbol{k}_0 00) \to (\boldsymbol{k}_n 0n)$ respectively. These are easily calculated in approximation (XIX.152) where we retain only the simple scattering terms in the expansion of T. In the transition $(\boldsymbol{k}_0 00) \to (\boldsymbol{k}_n n0)$, atom 1 is excited, atom 2 remains in its ground state; the contribution of T_2 obviously vanishes [cf. eq. (XIX.151)]; the contribution of T_1 gives $d\sigma_n/d\Omega$, the cross section of the process $(\boldsymbol{k}_0 0) \to (\boldsymbol{k}_n n)$, the elastic scattering of the particle by a single atom. The transition $(\boldsymbol{k}_0 00) \to (\boldsymbol{k}_n 0n)$ corresponds to the inverse situation: the contribution of T_1 vanishes, that of T_2 gives $d\sigma_n/d\Omega$. Finally

$$\frac{d\Sigma_n}{d\Omega} \equiv \frac{d\Sigma_{n0}}{d\Omega} + \frac{d\Sigma_{0n}}{d\Omega} = 2\frac{d\sigma_n}{d\Omega}.$$

[1]) This result is a confirmation and a more exact statement of the remarks of § X.2 regarding the conditions in which a target may be considered as formed of independent scatterers.

Contrary to what happens in elastic scattering, each atom acts as if it were alone: the waves scattered inelastically by atom 1 and by atom 2 correspond to different channels and cannot interfere.

26. Multiple Scattering

Consider now the transition $(\boldsymbol{k}_0 00) \to (\boldsymbol{k}pq)$ in which atoms 1 and 2 pass from their ground states to their pth and qth excited states respectively

$$E = \frac{\hbar^2 k^2}{2m} + e_p + e_q.$$

The cross section for this transition will be denoted by $\mathrm{d}\Sigma_{pq}/\mathrm{d}\Omega$.

Since both atoms change their quantum state in the transition, the contribution from the simple scattering terms T_1 and T_2 vanishes, and we must push the expansion of T to the second order. This gives

$$\langle \boldsymbol{k}pq|T|\boldsymbol{k}_0 00\rangle \simeq A_{12} + A_{21}, \tag{XIX.156}$$

where

$$A_{ij} \equiv \langle \boldsymbol{k}pq|T_i G_0 T_j|\boldsymbol{k}_0 00\rangle. \tag{XIX.157}$$

The diagrams for these two double scattering amplitudes are given in Figure XIX.5

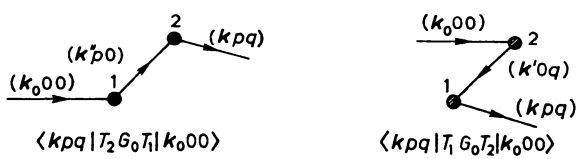


Fig. XIX.5. Diagram for the double scattering amplitudes.

We first calculate A_{21}. In order to take into account the condition $a \ll R$, we make the calculation in the $\{rmn\}$ representation. Taking into account that G_0 commutes with h_1 and h_2, we find

$$A_{21} = \iint f_2(\boldsymbol{r}'')\, g(\boldsymbol{r}'', \boldsymbol{r}')\, f_1(\boldsymbol{r}')\, \mathrm{d}\boldsymbol{r}''\, \mathrm{d}\boldsymbol{r}', \tag{XIX.158}$$

where

$$f_1(\boldsymbol{r}') \equiv \langle \boldsymbol{r}'p0|T_1|\boldsymbol{k}_0 00\rangle = \frac{1}{(2\pi)^3} \int \mathrm{e}^{\mathrm{i}\boldsymbol{k}'\cdot\boldsymbol{r}'} \langle \boldsymbol{k}'p0|T_1|\boldsymbol{k}_0 00\rangle\, \mathrm{d}\boldsymbol{k}' \tag{XIX.159}$$

$$f_2(\boldsymbol{r}'') \equiv \langle \boldsymbol{k}pq|T_2|\boldsymbol{r}''p0\rangle = \frac{1}{(2\pi)^3} \int \mathrm{e}^{-\mathrm{i}\boldsymbol{k}''\cdot\boldsymbol{r}''} \langle \boldsymbol{k}pq|T_2|\boldsymbol{k}''p0\rangle\, \mathrm{d}\boldsymbol{k}'' \tag{XIX.160}$$

$$g(\boldsymbol{r}'', \boldsymbol{r}') \equiv \langle \boldsymbol{r}''p0|G_0|\boldsymbol{r}'p0\rangle = -\frac{m}{2\pi\hbar^2} \frac{\mathrm{e}^{\mathrm{i}K_p|\boldsymbol{r}''-\boldsymbol{r}'|}}{|\boldsymbol{r}''-\boldsymbol{r}'|}. \tag{XIX.161}$$

In the last expression, K_p is the positive quantity given by

$$E = \frac{\hbar^2 K_p^2}{2m} + e_p + e_0.$$

We shall also use the vector $\mathbf{K}_p$, defined by

$$\mathbf{K}_p \equiv K_p \mathbf{R}/R.$$

Owing to the properties of T_1, the function $f_1(\mathbf{r}')$ is appreciably different to zero only in a domain of radius a about the origin; owing to the properties of T_2, $f_2(\mathbf{r}'')$ is appreciably different to zero only in a domain of radius a about the point $\mathbf{R}$. We can therefore replace $g(\mathbf{r}'', \mathbf{r}')$ in the integral (XIX.158) by the first term in its expansion in powers of $\mathbf{r}'/R$ and $(\mathbf{r}'' - \mathbf{R})/R$. In so doing we make an error of order a/R in the calculation of A_{21}. Since

$$|\mathbf{r}'' - \mathbf{r}'| \simeq R + \mathbf{R}\cdot(\mathbf{r}'' - \mathbf{R} - \mathbf{r}')/R + \ldots,$$

this leads to

$$g(\mathbf{r}'', \mathbf{r}') \simeq -\frac{m}{2\pi\hbar^2} \frac{e^{i\mathbf{K}_p\cdot(\mathbf{r}''-\mathbf{r}')}}{R}. \qquad \text{(XIX.162)}$$

Substituting expressions (XIX.159), (XIX.160) and (XIX.162) on the right-hand side of (XIX.158), and performing the integrations, we find:

$$A_{21} \simeq -(m/2\pi\hbar^2 R)\,\langle \mathbf{k}pq|T_2|\mathbf{K}_p p0\rangle\,\langle \mathbf{K}_p p0|T_1|\mathbf{k}_0 00\rangle. \qquad \text{(XIX.163)}$$

Similarly if we denote by $\mathbf{K}_q$ the vector defined by

$$\mathbf{K}_q \equiv -K_q \mathbf{R}/R, \qquad E = \frac{\hbar^2 K_q^2}{2m} + e_0 + e_q$$

we obtain for A_{12}, with an error of order a/R, the expression

$$A_{12} \simeq -(m/2\pi\hbar^2 R)\,\langle \mathbf{k}pq|T_1|\mathbf{K}_q 0q\rangle\,\langle \mathbf{K}_q 0q|T_2|\mathbf{k}_0 00\rangle. \qquad \text{(XIX.164)}$$

Note that the matrix elements of T_1 and T_2 occurring in expressions (XIX.163) and (XIX.164) correspond to transitions conserving the energy. By applying relations (XIX.150) and (XIX.151), we can therefore express A_{21} and A_{12} in terms of the transition amplitudes involved in the inelastic scattering by a single atom. This is also true for the cross section $\mathrm{d}\Sigma_{pq}/\mathrm{d}\Omega$ since

$$\begin{aligned}\frac{\mathrm{d}\Sigma_{pq}}{\mathrm{d}\Omega} &= \frac{2\pi}{\hbar v_0}\,|A_{21}+A_{12}|^2\,\varrho(E)\\ &= \frac{2\pi}{\hbar v_0}\,|A_{21}|^2\,\varrho(E)+\frac{2\pi}{\hbar v_0}\,|A_{12}|^2\,\varrho(E)+\frac{2\pi}{\hbar v_0}\,(A_{21}^*\,A_{12}+A_{21}\,A_{12}^*)\,\varrho(E).\end{aligned} \qquad \text{(XIX.165)}$$

We obtain the result

$$\begin{aligned} d\Sigma_{pq}/d\Omega = &(d\sigma_p(\boldsymbol{k}_0 \to \boldsymbol{K}_p)/R^2\,d\Omega)(d\sigma_q(\boldsymbol{K}_p \to \boldsymbol{k})/d\Omega) \\ &+ (d\sigma_q(\boldsymbol{k}_0 \to \boldsymbol{K}_q)/R^2\,d\Omega)(d\sigma_p(\boldsymbol{K}_q \to \boldsymbol{k})/d\Omega) \\ &+ \text{interference terms} \end{aligned} \qquad \text{(XIX.166)}$$

We leave to the reader the calculation of the interference terms. In this expression, the symbol $d\sigma_n(\boldsymbol{k} \to \boldsymbol{k}')/d\Omega$ denotes the cross section for the inelastic scattering $(\boldsymbol{k}0) \to (\boldsymbol{k}'n)$ of the particle by a single atom.

The first two terms of formula (XIX.166) are precisely those given by an elementary classical treatment. Classically, one can imagine the double scattering taking place in two different ways: either the particle is first scattered inelastically by atom 1 towards atom 2 and then scattered inelastically by atom 2 into the final direction, or, it is first scattered by atom 2 towards atom 1 and then by atom 1 into the final direction. The cross sections for these two processes are given respectively by the first and second terms of expression (XIX.166). To these we must add terms coming from the interference between the two types of scattered waves. Remarks analogous to those of § 25 may be made with respect to these interference terms. To observe them, the resolving power of the detector must be sufficient to distinguish angles of order λ/R.

If $\lambda \ll a$, the individual cross sections are essentially concentrated in the forward direction, the scattering amplitudes practically vanish for deflections greater than λ/a. Therefore – as is easily seen from inspection of equations (XIX.163–165) – double scattering can only occur if the two atoms are aligned to within λ/a in the incident direction, that is if $\boldsymbol{k}_0 \| \boldsymbol{R}$, or if $\boldsymbol{k}_0 \| (-\boldsymbol{R})$. In the first case A_{12} is negligible but A_{21} is appreciable for small deflections, that is, for $\boldsymbol{k} \| \boldsymbol{k}_0$ to within λ/a: the particle is first scattered by atom 1 and then by atom 2 into a direction nearly parallel to its incident direction. In the second case, the roles of atoms 1 and 2 are interchanged. The observation of the "trajectories" of ionizing particles in a Wilson cloud chamber is based on these results (cf. note 1, p. 142).

IV. VARIATIONAL CALCULATIONS OF TRANSITION AMPLITUDES

27. Stationary Expressions for the Phase Shift

The variational method has already been used for the determination of energy levels (Ch. XVIII). In the present section we give a brief

account of how it can be used to calculate phase shifts and, more generally, transition amplitudes. To do this, one must express these amplitudes as functionals of the scattering wave function which are stationary with respect to variations of this function about its correct value. The integral expressions for transition amplitudes given in the preceding section are not suitable for this purpose since they are not stationary. For example, expression (XIX.54) for T_l, considered as a functional of ψ_l, is not stationary when ψ_l is the correct solution of the radial equation; similarly, expression (XIX.127) for $T_{a\to b}$, considered as a functional of $\Psi_b^{(-)}$, is not stationary when $\Psi_b^{(-)}$ varies about its correct value. Several stationary expressions have been proposed for these quantities. Here we give the one, due to Schwinger [1]), which proves to be the most useful.

In this paragraph, we consider the case of a particle in a central potential and look for a stationary expression for the coefficient T_l of the expansion of the transition amplitude in spherical harmonics [expansion (XIX.51)]. Except for a few modifications to be described below we follow the notation of § 9.

If $\psi_a^{(+)}$ is the complete stationary scattering wave, the partial wave ψ_l satisfies the integral equation (XIX.56). With a view to later generalizations, we write the latter in the form

$$\psi_l(r) = j_l(kr) + g_l^{(+)} V \psi_l, \tag{XIX.167}$$

where $g_l^{(+)}$ is the integral operator whose kernel is the Green's function

$$g_l^{(+)}(r, r') \equiv -(2mk/\hbar^2)\, j_l(kr_<)\, h_l^{(+)}(kr_>), \tag{XIX.168}$$

in other words:

$$g_l^{(+)} V \psi_l \equiv \int_0^\infty g_l^{(+)}(r, r')\, V(r')\, \psi_l(r')\, r'^2\, \mathrm{d}r'.$$

We shall also use the notation $\langle \varphi_1, \varphi_2 \rangle$ to denote the scalar product of two radial functions φ_1, φ_2:

$$\langle \varphi_1, \varphi_2 \rangle \equiv \int_0^\infty \varphi_1^*(r)\, \varphi_2(r)\, r^2\, \mathrm{d}r.$$

---

[1]) See notably B. Lippmann and J. Schwinger, *loc. cit.*; a discussion of the respective merits of the different variational expressions proposed for the calculation of phase shifts is given by M. Moe and D. S. Saxon, Phys. Rev. 111 (1958) 950.

The integral form (XIX.54) for T_l can therefore be written

$$T_l = \langle j_l, V\psi_l \rangle. \tag{XIX.169}$$

Let us now put

$$A[\psi] \equiv \langle j_l, V\psi \rangle = \langle \psi^*, Vj_l \rangle = \int_0^\infty j_l(kr)\, V(r)\, \psi(r)\, r^2\, \mathrm{d}r, \tag{XIX.170}$$

$$B[\psi] \equiv \langle \psi^*, (V - Vg_l^{(+)}\, V)\, \psi \rangle = \int_0^\infty \psi^2(r) V(r)\, r^2\, \mathrm{d}r - \int_0^\infty \int_0^\infty (\psi(r)\, V(r)\, g_l^{(+)}(r, r')\, V(r')\, \psi(r'))\, r^2\, \mathrm{d}r\, r'^2\, \mathrm{d}r', \tag{XIX.171}$$

and consider the functional

$$\mathscr{T}_l[\psi] \equiv \frac{A^2}{B} \tag{XIX.172}$$

depending on the function ψ. The domain of variation of ψ is defined only by the condition that $r^2|\psi|^2$ be locally integrable.

As an obvious consequence of (XIX.167) and (XIX.169) we have

$$\mathscr{T}_l\,[\psi_l] = T_l.$$

Note too that $\mathscr{T}_l$ *depends neither on the normalization of the function* ψ (it does not change when ψ is multiplied by an arbitrary constant), *nor on the value taken by* ψ *in regions where* $V(r) = 0$. $\mathscr{T}_l$ therefore takes the same value T_l for any function ψ_l satisfying the following less restrictive condition than eq. (XIX.167):

$$\psi_l(r) = Cj_l(kr) + g_l^{(+)}\, V\psi_l \quad \text{if} \quad V(r) \neq 0 \tag{XIX.167a}$$

(C, arbitrary constant).

Let us calculate the variation of $\delta\mathscr{T}_l$ as a function of $\delta\psi$. From definition (XIX.172),

$$\delta\mathscr{T}_l = \frac{2A}{B}\,\delta A - \frac{A^2}{B^2}\,\delta B.$$

Now

$$\delta A = \int_0^\infty \delta\psi\, V(r)\, j_l(kr)\, r^2\, \mathrm{d}r = \langle \delta\psi^*, Vj_l \rangle$$

and, taking into account that V is real and $g_l^{(+)}(r, r')$ symmetrical in r and r',

$$\delta B = 2\langle \delta\psi^*, (V - Vg_l^{(+)} V)\, \psi \rangle.$$

We may therefore write

$$\delta\mathscr{T}_l = \frac{2A}{B^2}\langle\delta\psi^*, F\rangle,$$

where

$$F(r) = B[\psi]V(r)j_l(kr) - A[\psi]V(r)(\psi - g_l^{(+)}V\psi).$$

In order to have $\delta\mathscr{T}_l = 0$ for any $\delta\psi$, it is necessary and sufficient that $F(r)$ vanish. For this it is necessary that $j_l(kr)$ and $\psi - g_l^{(+)}V\psi$ be proportional at any point where $V(r)$ is not zero, that is, that ψ be one of the functions ψ_l obeying equation (XIX.167*a*). One easily verifies that this is also a sufficient condition. *Thus, the stationary value of $\mathscr{T}_l$ is equal to the amplitude that we wish to calculate*:

$$T_l = \mathscr{T}_l|_{\text{st}} \qquad \text{(XIX.173)}$$

In order to introduce only real functions into the calculation, we separate the real and imaginary parts of $g_l^{(+)}(r, r')$. The real part is the Green's function

$$g_l^{(1)}(r, r') \equiv -(2mk/\hbar^2)\,j_l(kr_<)\,n_l(kr_>)$$

and we have

$$g_l^{(+)}(r, r') = g_l^{(1)}(r, r') - \mathrm{i}(2mk/\hbar^2)\,j_l(kr)\,j_l(kr').$$

Substituting this expression in the definition of $\mathscr{T}_l$ (i.e. in expression (XIX.171) for $B[\psi]$), we can rewrite (XIX.171) in the equivalent form

$$T_l^{-1} = \mathscr{T}_l^{-1}\Big|_{\text{st}} = \frac{\langle\psi^*, (V - Vg_l^{(1)}V)\,\psi\rangle}{\langle j_l, V\psi\rangle^2}\Big|_{\text{st}} + \mathrm{i}\,\frac{2mk}{\hbar^2}$$

and since from equation (XIX.52)

$$T_l^{-1} - \mathrm{i}(2mk/\hbar^2) = -2mk\cot\delta_l/\hbar^2,$$

we obtain the following stationary expression for $k\cot\delta_l$:

$$k\cot\delta_l = -\frac{\hbar^2}{2m}\,\frac{\langle\psi^*, (V - Vg_l^{(1)}V)\,\psi\rangle}{\langle j_l, V\psi\rangle^2}\Big|_{\text{st}}$$

$$= -\frac{\hbar^2}{2m}\,\frac{\int_0^\infty \psi^2 Vr^2\,\mathrm{d}r - \int_0^\infty\int_0^\infty (\psi(r)\,V(r)\,g_l^{(1)}(r, r')\,V(r')\,\psi(r'))\,r^2\,\mathrm{d}r\,r'^2\,\mathrm{d}r'}{\left[\int_0^\infty j_l(kr)\,V(r)\,\psi(r)\,r^2\,\mathrm{d}r\right]^2}\Big|_{\text{st}}$$

(XIX.174)

It can easily be shown that the function ψ_l for which the right-hand side of equation (XIX.174) is stationary can be taken to be real, and that this reality condition for the trial functions does not modify the variational property obtained.

28. The Variational Calculation of Phase Shifts. Discussion

Equation (XIX.174) is the starting point of a variational method for calculating the phase shifts. The functional on the right-hand side is calculated by substituting for ψ a trial function φ depending on several parameters and taking the stationary value of the function with respect to variations of these parameters. The closer the trial function φ to the exact solution ψ the closer will be the resulting approximate value for $k \cot \delta_l$ to the exact one. As has already been pointed out, and here we have one of the chief virtues of the Schwinger variational method, the result is independent of the normalization of φ and of the values taken by φ in regions where the potential V vanishes. For the method to be precise it is therefore sufficient to take a trial function whose general form is the same as the exact solution in the region of the potential. The estimate of the error is based essentially on these semi-quantitative considerations and is, as in the variational calculation of energy levels, largely empirical.

If we limit ourselves to substituting the free wave $j_l(kr)$ for ψ in the right-hand side of (XIX.174), we find

$$k \cot \delta_l \equiv -\frac{\hbar^2}{2m} \frac{1-\Delta_l}{\langle j_l, V j_l \rangle} \tag{XIX.175}$$

putting

$$\Delta_l \equiv \frac{\langle j_l, V g_l^{(1)} V j_l \rangle}{\langle j_l, V j_l \rangle}. \tag{XIX.176}$$

This formula is *a priori* more exact than the one given by the Born approximation [eq. (X.75)]. In the limit when $\Delta_l \ll 1$, it is equivalent to the second-order Born approximation, but when Δ_l ceases to be small it is often much better.

To illustrate the comparison between these two methods let us treat the s-wave scattering by a square well in the low energy limit. We take

$$V(r) = \begin{cases} -V_0 & r < r_0 \\ 0 & r > r_0. \end{cases} \tag{XIX.177}$$

We shall calculate the scattering lengths

$$a = -\lim_{k\to 0} (k \cot \delta)^{-1}$$

by the different methods cited above and compare the results with those given by the exact calculation. The calculations are not difficult. If we put $b = (2m V_0 r_0^2/\hbar^2)^{\frac{1}{2}}$, the results are given by the following formulas:

exact calculation: $$a = -\left(\frac{\tan b}{b} - 1\right) r_0$$

variational calculation [formula (XIX.175)]: $$a_{\text{var}} = -\frac{\frac{1}{3}b^2}{1-\frac{2}{5}b^2}\, r_0$$

Born approximation: $$a_{\text{B}} = -\tfrac{1}{3}\, b^2\, r_0$$

second-order Born approximation: $$a_{\text{B}}^{(2)} = -(\tfrac{1}{3}\, b^2 + \tfrac{2}{15}\, b^4)\, r_0.$$

TABLE XIX.1

Comparison of various calculations of the scattering length as a function of the parameter

$$b = (2m V_0 r_0^2/\hbar^2)^{\frac{1}{2}}.$$

r_0 has been taken as the unit of length.

b	Exact calculation a	Stationary formula (XIX.175) a_{var}	Born Approximation 2nd order $a_{\text{B}}^{(2)}$	Born Approximation 1st order $a_{\text{B}}^{(1)}$
0	0	0	0	0
$0.1\,\pi$	− 0.034	− 0.034	− 0.034	− 0.033
$0.2\,\pi$	− 0.156	− 0.156	− 0.152	− 0.132
$0.3\,\pi$	− 0.460	− 0.459	− 0.401	− 0.296
$0.4\,\pi$	− 1.449	− 1.428	− 0.859	
$\frac{\pi}{2}$	∞ (*)	− 63.2 (**)	− 1.63	
$0.6\,\pi$	2.63	2.81	− 2.9	
$0.7\,\pi$	1.63	1.72		
$0.8\,\pi$	1.29	1.38		
$0.9\,\pi$	1.11	1.21		
π	1.00	1.12		

(*) $a = \infty$ corresponds to the existence of a bound state of zero energy.
(**) a_{var} becomes infinite and changes sign for $b = 0.503\,\pi$.

The validity criterion for the Born approximation is given by inequality (XIX.43), namely:

$$b \ll 1$$

(b is the measure of the depth of the potential, that is, of the number of bound s-states). The numerical results corresponding to the four formulas above are given in Table (XIX.1). Note that a_{var} remains an excellent approximation for relatively large value of b, up to and including the region $\frac{1}{2}\pi < b < \pi$ in which the Born expansion certainly does not converge.

29. Extension to Complex Collisions

The preceding theory can be extended to more general collision problems. We shall not enter into the details of such possible extensions but shall limit ourselves to giving a stationary expression for the transition matrix in the case of the elastic or inelastic scattering of two complex particles.

Unless otherwise stated, we follow the notation of Section III. The amplitude for the transition $a \to b$ is given by

$$T_{a\to b} = \langle \Phi_b | V_\beta | \Psi_a^{(+)} \rangle = \langle \Psi_b^{(-)} | V_\alpha | \Phi_a \rangle.$$

These expressions are not stationary with respect to variations of $\Psi_a^{(+)}$ or $\Psi_b^{(-)}$. Since we are concerned with a scattering process, the unperturbed Hamiltonians are equal; we therefore put

$$H_\alpha = H_\beta = H_0 \qquad V_\alpha = V_\beta = V$$
$$G_0^{(+)} = G_0^{(-)\dagger} = [E - H_0 + i\varepsilon]^{-1}.$$

One knows that

$$|\Phi_a\rangle = (1 - G_0^{(+)} V)\,|\Psi_0^{(+)}\rangle$$
$$\langle \Phi_b| = \langle \Psi^{(-)}|\,(1 - V G_0^{(+)}).$$

The amplitude $T_{a\to b}$ is also given by the expression

$$T_{a\to b} = \frac{\langle \Psi_b^{(-)}|V|\Phi_a\rangle\,\langle \Phi_b|V|\Psi_a^{(+)}\rangle}{\langle \Psi_b^{(-)}|(V - V\,G_0^{(+)}\,V)|\Psi_a^{(+)}\rangle}. \qquad \text{(XIX.178)}$$

This expression is stationary with respect to independent variations of the functions $\Psi_a^{(+)}$ and $\Psi_b^{(-)}$. It is the generalization of the stationary expression for T_l given in § 27.

Except in the simplest cases, such as the collision of two elementary

particles, one encounters a major difficulty in the practical use of this expression, that of obtaining an explicit expression for the Green function $G_0^{(+)}$. However, it may sometimes be useful in arguments where an explicit use of $G_0^{(+)}$ is not required [1]). In any case its applications are rather limited.

V. GENERAL PROPERTIES OF THE TRANSITION MATRIX

Certain properties of the T matrix follow directly from the characteristic properties of the Hamiltonian governing the collision. Some of them have already been indicated in the preceding sections. However their character of great generality has not been stressed. This question will systematically be investigated in the present section.

30. Conservation of Flux. Unitarity of the S Matrix

Certain properties of the T matrix are a simple consequence of the Hermitean character of the Hamiltonian H governing the collision. Among these are the integral representations (XIX.121) and (XIX.122). By using the same method as in § 19 we shall obtain two new relations, called conservation of flux relations.

Following the notations of § 19 we consider two stationary waves $\Psi_a^{(+)}$, $\Psi_b^{(+)}$ corresponding to the *same* value E of the energy. We have

$$\Psi_b^{(+)*}(H\Psi_a^{(+)}) - (H\Psi_b^{(+)})^*\Psi_a^{(+)} = 0.$$

The quantity obtained by summing over the spins and integrating over a finite volume of configuration space therefore vanishes. By applying Green's theorem this quantity can be transformed into a surface integral, which takes the form of a sum of terms relative to the different channels in the limit when that surface tends to infinity (cf. note, p. 837). We thereby obtain

$$0 = \sum_\nu \frac{\hbar^2}{2M_\nu} \lim_{R_\nu \to \infty} \{\Psi_b^{(+)*}, \Psi_a^{(+)}\}_\nu. \qquad \text{(XIX.179)}$$

The right-hand side of this equation is to be compared with that of

---

[1]) Cf. B. Lippmann and J. Schwinger, *loc. cit.*

(XIX.119). It can be evaluated by replacing the functions $\Psi_b^{(+)}$, $\Psi_a^{(+)}$ by their asymptotic form in each channel, giving

$$0 = \frac{\hbar^2}{2M_\beta} \lim_{R_\beta \to 0} \left\{ e^{-ik_b \cdot r_\beta}, f_{a\beta}^{(+)}(\Omega_\beta) \frac{e^{ik_b r_\beta}}{r_\beta} \right\}_\beta$$
$$+ \frac{\hbar^2}{2M_\alpha} \lim_{R_\alpha \to 0} \left\{ f_{b\alpha}^{(+)*}(\Omega_\alpha) \frac{e^{-ik_a r_\alpha}}{r_\alpha}, e^{ik_a \cdot r_\alpha} \right\}_\alpha$$
$$+ \sum_\nu \frac{\hbar^2}{2M_\nu} \lim_{R_\nu \to 0} \left\{ f_{b\nu}^{(+)*}(\Omega_\nu) \frac{e^{-ik_n r_\nu}}{r_\nu}, f_{a\nu}^{(+)}(\Omega_\nu) \frac{e^{ik_n r_\nu}}{r_\nu} \right\}_\nu .$$

The evaluation of the contribution of the first two terms is patterned on the work of § 2. The third term is a sum over all the channels, α and β included, which is easily evaluated from the definition of the symbol $\{...\}_\nu$; contrary to what occurs in the calculation of § 19, the contribution from this term does not vanish. We finally obtain

$$0 = -\frac{2\pi\hbar^2}{M_\beta} f_{a\beta}^{(+)}(\Omega_b) + \frac{2\pi\hbar^2}{M_\alpha} f_{b\alpha}^{(+)*}(\Omega_a) + \sum_\nu \frac{i\hbar^2 k_n}{M_\nu} \int f_{b\nu}^{(+)*}(\Omega_\nu) f_{a\nu}^{(+)}(\Omega_\nu) \, d\Omega_\nu .$$

Replacing the $f(\Omega)$ by the elements of the T matrix in accordance with definition (XIX.114), we obtain

$$0 = \langle b|T|a\rangle - \langle b|T^\dagger|a\rangle + 2\pi i \sum_\nu \varrho_\nu(E) \int \langle b|T^\dagger|n\rangle \, d\Omega_\nu \, \langle n|T|a\rangle . \quad \text{(XIX.180)}$$

Here we have denoted by n the plane wave in channel ν directed along Ω_ν and used the definition of the density of states

$$\varrho_\nu(E) = \frac{M_\nu k_n}{(2\pi)^3 \hbar^2} .$$

$T^\dagger$ is given by the usual definition of Hermitean conjugation:

$$\langle m|T^\dagger|n\rangle = \langle n|T|m\rangle^* .$$

The preceding work can also be carried through with the $\Psi^{(-)}$ in place of the $\Psi^{(+)}$. This gives the relation

$$0 = \sum_\nu \frac{\hbar^2}{2M_\nu} \lim_{R_\nu \to \infty} \{\Psi_a^{(-)}, \Psi_b^{(-)*}\}_\nu \qquad \text{(XIX.179')}$$

which is obtained from relation (XIX.179) by exchanging a and b and replacing $\Psi^{(+)}$ by $\Psi^{(-)*}$. To evaluate the right-hand side, it suffices to make the necessary substitutions in the preceding cal-

culation, and to use definition (XIX.114′) of the matrix elements of T in place of (XIX.114). One obtains:

$$0 = \langle b|T|a\rangle - \langle b|T^\dagger|a\rangle + 2\pi i \sum_\nu \varrho_\nu(E) \int \langle b|T|n\rangle \, d\Omega_\nu \langle n|T^\dagger|a\rangle. \quad \text{(XIX.180′)}$$

Let us give a second, more formal, demonstration of relations (XIX.180) and (XIX.180′). From the relations

$$\langle b|T|a\rangle = \langle \Phi_b|V_\beta|\Phi_a\rangle + \langle \Phi_b|V_\beta \frac{1}{E - H + i\varepsilon} V_\alpha|\Phi_a\rangle$$

$$\langle b|T^\dagger|a\rangle = \langle \Phi_b|V_\alpha|\Phi_a\rangle + \langle \Phi_b|V_\beta \frac{1}{E - H - i\varepsilon} V_\alpha|\Phi_a\rangle$$

we obtain, subtracting both sides and applying (XIX.124) and (A.15c):

$$\langle b|(T - T^\dagger)|a\rangle = -2\pi i \langle \Phi_b|V_\beta \delta(E - H) V_\alpha|\Phi_a\rangle. \quad \text{(XIX.181)}$$

We suppose that the set of stationary waves $\Psi_n^{(-)}$ *extended over the whole of the spectrum of* H – and complemented, if it exists, by a complete suitably normalized set of eigenvectors of the discrete spectrum – forms a complete orthogonal set of eigenvectors of H, and therefore that we have the closure relation

$$(2\pi)^{-3} \sum_n |\Psi_n^{(-)}\rangle\langle\Psi_n^{(-)}| = 1. \quad \text{(XIX.182)}$$

[The sum being extended to all the energies of the spectrum of H, discrete spectrum included.] We shall also suppose that the $\Psi_n^{(+)}$ satisfy the closure relation

$$(2\pi)^{-3} \sum_n |\Psi_n^{(+)}\rangle\langle\Psi_n^{(+)}| = 1. \quad \text{(XIX.182′)}$$

Using (XIX.182), we can put the scalar product on the right-hand side of (XIX.181) in the form

$$\langle \Phi_b|V_\beta \delta(E-H) V_\alpha|\Phi_a\rangle \equiv (2\pi)^{-3} \sum_n \langle \Phi_b|V_\beta|\Psi_n^{(-)}\rangle \delta(E - E_n) \langle \Psi_n^{(-)}|V_\beta|\Phi_a\rangle$$

$$= \sum_\nu \varrho_\nu(E) \int \langle b|T^\dagger|n\rangle \, d\Omega_\nu \, \langle n|T|a\rangle$$

from which we deduce relation (XIX.180). Relation (XIX.180′) is obtained in the same way using (XIX.182′).

Equations (XIX.180) and (XIX.180′) are satisfied by the matrix

elements of T and $T^\dagger$ for any a and b. They can be written more simply in the form:

$$T - T^\dagger + 2\pi i T^\dagger T = 0 \qquad \text{(XIX.183)}$$

$$T - T^\dagger + 2\pi i T T^\dagger = 0 \qquad \text{(XIX.183')}$$

with a suitable summation convention for the dummy indices in the matrix products $T^\dagger T$ and $TT^\dagger$, i.e. by weighting each index ν by the factor $\varrho_\nu(E)$ giving the density of states in the channel ν. Note that the matrix elements appearing here are all taken between states of energy E.

We now introduce the S matrix:

$$S \equiv 1 - 2\pi i T. \qquad \text{(XIX.184)}$$

This matrix is characteristic of collisions of energy E and its elements, like those of T, are defined between states of energy E. Expressing T and $T^\dagger$ in equations (XIX.180) and (XIX.180') in terms of S and $S^\dagger$ we find:

$$S^\dagger S = S S^\dagger = 1. \qquad \text{(XIX.185)}$$

The S matrix is unitary.

It can also be shown that if $U(t, t')$ is the evolution operator for the system then

$$S = \lim_{\substack{t \to +\infty \\ t' \to -\infty}} U(t, t').$$

The demonstration of this requires great care in passing to the limit; it will not be given here [cf. references note p. 801].

31. The Bohr–Peierls–Placzek Relation (Optical Theorem)

Consider relation (XIX.180) in the particular case when $a = b$ [N.B. relations (XIX.180) and (XIX.180') are identical when $a = b$]. We then have

$$\langle a|T|a\rangle - \langle a|T|a\rangle^* = -2\pi i \sum_\nu \int |\langle n|T|a\rangle|^2 \varrho_\nu(E)\, \mathrm{d}\Omega_\nu. \qquad \text{(XIX.186)}$$

The integrand on the right-hand side is to within a factor $2\pi/\hbar v_a$ the cross section $\mathrm{d}\sigma_{a\to n}/\mathrm{d}\Omega_\nu$, and therefore

$$\sum_\nu \int |\langle n|T|a\rangle|^2 \varrho_\nu(E)\, \mathrm{d}\Omega_\nu = \frac{\hbar v_a}{2\pi} \sum_\nu \int \frac{\mathrm{d}\sigma_{a\to n}}{\mathrm{d}\Omega_\nu}\, \mathrm{d}\Omega_\nu$$

$$= \frac{\hbar v_a}{2\pi}\, \sigma_a^{\mathrm{tot}},$$

where σ_a^{tot} is the total cross section for the collision corresponding to initial state a. Returning to (XIX.186) we have

$$\sigma_a^{\text{tot}} = -\frac{2}{\hbar v_a} \operatorname{Im} T_{aa}.$$

To put this expression in a more familiar form we express T_{aa} in terms of the scattering amplitude in the forward direction $f_{aa}^{(+)}(\Omega_a)$, which we denote simply by $f_a(0)$ [cf. eq. (XIX.114)]:

$$T_{aa} \equiv \langle a|T|a\rangle = -\frac{2\pi\hbar^2}{M_\alpha} f_a(0).$$

Thus

$$\boxed{\sigma_a^{\text{tot}} = \frac{4\pi}{k_a} \operatorname{Im} f_a(0)} \qquad \text{(XIX.187)}$$

This is the Bohr–Peierls–Placzek formula relating the total cross section to the imaginary part of the scattering amplitude in the forward direction.

32. Microreversibility

If H is invariant under time reversal, the T matrix satisfies the *microreversibility relation* (cf. § XV.20):

$$\boxed{T_{Kb\to Ka} = T_{a\to b}} \qquad \text{(XIX.188)}$$

The demonstration of this relation has already been given for the simple case of the scattering of a particle by a real potential (§ 3 and 14). The formal demonstration of § 14 can also be used in the general case with very little modification. K being the time-reversal operator, the initial and final states that are obtained from a and b by time reversal (reversal of momenta and spins) will be denoted by Ka and Kb respectively; more precisely

$$\Phi_{Kb} \equiv K\Phi_b \qquad \Phi_{Ka} \equiv K\Phi_a.$$

To prove relation (XIX.188), we make use of expression (XIX.142) giving the amplitude for a given transition:

$$T_{Kb\to Ka} = \langle \Phi_{Ka}| \left[V_\alpha + V_\alpha \frac{1}{E - H + i\varepsilon} V_\beta \right] |\Phi_{Kb}\rangle$$

$$= (\langle \Phi_a|K^\dagger) \left[V_\alpha + V_\alpha \frac{1}{E - H + i\varepsilon} V_\beta \right] (K|\Phi_b\rangle).$$

Since K is antiunitary and commutes with H, V_α and V_β this gives

$$T_{Kb\to Ka} = \langle \Phi_b | \left[V_\alpha + V_\beta \frac{1}{E - H + i\varepsilon} V_\alpha \right] | \Phi_a \rangle$$
$$= \langle \Phi_b | \left[V_\beta + V_\beta \frac{1}{E - H + i\varepsilon} V_\alpha \right] \Phi_a \rangle.$$

The second line was obtained by applying (XIX.124), and is equal to the amplitude $T_{a\to b}$ as given by expression (XIX.142). Q.E.D.

By squaring both sides of (XIX.188) and expressing the result in terms of the cross sections $d\sigma_{a\to b}/d\Omega_\beta$ and $d\sigma_{Kb\to Ka}/d\Omega_\alpha$, we obtain the microreversibility relation for cross sections:

$$v_b \, d\sigma_{Kb\to Ka}/\varrho_a(E)\, d\Omega_\alpha = v_a \, d\sigma_{a\to b}/\varrho_b(E)\, d\Omega_\beta. \quad \text{(XIX.189)}$$

Here we have used the fact that $v_{Kb} = v_b$, $\varrho_{Ka} = \varrho_a$. From the definition of velocity and density of states [eq. (XVII.52)] we have for each channel

$$v_n \varrho_n = p_n^2/(2\pi\hbar)^3 = k_n^2/(2\pi)^3 \hbar.$$

Therefore, relation (XIX.189) can be written, with the convention that $d\Omega_\alpha = d\Omega_\beta = d\Omega$,

$$\boxed{k_b^2 \, d\sigma_{Kb\to Ka} = k_a^2 \, d\sigma_{a\to b}} \quad \text{(XIX.190)}$$

33. Invariance Properties of the T Matrix

Let us now investigate the consequences of the invariance of H with respect to rotations, reflections and more generally with respect to transformations represented by unitary operators. We successively consider the consequences of invariance with respect to a given transformation and with respect to a specified group of transformations.

If H, H_α and H_β are invariant in a given unitary transformation X one has the relation (cf. § XV.13)

$$T_{Xa\to Xb} = T_{a\to b}. \quad \text{(XIX.191)}$$

[N.B. This relation differs from the microreversibility relation by the sense of the reaction corresponding to the amplitude on the left-hand side.]

To prove (XIX.191), one needs only to observe that according to definition (XIX.144) the operator $T^{\beta\alpha}$ is also invariant in the transformation X:

$$X^\dagger T^{\beta\alpha} X = T^{\beta\alpha},$$

and therefore

$$T_{Xa\to Xb} \equiv \langle \Phi_b | X^\dagger T^{\beta\alpha} X | \Phi_a \rangle = \langle \Phi_b | T^{\beta\alpha} | \Phi_a \rangle \equiv T_{a\to b}.$$

From (XIX.191), one easily obtains the following relation between cross sections:

$$d\sigma_{Xa\to Xb} = d\sigma_{a\to b} \qquad \text{(XIX.192)}$$

to be compared with (XIX.190).

Suppose now that the total Hamiltonian H and the different "unperturbed" Hamiltonians $H_\alpha, H_\beta, \ldots$, are invariant with respect to the transformations of a certain group G. As we shall see this results in certain properties of the T matrix analogous to those obtained in § XV.11 for the matrices representing observables invariant with respect to G.

To see this, we adopt the notation of § XV.11 and consider the two sets of observables J and M defined in that paragraph. Let ν be one of the channels open to collisions of energy E. H_ν, J and M form a set of commuting observables. Denote by $|\nu E_\nu \tau_\nu j\mu\rangle$ the vectors of a common basis (we shall assume that it is a standard basis suited to the group G, although this restriction is not essential). The index ν indicates the channel considered, E_ν is the eigenvalue of H_ν, j and μ denote the eigenvalues of J and M respectively, τ_ν is an additional quantum number which varies in a domain depending on the values of E_ν, j, and μ. These vectors are orthonormal:

$$\langle \nu E_\nu \tau_\nu j\mu | \nu E'_\nu \tau'_\nu j'\mu' \rangle = \delta(E_\nu - E'_\nu)\, \delta_{\tau_\nu \tau'_\nu}\, \delta_{jj'}\, \delta_{\mu\mu'}.$$

For simplicity we shall suppose the index τ_ν discrete. This is the case if we are dealing with a 2-particle channel and if the group G is the rotation group or contains the rotation group as a subgroup. τ_ν then takes a *finite* number of discrete values.

Consider now a collision at a fixed energy E and the transitions from a certain channel ν to another channel ν'; according to the

hypotheses made above, the corresponding operator $T^{\nu\nu'}$ [definition (XIX.144)] is invariant with respect to the transformations of the group G. By an obvious generalization of property (XV.52) we have

$$\langle \nu E \tau_\nu j \mu | T^{\nu\nu'} | \nu' E \tau_{\nu'} j' \mu' \rangle = \delta_{jj'} \delta_{\mu\mu'} \langle \nu \tau_\nu | T^{(j)} | \nu' \tau_{\nu'} \rangle, \qquad \text{(XIX.193)}$$

where $\langle \nu \tau_\nu | T^{(j)} | \nu' \tau_{\nu'} \rangle$ depends on the quantum numbers j, τ_ν and $\tau_{\nu'}$, but not on μ. Equation (XIX.193) sums up all the consequences of the invariance with respect to the group G.

To each pair $(\nu\nu')$ of open channels, there therefore corresponds a certain number of coefficients $\langle \nu \tau_\nu | T^{(j)} | \nu' \tau_{\nu'} \rangle$. Those of these coefficients corresponding to a certain given value of j (j and E fixed; ν, τ_ν, ν' and $\tau_{\nu'}$ variable) form a certain square matrix which we denote by $T^{(j)}$. This matrix is analogous to the T matrix but the number of its dimensions is in general much smaller. In particular if G contains the rotation group and if none of the open channels has more than two particles, $T^{(j)}$ is a finite matrix; if in addition there is but one open channel and if the two particles have spin zero, $T^{(j)}$ has only one dimension, i.e. it is an ordinary number, equal to within an easily determined factor to the coefficient T_j defined by expansion (XIX.51).

It is easy to expand the transition amplitude $T_{a \to b}$ in a series of elements of the $T^{(j)}$ matrices. To this end we start from definition (XIX.145), which we transform using the closure relations for the standard bases corresponding to channels α and β, and make use of property (XIX.193). The coefficients of the expansion depend on the projections of the plane waves Φ_a and Φ_b on the basis vectors in question. Φ_n denoting a plane wave of energy E in channel ν, we can write

$$\langle \Phi_n | \nu E_\nu \tau_\nu j \mu \rangle = \delta(E - E_\nu) \langle \Phi_n | \nu \tau_\nu j \mu \rangle. \qquad \text{(XIX.194)}$$

With this notation, we find, after a brief calculation

$$T_{a \to b} = \sum_{j\mu} \sum_{\tau_\alpha \tau_\beta} \langle \Phi_b | \beta \tau_\beta j \mu \rangle \langle \beta \tau_\beta | T^{(j)} | \alpha \tau_\alpha \rangle \langle \alpha \tau_\alpha j \mu | \Phi_a \rangle. \qquad \text{(XIX.195)}$$

This expansion is the generalization of (XIX.51). Like the latter it is particularly useful when convergence is rapid; this usually occurs in nuclear collisions when the wavelength in the entrance and exit channels is large with respect to the range of nuclear forces, and it is here that it is most often used [cf. references note 1, p. 832].

EXERCISES AND PROBLEMS

1. Demonstrate the asymptotic property

$$\exp(i\boldsymbol{q}\cdot\boldsymbol{r}) \underset{r\to\infty}{\sim} \frac{2\pi}{iqr}\left[\delta(\Omega_r-\Omega_q)\,e^{iqr}-\delta(\Omega_r+\Omega_q)\,e^{-iqr}\right]+O\left(\frac{1}{r^2}\right).$$

(*i*) using the expansion of the plane wave in spherical harmonics and the closure relations for the latter;

(*ii*) looking directly for the asymptotic form of

$$\int \exp(i\boldsymbol{q}\cdot\boldsymbol{r})\,\varphi(\Omega_r)\,d\Omega_r.$$

[$\varphi(\Omega)$, continuous function of the direction $\Omega \equiv (\theta\varphi)$; Ω_r and Ω_q, respective directions of the vectors $\boldsymbol{r}$ and $\boldsymbol{q}$.]

2. Show that for two plane waves $\exp(i\boldsymbol{q}_1\cdot\boldsymbol{r})$, $\exp(i\boldsymbol{q}_2\cdot\boldsymbol{r})$ of the same wavelength ($q_1 = q_2$), we have the exact relation

$$\{\exp(i\boldsymbol{q}_1\cdot\boldsymbol{r}),\ \exp(i\boldsymbol{q}_2\cdot\boldsymbol{r})\} = 0$$

whatever R. [definition of $\{\dots\}$ of § 2].

3. Calculate the scattering cross section in the Born approximation

(*i*) for the square well of depth V_0 and radius a;

(*ii*) for the Gaussian potential

$$V = V_0 \exp(-r^2/a^2);$$

(*iii*) for the Yukawa potential

$$V = \frac{V_0 \exp(-Kr)}{Kr}.$$

Verify for these three particular cases the assertions of § 4 concerning the general form of the high energy angular distribution; determine the asymptotic form of the total cross section.

4. Consider the Hamiltonian of a particle of mass m in a short range potential $V(r)$, ($rV \underset{r\to\infty}{\to} 0$)

$$H \equiv -(\hbar^2/2m)\Delta + V(r).$$

At energy $E = \hbar^2k^2/2m$, the radial equation of order l is

$$\left[\frac{d^2}{dr^2} - \frac{l(l+1)}{r^2} - \frac{2m}{\hbar^2}V + k^2\right] y_l = 0.$$

The regular solution is denoted by $F_l(k;r)$, the "outgoing" and "incoming"

solutions by $u_l^{(+)}(k;r)$ and $u_l^{(-)}(k;r)$ respectively. These solutions are normalized so that

$$F_l(k;r) \underset{r\to\infty}{\sim} \sin(kr - \tfrac{1}{2}l\pi + \delta_l)$$

$$u_l^{(\pm)}(k;r) \underset{r\to\infty}{\sim} \exp[\pm i(kr - \tfrac{1}{2}l\pi)].$$

(*i*) Show that the stationary waves $\psi_{\mathbf{k}}^{(\pm)}$ (definitions of § 2) are given by the expansions

$$\Psi_{\mathbf{k}}^{(\pm)} = \frac{4\pi}{kr}\sum_{lm} i^l \exp(\pm i\delta_l)\, Y_l^{m*}(\hat{\mathbf{k}})\, Y_l^m(\hat{\mathbf{r}})\, F_l(k;r).$$

(*ii*) Show that the functions $G^{(\pm)}(\mathbf{r},\mathbf{r}')$ defined by

$$\frac{\hbar^2}{2m}G^{(\pm)}(\mathbf{r},\mathbf{r}') = -\frac{1}{krr'}\sum_{lm}\exp(\pm i\delta_l)Y_l^{m*}(\hat{\mathbf{r}})Y_l^m(\hat{\mathbf{r}}')F_l(k;r_<)u_l^{(\pm)}(k;r_>)$$

are the Green functions for the operator $E - H$, that is, functions symmetrical in $\mathbf{r}$ and $\mathbf{r}'$ satisfying the equation

$$(E - H)\,G^{(\pm)}(\mathbf{r},\mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}')$$

and that their asymptotic behavior is given by

$$G^{(\pm)}(\mathbf{r},\mathbf{r}') \underset{r\to\infty}{\sim} -\frac{2m}{\hbar^2}\frac{\exp(\pm ikr)}{r}\,\psi_{\pm\mathbf{k}}^{(\mp)}(\mathbf{r}').$$
$$(\mathbf{k} = k\mathbf{r}/r)$$

[These formulas also apply when V behaves asymptotically like $1/r$ if we suitably re-define F_l and $u_l^{(\pm)}$; the asymptotic form of $G^{(\pm)}$ must be re-defined accordingly.]

5. Prove formula (XIX.151).

"I am black, but comely, . . ."
(*Song of Songs*, 1 : 5)

PART FIVE

ELEMENTS OF RELATIVISTIC QUANTUM MECHANICS

CHAPTER XX

THE DIRAC EQUATION

I. GENERAL INTRODUCTION

1. Relativistic Quantum Mechanics [1])

All of the applications made up to the present have been based on the Schrödinger equation. This equation, deduced by the correspondence principle from the Hamiltonian formalism of non-relativistic Classical Mechanics, has all the invariance properties of the Hamiltonian from which it derives. In particular, if the system is isolated, it is invariant under spatial rotations and translations. It can also be shown that it is invariant under Galilean transformations (cf. Problem XV.7). Therefore, the physical properties predicted by the Schrödinger theory are invariant in a Galilean change of referential, but they do not have the invariance under a Lorentz change of referential required by the principle of relativity. Since the Galilean transformation approximates to the Lorentz transformation only in the limit of small velocities, one expects – and experiment verifies – that this theory will correctly describe phenomena only when the velocities of the particles involved are negligible beside the velocity of light: $v \ll c$. In particular, all phenomena concerning the interaction between light and matter, such as emission, absorption or scattering of photons, is outside the framework of non-relativistic Quantum Mechanics.

One of the main difficulties in elaborating relativistic Quantum Mechanics comes from the fact that the *law of conservation of the number of particles ceases in general to be true.* Due to the equivalence of mass and energy, one of the most important consequences of relativity, there can be creation or absorption of particles whenever the interactions give rise to energy transfers equal or superior to the rest masses of these particles. To be a complete theory, Relativistic Quantum Mechanics must encompass in a single scheme dynamical

[1]) Knowledge of Sections I and II of Appendix D is recommended for reading this chapter.

states differing not only by the quantum state, but also by the *number* and the *nature* of the elementary particles of which they are composed. For this, we must turn to the concept of the quantized field, whence the name of Quantum Field Theory currently given to Relativistic Quantum Mechanics. This theory, in its present form, is exempt neither of difficulties nor even of contradictions, but it accounts for a very large body of experimental facts.

The fifth and last part of this book is designed to serve as an introduction to Quantum Field Theory and at the same time to furnish elementary methods for calculating certain relativistic effects concerning the dynamics of the electron and the interaction between the electromagnetic field and charged particles.

It is made up of two chapters.

The present chapter, the first of the two, is devoted to one of the simplest problems in Relativistic Quantum Mechanics, the problem of a particle of spin $\frac{1}{2}$ in a given force field. One of the most important examples is the electron in an electromagnetic field. The field is not quantized and one tries to describe the evolution of the system with a wave equation having the invariance properties required by the principle of relativity. This equation must also satisfy the correspondence principle and give the Pauli theory in the non-relativistic approximation. Such an equation exists: it is called the Dirac equation. After reviewing the Lorentz Group and Classical Relativistic Dynamics (Section I) we establish the Dirac equation (Section II), and make a detailed study of its invariance properties (Section III). In the remainder of this chapter we discuss the physical significance of the theory, and, in the course of reviewing its principal applications, examine how it is situated with respect to Classical Dynamics (Section IV), non-relativistic Quantum Mechanics (Section V) and Quantum Field Theory (Section VI).

The second chapter is devoted to the concept of the quantized field, and to the elementary Quantum theory of electromagnetic radiation and its interaction with atomic and nuclear systems.

2. Notation, Various Conventions and Definitions

UNITS. Except for a few obvious exceptions, all expressions appearing in what follows are written with

$$\hbar = c = 1.$$

With this particular choice of units, time appears to have the dimension of a length; energies, momenta and masses the dimension of an inverse length; electric charges appear as dimensionless quantities ($e^2 \equiv e^2/\hbar c \simeq 1/137$). The general expressions may be re-established by simple considerations of homogeneity.

COORDINATES. Specification of an instant t and a point $\boldsymbol{r} \equiv (x, y, z)$ of ordinary space defines a point of space-time. We denote the coordinates of this point by x^0, x^1, x^2, x^3; $x^0 \equiv ct$ is the time coordinate; x^1, x^2, x^3 the three spatial coordinates: $x^1 \equiv x$, $x^2 \equiv y$, $x^3 \equiv z$. More generally, we use the indices 0, 1, 2, 3 to denote the components of four-vectors or tensors along the axes t, x, y, z respectively. Greek indices denote the space-time components of four-vectors or tensors and therefore take the four values 0, 1, 2, 3; roman indices are reserved for the components of ordinary space and therefore take the three values 1, 2, 3. Thus:

$$x^\mu \equiv (x^0, x^k) \equiv (x^0, x^1, x^2, x^3)$$
$$(\mu = 0, 1, 2, 3) \qquad (k = 1, 2, 3).$$

METRIC TENSOR, COVARIANT AND CONTRAVARIANT INDICES

The space-time metric is a pseudo-euclidian metric defined by the *metric tensor*

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

or again

$$g_{00} = 1, \quad g_{kk} = -1, \quad g_{\mu\nu} = 0 \text{ if } \mu \neq \nu. \tag{XX.1}$$

We distinguish between *covariant* vectors (that transform like $\partial/\partial x^\mu$) and *contravariant* vectors (that transform like x^μ), and more generally between covariant tensor components and contravariant tensor components. Following the usual convention, covariant indices are placed as subscripts, contravariant indices as superscripts. Thus a^μ denotes a contravariant vector. The corresponding covariant vector a_μ is obtained by application of the metric tensor:

$$a_\mu = \sum g_{\mu\nu} a^\nu,$$

which gives

$$a_0 = a^0, \qquad a_k = -a^k.$$

We shall always follow *the convention of summing over repeated indices*. With this convention the preceding relation becomes simply

$$a_\mu = g_{\mu\nu} a^\nu.$$

Similarly, indices are raised by applying the tensor $g^{\mu\nu}$:

$$a^\mu = g^{\mu\nu} a_\nu.$$

In the present case, we also have

$$g^{\mu\nu} = g_{\mu\nu}.$$

In addition:

$$g_\mu{}^\nu = g_{\mu\varrho} g^{\varrho\nu} = g^\mu{}_\nu = \delta_\mu{}^\nu,$$

where $\delta_\mu{}^\nu$ is the Kronecker symbol:

$$\delta_\mu{}^\nu = \begin{cases} 1 & \text{if} \quad \mu = \nu \\ 0 & \text{if} \quad \mu \neq \nu. \end{cases}$$

THREE-VECTORS, FOUR-VECTORS, SCALAR PRODUCT

For three-vectors, or vectors of ordinary space, we retain the usual notation; each of them is denoted by a bold-face letter and its length by the corresponding character in ordinary print.

The three space components of a contravariant four-vector a^μ form a three-vector. With the above notations, we therefore have

$$a^\mu \equiv (a^0, a^1, a^2, a^3) \equiv (a^0, \boldsymbol{a}) \quad \boldsymbol{a} \equiv (a_x, a_y, a_z)$$

$$a^1 = a_x \quad a^2 = a_y \quad a^3 = a_z \quad a \equiv (\boldsymbol{a}\cdot\boldsymbol{a})^{\frac{1}{2}} \equiv [a_x{}^2 + a_y{}^2 + a_z{}^2]^{\frac{1}{2}}.$$

When no confusion is possible with the length of the three-vector $\boldsymbol{a}$, we shall sometimes omit the index and denote the four-vector a^μ simply by a.

The scalar product of two four-vectors a^μ and b^μ is obtained by contracting the contravariant components of the one with the covariant components of the other, i.e. it is given by either $a_\mu b^\mu$, or $a^\mu b_\mu$:

$$a_\mu b^\mu = a^\mu b_\mu = a^0 b^0 - \boldsymbol{a}\cdot\boldsymbol{b} \tag{XX.2}$$

The *norm* of a^μ is $a_\mu a^\mu = (a^0)^2 - a^2$.

CLASSIFICATIONS OF THE FOUR-VECTORS

The four-vectors may be put into three classes, according to the sign of their norm

$$\begin{aligned} a_\mu a^\mu &< 0 \qquad a^\mu = \text{space-like vector} \\ a_\mu a^\mu &= 0 \qquad a^\mu = \text{null vector} \\ a_\mu a^\mu &> 0 \qquad a^\mu = \text{time-like vector} \end{aligned}$$

This classification corresponds to the position of the vector with respect to the light cone $x_\mu x^\mu = 0$. The two latter cases can be further classified according to the sign of the time component:

$a^0 > 0$ the vector points towards the future;
$a^0 < 0$ the vector points towards the past.

GRADIENT, DIFFERENTIAL OPERATORS

We retain the notation $\nabla \equiv (\partial/\partial x, \partial/\partial y, \partial/\partial z)$ and $\Delta \equiv \nabla \cdot \nabla$.

The four partial-differentiation operators $\partial/\partial x^\mu$ form a *covariant* vector, called the gradient operator, which we denote by the symbol ∂_μ:

$$\begin{aligned} \partial_\mu \equiv \partial/\partial x^\mu &\equiv (\partial/\partial x^0, \partial/\partial x^1, \partial/\partial x^2, \partial/\partial x^3) \\ &\equiv (\partial/\partial ct, \nabla). \end{aligned} \tag{XX.3}$$

We shall also make use of the "contravariant gradient":

$$\partial^\mu \equiv g^{\mu\nu}\partial_\nu \equiv (\partial/\partial ct, -\nabla). \tag{XX.4}$$

The Dalembertian $\Box$ is defined [1]) by (cf. § II.12):

$$\Box \equiv \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \Delta \equiv \partial_\mu \partial^\mu. \tag{XX.5}$$

THE $\varepsilon^{\lambda\mu\nu\varrho}$ TENSOR

$\varepsilon^{\lambda\mu\nu\varrho}$ denotes the completely antisymmetrical tensor with four indices, the components of which are equal to 0 if two of the indices are equal, to $+1$ if $(\lambda\mu\nu\varrho)$ is an even permutation of (0, 1, 2, 3), and to -1 if $(\lambda\mu\nu\varrho)$ is an odd permutation of (0, 1, 2, 3).

ELECTROMAGNETIC FIELD

The electromagnetic potential is made up of a vector term $\boldsymbol{A}(\boldsymbol{r}, t)$ and a scalar potential $\varphi(\boldsymbol{r}, t)$ which form a four-vector A^μ:

$$A^\mu \equiv (\varphi, \boldsymbol{A}). \tag{XX.6}$$

---

[1]) Many authors use the symbol $\Box$ to denote the negative of the operator defined here.

The electric field $\mathscr{E}$ and the magnetic field $\mathscr{H}$ are given by

$$\mathscr{E} = -\nabla\varphi - \partial A/\partial x^0, \qquad \mathscr{H} = \operatorname{curl} \boldsymbol{A}. \tag{XX.7}$$

The components of $\mathscr{E}$ and $\mathscr{H}$ form an antisymmetrical space-time tensor, $F_{\mu\nu}$, according to the definition

$$F_{\mu\nu} = \frac{\partial A_\nu}{\partial x^\mu} - \frac{\partial A_\mu}{\partial x^\nu} \tag{XX.8}$$

giving

$$F_{\mu\nu} \equiv \begin{pmatrix} 0 & \mathscr{E}_x & \mathscr{E}_y & \mathscr{E}_z \\ -\mathscr{E}_x & 0 & -\mathscr{H}_z & \mathscr{H}_y \\ -\mathscr{E}_y & \mathscr{H}_z & 0 & -\mathscr{H}_x \\ -\mathscr{E}_z & -\mathscr{H}_y & \mathscr{H}_x & 0 \end{pmatrix} \tag{XX.9}$$

We shall also use the four-vector operator D_μ defined by

$$D_\mu \equiv \partial_\mu + \mathrm{i}eA_\mu \equiv \left(\frac{\partial}{\partial x^0} + \mathrm{i}e\varphi,\ \nabla - \mathrm{i}e\boldsymbol{A}\right). \tag{XX.10}$$

3. The Lorentz Group

A Lorentz change of referential is a *real, linear* transformation of the coordinates *conserving the norm* of the intervals between the different points of space-time. In such a transformation, the new coordinates x'^μ of a space-time point are obtained from the old ones x^μ by the relation

$$x'^\mu = \Omega^\mu{}_\nu x^\nu + a^\mu.$$

The real vector a^μ represents a simple translation of the space-time axes. In what follows, we shall treat the translations separately and give the name of Lorentz transformation to the homogeneous transformations $(a^\mu = 0)$ [1]:

$$x'^\mu = \Omega^\mu{}_\nu\, x^\nu. \tag{XX.11}$$

By raising or lowering the indices, we can obtain the matrices $\Omega_\mu{}^\nu$, $\Omega^{\mu\nu}$, $\Omega_{\mu\nu}$ from the matrix $\Omega^\mu{}_\nu$ (for example: $\Omega^{\mu\nu} = g^{\nu\varrho}\Omega^\mu{}_\varrho$). Specification of any one of these matrices defines the Lorentz trans-

[1] The group formed by all of the Lorentz transformations including the translations is commonly called the *inhomogeneous Lorentz group*, or *Poincaré group*.

formation in question. The condition of reality and of invariance of the norm are written

$$\Omega_{\mu\nu}{}^{*} = \Omega_{\mu\nu} \tag{XX.12}$$

$$\Omega_{\mu\nu}\Omega^{\mu\lambda} = \Omega_{\nu\mu}\Omega^{\lambda\mu} = \delta_{\nu}{}^{\lambda}. \tag{XX.13}$$

It follows that

$$\det |\Omega^{\mu}{}_{\nu}| = \pm 1 \tag{XX.14}$$

and the inverse transformation is written

$$x^{\mu} = x'^{\nu}\Omega_{\nu}{}^{\mu}. \tag{XX.15}$$

These transformations form a group, *the complete Lorentz group*. It is the group of real linear transformations conserving scalar products between four-vectors.

If $\Omega^{00} > 0$, the transformation conserves the sense of time-like vectors, that is, it conserves the sign of the time component of these vectors; it is then called orthochronous, and the set of these particular Lorentz transformations is called *the orthochronous Lorentz group*.

If *in addition* $\det |\Omega^{\mu}{}_{\nu}| = +1$, the transformation also conserves the sense of Cartesian systems in ordinary space, it is then called a proper Lorentz transformation. The ensemble of these transformation forms a group, *the proper Lorentz group*, which we denote by $\mathscr{L}_0$.

All transformations of the proper group may be considered as a succession of infinitesimal transformations. The matrix $\Omega_{\mu\nu}$ of an *infinitesimal Lorentz transformation* is of the form

$$g_{\mu\nu} + \omega_{\mu\nu},$$

where the quantities $\omega_{\mu\nu}$ are infinitesimals. Conditions (XX.12) and (XX.13) give

$$\omega_{\mu\nu} = \omega_{\mu\nu}{}^{*}, \qquad \omega_{\mu\nu} + \omega_{\nu\mu} = 0. \tag{XX.16}$$

$\omega_{\mu\nu}$ is therefore a real antisymmetrical tensor.

Put

$$Z_{\mu\nu}^{(\alpha\beta)} = -Z_{\mu\nu}^{(\beta\alpha)} = g_{\mu\alpha}g_{\nu\beta} - g_{\mu\beta}g_{\nu\alpha}. \tag{XX.17}$$

$Z_{\mu\nu}^{(\alpha\beta)}$ is an antisymmetrical tensor whose only non-vanishing elements are the two elements $\mu = \alpha, \nu = \beta$ and $\mu = \beta, \nu = \alpha$; one of which is equal to $+1$, the other to -1. ε being an infinitesimal quantity,

$$g_{\mu\nu} - \varepsilon Z_{\mu\nu}^{(\alpha\beta)}$$

is the matrix of a particular infinitesimal Lorentz transformation, the "rotation" through an angle ε in the $x^\alpha x^\beta$ plane.

There exists in all six infinitesimal transformations of this type. The "rotations" in the planes x^1x^2, x^2x^3 and x^3x^1 are spatial rotations of angle ε about the axes Oz, Ox, Oy respectively, the "rotations" in the planes x^1x^0, x^2x^0, x^3x^0 are special Lorentz transformations of velocity ε in the directions Ox, Oy, Oz respectively [1]).

In addition to infinitesimal transformations, one can define different types of *reflection*, notably the spatial reflection s ($x^0 = x^0$, $x^k = -x^k$) and the time reflection t ($x^0 = -x^0$, $x^k = x^k$). The orthochronous group is made up of the transformations $\mathscr{L}_0$, of the reflection s and of their products $s\mathscr{L}_0$. The complete group is formed of the transformations $\mathscr{L}_0$, $s\mathscr{L}_0$, $t\mathscr{L}_0$, and $st\mathscr{L}_0$. The properties of these four sheets of the complete group are summed up in the following table:

Sheet	Det $\lvert \Omega^\mu_\nu \rvert$	Ω^{00}	group
$\mathscr{L}_0$	$+1$	>0	proper, orthochronous, complete
$s\mathscr{L}_0$	-1	>0	orthochronous, complete
$t\mathscr{L}_0$	-1	<0	complete
$st\mathscr{L}_0$	$+1$	<0	complete

---

[1]) If the new referential is obtained from the old one by a rotation through a finite angle φ about Oz, one has:

$$x'^1 = x^1 \cos\varphi + x^2 \sin\varphi, \quad x'^2 = x^2 \cos\varphi - x^1 \sin\varphi, \quad x'^3 = x^3, \quad x'^0 = x^0.$$

If it is obtained by a special Lorentz transformation of velocity $v = \tanh\varphi$ directed along Ox, one has

$$x'^1 = x^1 \cosh\varphi - x^0 \sinh\varphi, \quad x'^0 = x^0 \cosh\varphi - x^1 \sinh\varphi, \quad x'^2 = x^2, \quad x'^3 = x^3.$$

The transformations considered above correspond to the case when $\varphi = \varepsilon \equiv$ infinitesimal.

4. Classical Relativistic Dynamics

Let us recall the dynamical properties of a classical particle of rest mass m and charge e in an electromagnetic field $(\varphi, \mathbf{A})$.

Let $\mathbf{v}$ be the velocity of the particle:

$$\mathbf{v} \equiv \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}. \tag{XX.18}$$

We define the relativistic mass M and the mechanical momentum [1] $\boldsymbol{\pi}$ by:

$$M \equiv \frac{m}{\sqrt{1-v^2}}, \qquad \boldsymbol{\pi} \equiv M\mathbf{v} \tag{XX.19}$$

$(M, \boldsymbol{\pi})$ is a certain four-vector π^μ of norm m^2:

$$M^2 - \boldsymbol{\pi}^2 = m^2 \tag{XX.20}$$

and pointing into the future $(M > 0)$.

In the absence of a field, the particle follows a uniform rectilinear motion: $\mathbf{v} = Cst.$

In the presence of an electromagnetic field, the trajectory followed by the particle satisfies the equation

$$\frac{\mathrm{d}\boldsymbol{\pi}}{\mathrm{d}t} = e[\mathscr{E} + \mathbf{v} \times \mathscr{H}] \equiv \mathbf{F}. \tag{XX.21}$$

This is the fundamental equation of the relativistic dynamics of a material point. The vector $\mathbf{F}$ is called the *Lorentz force*.

From (XX.21) we have the equations

$$\frac{\mathrm{d}M}{\mathrm{d}t} = (\mathbf{v} \cdot \mathbf{F}) = e(\mathbf{v} \cdot \mathscr{E}), \tag{XX.21'}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{r} \times \boldsymbol{\pi}) = \mathbf{r} \times \mathbf{F} \tag{XX.22}$$

giving respectively the law of motion of the mass and of the moment of the mechanical momentum.

These relations can be put in covariant form by introducing the proper time τ of the particle, in accordance with the definition

$$\begin{aligned} \mathrm{d}\tau &= (\mathrm{d}x^\mu \, \mathrm{d}x_\mu)^{\frac{1}{2}} \\ &= \sqrt{1-v^2}\, \mathrm{d}t. \end{aligned}$$

[1] Not to be confused with the momentum which in this book means the Lagrange canonical conjugate of the coordinates (cf. note, p. 54, vol. I).

One defines the four-velocity

$$u^\mu \equiv \mathrm{d}x^\mu/\mathrm{d}\tau \equiv (\mathrm{d}t/\mathrm{d}\tau, \boldsymbol{v}\,\mathrm{d}t/\mathrm{d}\tau) \qquad (u^\mu u_\mu = 1)$$

whose product with m gives the mechanical four-momentum

$$\pi^\mu \equiv m u^\mu \equiv (M, \boldsymbol{\pi}).$$

Equations (XX.21) and (XX.21′) are equivalent to the formally covariant equation

$$\frac{\mathrm{d}\pi^\mu}{\mathrm{d}\tau} = eF^{\mu\nu}u_\nu, \tag{XX.23}$$

or

$$\frac{\mathrm{d}u^\mu}{\mathrm{d}\tau} = \frac{e}{m}F^{\mu\nu}u_\nu$$

$F^{\mu\nu}$ is the electromagnetic tensor [eqs. (XX.8–9)].

These laws of motion can be deduced from a Lagrangian or Hamiltonian formalism (cf. Problem I.5). The momentum $\boldsymbol{p}$ and the energy E form a four-vector p^μ, related to the four-vector π^μ by the relation

$$p^\mu = \pi^\mu + eA^\mu \tag{XX.24}$$

i.e.

$$E = M + e\varphi, \qquad \boldsymbol{p} = \boldsymbol{\pi} + e\boldsymbol{A}.$$

The Hamiltonian function is defined by

$$H \equiv e\varphi + \sqrt{(\boldsymbol{p} - e\boldsymbol{A})^2 + m^2} \tag{XX.25}$$

in accordance with relations (XX.24) and (XX.20). From (XX.25) we obtain Hamilton's canonical equations

$$\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} = \frac{\boldsymbol{\pi}}{M}, \qquad \frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t} = -e\,\mathrm{grad}\,(\varphi - \boldsymbol{v}\cdot\boldsymbol{A}).$$

The first of these is the definition of velocity. The second is equivalent to (XX.21) as may easily be verified using the definitions of $\mathscr{E}$ and $\mathscr{H}$ [eq. (XX.7)] and the fact that

$$\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t} = \left(\frac{\partial}{\partial t} + \boldsymbol{v}\cdot\mathrm{grad}\right)\boldsymbol{A}.$$

II. THE DIRAC AND KLEIN–GORDON EQUATIONS

5. The Klein–Gordon Equation

Since the problem of finding a relativistic wave equation for the electron is complicated by the existence of spin, we first look for a

relativistic wave equation for a particle of spin 0, a π meson for example. Since such a particle has no internal degrees of freedom, its wave function Ψ depends only on the variables $\boldsymbol{r}$ and t. Let m be its mass and e its charge, and suppose that it is moving in the electromagnetic potential $A^\mu \equiv (\varphi, \boldsymbol{A})$.

To find the wave equation we proceed empirically using the correspondence principle; this will guarantee that we obtain the classical laws of motion when the classical approximation is valid.

We recall that the Schrödinger correspondence rule is given by

$$E \to \mathrm{i}\frac{\partial}{\partial t}, \qquad \boldsymbol{p} \to -\mathrm{i}\nabla. \tag{XX.26}$$

Putting $p^\mu \equiv (E, \boldsymbol{p})$, this rule can be written more simply:

$$p^\mu \to \mathrm{i}\partial^\mu. \tag{XX.26'}$$

From expression (XX.25) for the Hamiltonian we obtain

$$E = e\varphi + \sqrt{(\boldsymbol{p} - e\boldsymbol{A})^2 + m^2} \tag{XX.27}$$

from which we obtain, by rule (XX.26), the wave equation

$$\left(\mathrm{i}\frac{\partial}{\partial t} - e\varphi\right)\Psi = \left[\left(\frac{1}{\mathrm{i}}\nabla - e\boldsymbol{A}\right)^2 + m^2\right]^{\frac{1}{2}}\Psi.$$

This equation has two serious drawbacks. First, the dissymmetry between the space and time coordinates is such that relativistic invariance and its consequences are not clearly exhibited. Second, the operator on the right-hand side is a square root, which is practically untractable except when the field $\boldsymbol{A}$ vanishes.

One avoids these two difficulties by taking relation (XX.20) as the starting point of the correspondence operation, giving

$$(E - e\varphi)^2 - (\boldsymbol{p} - e\boldsymbol{A})^2 = m^2. \tag{XX.28}$$

This relation is not equivalent to (XX.27), but to the more general relation

$$E = e\varphi \pm \sqrt{(\boldsymbol{p} - e\boldsymbol{A})^2 + m^2}. \tag{XX.29}$$

Only the + sign corresponds to real classical solutions; the − sign represents solutions of negative mass without any physical significance. Thus by taking (XX.28) as a starting point we introduce *parasitic solutions of negative mass.*

The correspondence operation applied to (XX.28) gives the *Klein–Gordon equation*:

$$\left[\left(\mathrm{i}\frac{\partial}{\partial t} - e\varphi\right)^2 - \left(\frac{1}{\mathrm{i}}\nabla - e\boldsymbol{A}\right)^2\right]\Psi = m^2\Psi \qquad \text{(XX.30)}$$

which can also be written in the form

$$(D_\mu D^\mu + m^2)\,\Psi \equiv [(\partial_\mu + \mathrm{i}eA_\mu)(\partial^\mu + \mathrm{i}eA^\mu) + m^2]\,\Psi = 0, \quad \text{(XX.30')}$$

where its relativistic invariance is evident.

Let us briefly consider the interpretation of this equation [1]). To simplify the discussion we limit ourselves to the case when the field vanishes. We then have simply (cf. § II.12):

$$(\Box + m^2)\,\Psi = 0. \qquad \text{(XX.31)}$$

This is a second-order differential equation with respect to the time, and we must therefore know both Ψ and $\partial\Psi/\partial t$ at the initial time for Ψ to be completely determined at any later time. This difficulty is easily surmounted if we postulate that the dynamical state of the system at a given time is represented not by the single function Ψ but by the set of two functions Ψ and $\partial\Psi/\partial t$ or by the two linear combinations:

$$\Phi = \Psi + \frac{\mathrm{i}}{m}\frac{\partial\Psi}{\partial t}, \qquad \chi = \Psi - \frac{\mathrm{i}}{m}\frac{\partial\Psi}{\partial t}.$$

This is equivalent to postulating that the state of the system is represented by a wave function with two components, Φ and χ. This wave function must obey a differential equation of the first order with respect to time which is easily deduced from the Klein–Gordon equation. In the non-relativistic limit, the energy of the particle is nearly equal to its rest mass m, so that

$$\mathrm{i}\frac{\partial\Psi}{\partial t} \simeq m\Psi$$

and therefore, $\chi \ll \Phi$. One of the two components becoming negligible beside the other, we obtain the non-relativistic Schrödinger theory in which the dynamical state of a particle of spin 0 is represented by a one-component wave function.

[1]) For a more complete account, see H. Feshbach and F. Villars, Rev. Mod. Phys. 30 (1958) 24, where a list of the main references will also be found.

In order to interpret the wave function, we must define a position probability density P and a current probability density $\boldsymbol{j}$ satisfying the equation of continuity (cf. § IV.4):

$$\frac{\partial P}{\partial t} + \nabla \cdot \boldsymbol{j} = 0 \tag{XX.32}$$

or with the notation $j^\mu \equiv (P, \boldsymbol{j})$

$$\partial_\mu j^\mu = 0. \tag{XX.33}$$

Since Ψ and Ψ^* both satisfy equation (XX.31),

$$\Psi^* (\Box \Psi) - (\Box \Psi^*) \Psi = 0$$

which gives, using the definition of the Dalembertian,

$$\partial_\mu [\Psi^* (\partial^\mu \Psi) - (\partial^\mu \Psi^*) \Psi] = 0.$$

The equation of continuity is satisfied if we take j^μ proportional to the bracket on the left-hand side. The proportionality constant is fixed so as to recover the usual definition in the non-relativistic limit:

$$j^\mu = \frac{\mathrm{i}}{2m} [\Psi^* (\partial^\mu \Psi) - (\partial^\mu \Psi^*) \Psi],$$

i.e.

$$\begin{aligned} P(\boldsymbol{r}, t) &= \frac{\mathrm{i}}{2m} \left[\Psi^* \frac{\partial \Psi}{\partial t} - \frac{\partial \Psi^*}{\partial t} \Psi \right] \\ \boldsymbol{j}(\boldsymbol{r}, t) &= \frac{1}{2\mathrm{i}m} [\Psi^* (\nabla \Psi) - (\nabla \Psi^*) \Psi]. \end{aligned} \tag{XX.34}$$

Examining (XX.34) we see that *the density $P(\boldsymbol{r}, t)$ is not positive-definite.* Here we have one of the major difficulties with the Klein–Gordon equation.

Another difficulty, related to the preceding one, is due to the possibility of "negative energy solutions". If, for example, we look for the plane wave solutions of the equation in the absence of a field,

$$\Psi = \exp [-\mathrm{i}(Et - \boldsymbol{p} \cdot \boldsymbol{r})]$$

we obtain, substituting this expression in (XX.31), the condition:

$$E = \pm \sqrt{\boldsymbol{p}^2 + m^2}.$$

There therefore exist solutions of negative energy $-\sqrt{\boldsymbol{p}^2 + m^2}$. Their presence is obviously due to the above-mentioned introduction of

negative masses into the theory (it would be more correct to call them negative mass solutions; however when the field is null, the distinction between mass and energy is illusory).

Following Pauli and Weisskopf [1]), we deal with these difficulties by modifying the interpretation of the four-vector j^μ and the definition of average values. According to this reinterpretation of the theory, ej^μ denotes the current-density four-vector; in particular $eP(\boldsymbol{r}, t)$ is the electric charge density. Equation (XX.33) is therefore an equation for the conservation of the charge. On the other hand, the number of particles is not conserved; this is explained by the possibility of annihilation and creation of pairs of particles of opposite charge, phenomena that only Field Theory accounts for in a satisfactory manner. When formulated in this way, *the theory is therefore a one-charge theory and not a one-particle theory.* In the Dirac theory, on the contrary, one is able to define a positive-definite density P, however, we shall see that the difficulty with negative energies remains, and that the Dirac theory cannot be considered in an entirely satisfying fashion as being a one-particle theory either (Section VI).

6. The Dirac Equation

Let us now attempt to form a relativistic wave equation for the electron. Following Dirac, we proceed by analogy with non-relativistic Quantum Mechanics.

Just as the electron of the non-relativistic theory is represented by a two-component spinor which transforms under rotation like an angular momentum of value $\frac{1}{2}$, the electron of the relativistic theory must be represented by a wave function of several components having a certain well-defined variance with respect to the larger group of Lorentz transformations. We denote the s component of the wave Ψ by $\psi_s(\boldsymbol{r}, t)$. Ψ may be written in the form of a column matrix:

$$\Psi = \begin{vmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_N \end{vmatrix}$$

As in the non-relativistic case, one may equally well regard the wave Ψ at a given time as a function of the orbital variables $\boldsymbol{r}$ and

[1]) W. Pauli and V. Weisskopf, Helv. Phys. Acta 7 (1934) 709. See also H. Feshback and F. Villars *loc. cit.* note p. 886.

the intrinsic, or spin, variables s ($s = 1, 2, \ldots, N$). Such a wave represents a certain state vector $|\psi(t)\rangle$; and the space of these states, $\mathscr{E}$, is the tensor product

$$\mathscr{E} = \mathscr{E}^{(0)} \otimes \mathscr{E}^{(s)}$$

of the orbital-variable space $\mathscr{E}^{(0)}$ by the spin-variable space $\mathscr{E}^{(s)}$; the wave Ψ represents this vector in a suitable representation:

$$\Psi(\mathbf{r}, s; t) \equiv \psi_s(\mathbf{r}, t) \equiv \langle \mathbf{r}s | \Psi(t) \rangle.$$

Continuing with the analogy, we define the position probability density by the formula

$$P(\mathbf{r}, t) = \sum_{s=1}^{N} |\psi_s|^2. \tag{XX.35}$$

With these hypotheses, the wave equation is necessarily of the form

$$\mathrm{i} \frac{\partial \Psi}{\partial t} = H_D \Psi, \tag{XX.36}$$

where H_D is a Hermitean operator of state-vector space. To see this, we note, on the one hand that since Ψ completely defines the dynamical state of the electron at each instant the wave equation must be of the first order with respect to time; on the other that H_D must be Hermitean in order to guarantee the self-consistency of our definition of $P(\mathbf{r}, t)$ (cf. § IV.3).

Since we are seeking a relativistic wave equation, we also require that it exhibit a certain formal symmetry between the spatial coordinates and the time namely that it also be of first order with respect to the spatial variables.

Let us first consider the case of an electron in the absence of field. The Hamiltonian must then be invariant under translation, thus independent of $\mathbf{r}$. Taking into account all the preceding hypotheses, it can therefore be written in the form

$$H_D = \boldsymbol{\alpha} \cdot \mathbf{p} + \beta m, \tag{XX.37}$$

where the operator $\mathbf{p}$ has the significance indicated by the correspondence rule (XX.26), i.e. $\mathbf{p} = -\mathrm{i}\nabla$, and where $\boldsymbol{\alpha} \equiv (\alpha_x, \alpha_y, \alpha_z)$ and β denote 4 Hermitean operators acting on the spin variables alone. If we adopt the notation $E \equiv \mathrm{i}\partial/\partial t$ the wave equation reads

$$[E - \boldsymbol{\alpha} \cdot \mathbf{p} - \beta m] \Psi = 0. \tag{XX.38}$$

To determine $\boldsymbol{\alpha}$ and β, we invoke the correspondence principle: we require the solutions of this equation to satisfy the Klein–Gordon equation, i.e. we require that

$$[E^2 - \boldsymbol{p}^2 - m^2]\Psi = 0. \tag{XX.39}$$

Multiplying on the left by the operator $[E + \boldsymbol{\alpha}\cdot\boldsymbol{p} + \beta m]$, eq. (XX.38) gives the second-order equation

$$[E^2 - \sum_k (\alpha^k)^2 (p^k)^2 - \beta^2 m^2 - \sum_{k<l} (\alpha^k \alpha^l + \alpha^l \alpha^k) p^k p^l - \sum_k (\alpha^k \beta + \beta \alpha^k) m p^k] \Psi = 0.$$

This equation and eq. (XX.39) are identical if the 4 operators β, $\boldsymbol{\alpha}$ anticommute and if their squares are equal to 1:

$$\begin{aligned} (\alpha^k)^2 = 1 \qquad & \alpha^k \alpha^l + \alpha^l \alpha^k = 0 \qquad (k \neq l) \\ \beta^2 = 1 \qquad & \alpha^k \beta + \beta\, \alpha^k = 0 \end{aligned} \tag{XX.40}$$

Equation (XX.38), in which the matrices β, $\boldsymbol{\alpha}$ are chosen to be Hermitean and to satisfy relation (XX.40), is called the Dirac equation.

From this equation for the free electron, we pass to the Dirac equation for an electron in the electromagnetic field $(\varphi, \boldsymbol{A})$ by making the substitution

$$E \to E - e\varphi, \qquad \boldsymbol{p} \to \boldsymbol{p} - e\boldsymbol{A} \tag{XX.41}$$

[e is the charge of the electron $(e<0)$]. One obtains:

$$[(E - e\varphi) - \boldsymbol{\alpha}\cdot(\boldsymbol{p} - e\boldsymbol{A}) - \beta m]\, \Psi = 0 \tag{XX.42}$$

i.e.

$$\left[\left(\mathrm{i}\frac{\partial}{\partial t} - e\varphi\right) - \boldsymbol{\alpha}\cdot(-\mathrm{i}\nabla - e\boldsymbol{A}) - \beta m\right] \Psi = 0. \tag{XX.43}$$

Comparing this with equation (XX.36) we find the following expression for the Dirac Hamiltonian in the presence of an external field:

$$H_D = e\varphi + \boldsymbol{\alpha}\cdot(\boldsymbol{p} - e\boldsymbol{A}) + \beta m. \tag{XX.44}$$

7. Construction of the Space $\mathscr{E}^{(s)}$. Dirac Representation

It remains to construct $\mathscr{E}^{(s)}$. The operators of this space are the 4 basic operators β, α_x, α_y, α_z and all of the functions that can be formed with them. $\mathscr{E}^{(s)}$ must be irreducible with respect to this set of operators.

To construct $\mathscr{E}^{(s)}$ we shall make use of the Hermitean character of the four basic operators and of relations (XX.40) defining their algebraic properties.

These properties are analogous to those of the three operators $\sigma_1, \sigma_2, \sigma_3$ of the non-relativistic spin $\frac{1}{2}$ theory. In this case, the spin-variable space $\mathscr{E}^{(\sigma)}$ has two dimensions. It can be constructed in the following way. Since σ_3 is Hermitean and $\sigma_3^2 = 1$ its only possible eigenvalues are ± 1. Moreover, to each eigenvector of σ_3 one can associate another eigenvector corresponding to the opposite eigenvalue. Consider for example, a vector $|+\rangle$ such that: $\sigma_3|+\rangle = |+\rangle$. Since σ_3 and σ_1 anticommute, the vector $|-\rangle \equiv \sigma_1|+\rangle$ has the property $\sigma_3|-\rangle = (-1)|-\rangle$. One has $\sigma_1|\pm\rangle = |\mp\rangle$ and $\sigma_3|\pm\rangle = (\pm 1)|\pm\rangle$. The space spanned by the vectors $|+\rangle$ and $|-\rangle$ is therefore invariant with respect to the operators σ_3 and σ_1 and with respect to functions of these operators (notably $\sigma_2 \equiv i\sigma_1\sigma_3$). From the very fashion in which it was constructed, it is irreducible. It is therefore the sought-for space $\mathscr{E}^{(\sigma)}$. In the representation with basis vectors $|+\rangle$ and $|-\rangle$ the operators σ_1, σ_2 and σ_3 are represented by the Pauli matrices [cf. § XIII.19 or formula (VII.65)].

To construct $\mathscr{E}^{(s)}$ we reduce the problem to the preceding one. We introduce the operators σ_x, σ_y, σ_z and ϱ_1, ϱ_2, ϱ_3 defined by:

$$\sigma_z = -i\alpha_x\alpha_y, \quad \sigma_x = -i\alpha_y\alpha_z, \quad \sigma_y = -i\alpha_z\alpha_x \tag{XX.45}$$

$$\varrho_3 = \beta, \quad \varrho_1 = \sigma_z\alpha_z = -i\alpha_x\alpha_y\alpha_z, \quad \varrho_2 = i\varrho_1\varrho_3 = -\beta\alpha_x\alpha_y\alpha_z; \tag{XX.46}$$

the four basic operators are expressed in terms of the ϱ and the σ by the formulas

$$\beta = \varrho_3, \qquad \alpha^k = \varrho_1\sigma^k. \tag{XX.47}$$

The construction of $\mathscr{E}^{(s)}$ therefore consists in the construction of a space irreducible with respect to the ϱ and the σ. Now it is easy to see that:

(*i*) each ϱ commutes with each σ;

(*ii*) the σ are three anticommuting Hermitean operators whose square is unity;

(*iii*) the ϱ are three anticommuting Hermitean operators whose square is unity.

Consequently (cf. § VIII.7):

(*i*) $\mathscr{E}^{(s)}$ is the tensor product

$$\mathscr{E}^{(s)} = \mathscr{E}^{(\varrho)} \otimes \mathscr{E}^{(\sigma)}$$

of a space $\mathscr{E}^{(\varrho)}$ irreducible with respect to the ϱ and a space $\mathscr{E}^{(\sigma)}$ irreducible with respect to the σ;

(*ii*) $\mathscr{E}^{(\sigma)}$ is a 2-dimensional space that can be constructed following the method indicated above;

(*iii*) $\mathscr{E}^{(\varrho)}$ is also a 2-dimensional space that can be constructed by the same method.

The space $\mathscr{E}^{(s)}$ therefore has 4 dimensions.

In the following sections we show that the σ are related to the spin, and the ϱ to the sign of the energy, the Dirac equation having, like the Klein–Gordon – and for the same reasons – negative energy solutions. In particular we shall see that $\boldsymbol{\alpha}$ is a (polar) vector operator and that $\boldsymbol{\sigma} \equiv (\sigma_x, \sigma_y, \sigma_z)$ is an (axial) vector operator; in addition one formally has

$$\boldsymbol{\alpha} \times \boldsymbol{\alpha} = 2\mathrm{i}\boldsymbol{\sigma}. \tag{XX.48}$$

The spin of the electron is $\frac{1}{2}\boldsymbol{\sigma}$. The sign of the energy is roughly given by the eigenvalue of $\beta \equiv \varrho_3$.

The dynamical state of the electron is therefore represented by a function Ψ having 4 components, i.e. twice as many components as the wave of the non-relativistic spin $\frac{1}{2}$ theory. The representation in which the ϱ and the σ are the Pauli matrices [cf. eq. (VII.65–66)] is called the *Dirac representation*; in this representation, each component corresponds to a given orientation of the spin along the axis Oz, and roughly to a given sign of the energy.

8. Covariant Form of the Dirac Equation

Equation (XX.43) is the Dirac equation as originally proposed by Dirac himself. It is in this form that it lends itself most easily to physical interpretation and to the study of the passage to the non-relativistic limit. We now propose to obtain a second form, more symmetrical with respect to the space and time coordinates and

therefore more convenient whenever questions of relativistic covariance play a preponderant role.

To this effect, we multiply both sides of (XX.43) on the left by β and put:

$$\gamma^\mu \equiv (\gamma^0, \gamma^1, \gamma^2, \gamma^3) \equiv (\gamma^0, \boldsymbol{\gamma}) \qquad \gamma^0 \equiv \beta \qquad \boldsymbol{\gamma} \equiv \beta\boldsymbol{\alpha}. \tag{XX.49}$$

We obtain

$$\boxed{[\mathrm{i}\gamma^\mu D_\mu - m]\,\Psi \equiv [\gamma^\mu(\mathrm{i}\partial_\mu - eA_\mu) - m]\,\Psi = 0} \tag{XX.50}$$

The properties of γ^μ are easily obtained from those of $\boldsymbol{\alpha}$ and β by applying definitions (XX.49). The ten relations (XX.40) give the ten equivalent relations

$$\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}. \tag{XX.51}$$

The Hermitean conditions on the $\boldsymbol{\alpha}$ and β are equivalent to the conditions

$$\gamma^{0\dagger} = \gamma^0, \qquad \gamma^{k\dagger} = -\gamma^k \tag{XX.52}$$

that may be written in the condensed form

$$\gamma^{\mu\dagger} = \gamma^0\gamma^\mu\gamma^0. \tag{XX.53}$$

It is convenient to extend the usual rule of raising and lowering of indices to the γ, and to put

$$\gamma_\mu = g_{\mu\nu}\gamma^\nu. \tag{XX.54}$$

Note that:

$$\gamma_0 = \gamma^0, \qquad \gamma_k = -\gamma^k \tag{XX.55}$$

$$\gamma^\mu = \gamma_\mu{}^\dagger = \gamma_\mu{}^{-1}. \tag{XX.56}$$

9. Adjoint Equation. Definition of the Current

We have defined a positive-definite position probability density [eq. (XX.35)]. As indicated above, the Hermitean character of the Dirac Hamiltonian guarantees that this definition is self-consistent. We shall now define a current density and show that the current defined with the solutions of the Dirac equation obeys an equation of continuity. We first give a complete discussion of this problem using the Dirac form, and then repeat the argument with the covariant form.

Suppose that we have chosen a particular representation of the β and $\boldsymbol{\alpha}$. The wave Ψ is then a certain column matrix with four lines

$$\Psi \equiv \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix}.$$

Denote its Hermitean conjugate by:

$$\Psi^\dagger \equiv (\psi_1^* \psi_2^* \psi_3^* \psi_4^*).$$

The operators of spin space are 4×4 matrices. One can define partial scalar products in which one sums over the spin variables alone. We shall indicate such scalar products by parentheses; thus the density P is written

$$P(\mathbf{r}, t) \equiv (\Psi^\dagger \Psi). \tag{XX.57}$$

As another illustration, denote by β_{st} the element in line s and column t of the matrix β $(s, t = 1, 2, 3, 4)$:

$$(\Psi^\dagger \beta \Psi) \equiv \sum_s \sum_t \psi_s^* \beta_{st} \psi_t.$$

Now if Ψ is a solution of the Dirac equation, that is, if

$$\begin{aligned} \mathrm{i} \frac{\partial \Psi}{\partial t} &= H_D \Psi \\ &= \left[e\varphi + \sum_k \alpha^k \left(-\mathrm{i} \frac{\partial}{\partial x^k} - eA^k \right) + \beta m \right] \Psi, \end{aligned} \tag{XX.58}$$

then $\Psi^\dagger$ is a solution of the Hermitean conjugate equation, that is of the equation obtained by taking the complex conjugate of (XX.58) and replacing each matrix by its transpose:

$$\begin{aligned} \mathrm{i} \frac{\partial \Psi^\dagger}{\partial t} &= -\Psi^\dagger H_D \\ &= -e\varphi \Psi^\dagger - \sum_k \left(\mathrm{i} \frac{\partial}{\partial x^k} - eA^k \right) \Psi^\dagger \alpha^k - m \Psi^\dagger \beta. \end{aligned} \tag{XX.59}$$

Scalar multiplication of (XX.58) on the left by $\Psi^\dagger$ and (XX.59) on the right by Ψ and then addition gives

$$\mathrm{i} \frac{\partial}{\partial t} (\Psi^\dagger \Psi) = -\mathrm{i} \sum_k \frac{\partial}{\partial x^k} (\Psi^\dagger \alpha^k \Psi). \tag{XX.60}$$

On the left-hand side we recognize the time derivative of the probability density, and on the right-hand side the divergence of a certain vector $\boldsymbol{j}(\boldsymbol{r}, t)$ defined by

$$\boldsymbol{j}(\boldsymbol{r}, t) \equiv (\Psi^\dagger \boldsymbol{\alpha} \Psi). \tag{XX.61}$$

$\boldsymbol{j}(\boldsymbol{r}, t)$ is the sought-for current density and (XX.60) is equivalent to the equation of continuity

$$\frac{\partial}{\partial t} P + \nabla \cdot \boldsymbol{j} = 0.$$

The above may be repeated starting from the covariant form of the Dirac equation [eq. (XX.50)]. The Hermitean conjugate of (XX.50) is

$$(-\mathrm{i}\partial_\mu - eA_\mu)\, \Psi^\dagger \gamma^{\mu\dagger} - m\, \Psi^\dagger = 0. \tag{XX.62}$$

[Here the symbol $\partial_\mu \Psi^\dagger \gamma^{\mu\dagger}$ represents the line matrix with four columns $(\partial \Psi^\dagger / \partial x^\mu)\gamma^{\mu\dagger}$.] It is convenient to put:

$$\overline{\Psi} = \Psi^\dagger \gamma^0, \qquad \Psi^\dagger = \overline{\Psi} \gamma^0. \tag{XX.63}$$

Taking relations (XX.53) into account, equation (XX.62) can then be put after multiplication on the left by γ^0 in the simpler form:

$$\boxed{(-\mathrm{i}\partial_\mu - eA_\mu)\, \overline{\Psi} \gamma^\mu - m \overline{\Psi} = 0} \tag{XX.64}$$

This equation is obviously equivalent to (XX.59). $\overline{\Psi}$ is called the *adjoint of* Ψ, and (XX.64) the *adjoint equation.*

Scalar multiplication of (XX.50) on the left by $\overline{\Psi}$ and (XX.64) on the right by Ψ and then subtraction, gives

$$\mathrm{i}\partial_\mu (\overline{\Psi} \gamma^\mu \Psi) = 0.$$

One defines the current density four-vector

$$j^\mu \equiv (\overline{\Psi} \gamma^\mu \Psi). \tag{XX.65}$$

The preceding equation is equivalent to the equation of continuity

$$\partial_\mu j^\mu = 0.$$

One easily verifies that $j^\mu \equiv (P, \boldsymbol{j})$. Thus, as expected, we obtain the equation of continuity written in its covariant form. In the next section, we show that the 4 components of j^μ indeed form a four-vector.

III. INVARIANCE PROPERTIES OF THE DIRAC EQUATION

10. Properties of the Dirac Matrices

As a preliminary to the study of the invariance properties of the Dirac equation we now make a systematic study of the properties of four 4×4 *matrices* $\gamma^\mu \equiv (\gamma^0, \gamma^1, \gamma^2, \gamma^3)$ satisfying the relations

$$\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}I, \tag{XX.66}$$

where I is the unit matrix. The matrix relations (XX.66) are analogous to the operator relations (XX.51); however, the matrices introduced here do not necessarily verify the unitarity conditions (XX.53). All of the following properties are consequences of relation (XX.66) alone.

γ^A MATRICES. As the γ^μ anticommute, and as their square is equal to $+I$ or $-I$, any product of several γ^μ is equal, to within a sign, to one of the 16 particular γ^A matrices given in Table 1. These have been grouped in 5 classes denoted by (S), (V), (T), (A) and (P) and containing 1, 4, 6, 4 and 1 elements respectively; the reasons for this classification will become clear at the end of this section (cf. § 14).

TABLE XX.1

The γ^A matrices

	Matrix	Explicit form	
		$(\gamma^A)^2 = I$	$(\gamma^A)^2 = -I$
(S)	$1 \equiv$	I	
(V)	$\gamma^\mu \equiv \{\gamma^0, \gamma^k\} \equiv$	γ^0	γ^1 γ^2 γ^3
(T)	$\gamma^{[\lambda\mu]} \equiv \{\gamma^k\gamma^0, \gamma^5\gamma^0\gamma^k\} \equiv$	$\gamma^1\gamma^0$ $\gamma^2\gamma^0$ $\gamma^3\gamma^0$	$\gamma^2\gamma^3$ $\gamma^3\gamma^1$ $\gamma^1\gamma^2$
(A)	$\gamma^{[\lambda\mu\nu]} \equiv \{\gamma^0\gamma^5, \gamma^k\gamma^5\} \equiv$	$\gamma^1\gamma^2\gamma^3$	$\gamma^0\gamma^2\gamma^3$ $\gamma^0\gamma^3\gamma^1$ $\gamma^0\gamma^1\gamma^2$
(P)	$\gamma^{[\lambda\mu\nu\varrho]} \equiv \gamma^5$		$\gamma^0\gamma^1\gamma^2\gamma^3$

It is clear that $(\gamma^A)^2$ is equal, according to the case, to I or to $-I$; the six matrices whose square is equal to I are grouped in the left-hand column; the ten matrices whose square is equal to $-I$ in the right-hand column.

Of all these matrices, only I commutes with all of the others. If

$\gamma^A \neq I$, it anticommutes with 8 of these 16 matrices and commutes with the 8 others.

In particular the matrix γ^5, defined by [1])

$$\gamma^5 \equiv \gamma^0\gamma^1\gamma^2\gamma^3 \tag{XX.67}$$

anticommutes with the γ^μ:

$$\gamma^5\gamma^\mu + \gamma^\mu\gamma^5 = 0 \tag{XX.68}$$

and its square is

$$(\gamma^5)^2 = -I. \tag{XX.69}$$

INVERSE MATRICES (γ_A). We define the γ_μ matrices by the relation:

$$\gamma_\mu = g_{\mu\nu}\gamma^\nu. \tag{XX.70}$$

Obviously:

$$\gamma^\mu = [\gamma_\mu]^{-1}.$$

It follows that the inverse of a γ^A is obtained by reversing the order of the γ^μ involved and replacing each of them by the corresponding γ_μ; we denote the expression thereby obtained by γ_A:

$$\gamma_A\gamma^A = \gamma^A\gamma_A = I. \tag{XX.71}$$

Following this method of constructing the inverse, one finds

$$\gamma_5 = \gamma_3\gamma_2\gamma_1\gamma_0.$$

TRACE AND DETERMINANT

$$\operatorname{Tr}\gamma^A = \begin{cases} 4 & \text{if } \gamma^A = I \\ 0 & \text{if } \gamma^A \neq I. \end{cases} \tag{XX.72}$$

To see this, suppose $\gamma^A \neq I$ and let γ^B be one of the 8 matrices which anticommute with γ^A:

$$\gamma^A = -\gamma^B\gamma^A\gamma_B.$$

We then have:

$$\operatorname{Tr}\gamma^A = -\operatorname{Tr}\gamma^B\gamma^A\gamma_B = -\operatorname{Tr}\gamma_B\gamma^B\gamma^A = -\operatorname{Tr}\gamma^A = 0.$$

We note in passing (Problem XX.3) that:

$$\det\gamma^A = 1.$$

REARRANGEMENT LEMMA. The following property may be verified by simple inspection:

If we multiply each matrix of the set of 16 γ^A matrices on the right

[1]) The index 4 is commonly used to denote the time component, according to the definition: $x^4 = ix^0 = ict$.

(or on the left) by a particular one of them, we obtain the same set of 16 matrices, except for possible changes in order and sign.

LINEAR INDEPENDENCE AND IRREDUCIBILITY. Using the rearrangement lemma and the properties of the trace, it can easily be shown that:

(*i*) The 16 γ^A matrices are linearly independent.

(*ii*) Any 4th order matrix M is a uniquely-defined linear combination of the γ^A:

$$M = \sum_A m_A \gamma^A \qquad m_A = \tfrac{1}{4} \operatorname{Tr} \gamma_A M.$$

(*iii*) Any matrix commuting with every γ^μ, and therefore with every γ^A, is a multiple of the unit matrix:

$$\text{if } [M, \gamma^\mu] = 0 \text{ for any } \mu, \; M = Cst. \times I.$$

FUNDAMENTAL THEOREM. *Let γ^μ and γ'^μ be two sets of 4 fourth-order matrices satisfying relation* (XX.66). *There exists a matrix S, non-singular* (det $S \neq 0$) *and defined to within a constant, such that*

$$\gamma^\mu = S \gamma'^\mu S^{-1} \qquad (\mu = 0, 1, 2, 3). \tag{XX.73}$$

To demonstrate this theorem we proceed in the following way. To each set γ^μ, γ'^μ there corresponds a set of 16 matrices γ^A, γ'^A whose definition and principle properties were given above; thus to each particular γ^A there corresponds a certain matrix γ'^A, the index A taking all 16 distinct values. Let F be a certain matrix and denote by S the matrix defined by

$$S \equiv \sum_A \gamma'^A F \gamma_A,$$

the sum $\sum_A$ being extended over the 16 possible values of the index A.

Denote a particular matrix by γ^B, its inverse by γ_B and the corresponding matrix in the other system by γ'^B; in virtue of the rearrangement lemma,

$$\gamma'^B S \gamma_B \equiv \sum_A \gamma'^B \gamma'^A F \gamma_A \gamma_B = \sum_A \gamma'^A F \gamma_A \equiv S,$$

therefore

$$\gamma'^B S = S \gamma^B. \tag{XX.74}$$

For relation (XX.73) to be verified by S, it remains to show that S has an inverse. To this effect, we introduce the matrix T, defined by

$$T \equiv \sum_A \gamma^A G \gamma_A',$$

where G is an arbitrary matrix. Reasoning in an analogous way one shows that

$$\gamma^B T = T\gamma'^B.$$

Therefore

$$\gamma^B TS = T\gamma'^B S = TS\gamma^B,$$

whatever γ^B; since TS commutes with any matrix γ^B, it is a multiple of unity: $TS = c \times I$. The multiplication constant is given by the formula:

$$c = \tfrac{1}{4}\operatorname{Tr} TS = \tfrac{1}{4}\sum_A \sum_B \operatorname{Tr} \gamma^A G \gamma'_A \gamma'^B F \gamma_B$$

$$= \tfrac{1}{4}\operatorname{Tr} G\left(\sum_A \sum_B \gamma'_A \gamma'^B F \gamma_B \gamma^A\right) = 4 \operatorname{Tr} GS.$$

Now F can always be chosen so that S has at least one non-zero element, for clearly the 16 γ^A matrices would not be linearly independent if S vanished for every F. Thus it is always possible to choose G such that

$$\operatorname{Tr} GS \equiv \sum_s \sum_t G_{st} S_{ts} = \tfrac{1}{4};$$

one then has $c = 1$ and therefore $T = S^{-1}$. Thus S does have an inverse and property (XX.73) is obtained by multiplying both sides of (XX.74) on the right by S^{-1}.

If another matrix, S', has the same property, $S^{-1}S'$ commutes with all of the γ^μ and therefore: $S^{-1}S' = Cst \times I$. Conversely, if S has the property (XX.73), then so has any multiple of S. Thus we have shown that the non-singular matrix S exists, and that it is defined to within a constant. Q.E.D.

UNITARY γ^μ MATRICES. If the matrices obeying relations (XX.66) are unitary:

$$\gamma_\mu \equiv \gamma^0 \gamma^\mu \gamma^0 = \gamma^{\mu\dagger}, \tag{XX.75}$$

all of the γ^A matrices are unitary and it follows that they are Hermitean or anti-Hermitean according as $(\gamma^A)^2$ is equal to $+I$ or $-I$.

The fundamental theorem is completed by the following corollary, the proof of which is left to the reader:

Let γ^μ and γ'^μ be two systems of 4 fourth-order unitary matrices satisfying relations (XX.66). *There exists a unitary matrix U, defined to within a phase, such that*: $\gamma'^\mu = U\gamma^\mu U^\dagger$ $(\mu = 0, 1, 2, 3)$.

COMPLEX CONJUGATION, B MATRIX. In particular, if the γ^μ are unitary and obey relations (XX.66), the 4 complex conjugate matrices

$\gamma^{\mu *}$ are also unitary and also obey relations (XX.66). The preceding corollary therefore applies: the γ^{μ} are obtained from the $\gamma^{\mu *}$ by a unitary transformation. We shall henceforth denote the matrix of that transformation by B (B is defined to within a phase):

$$\gamma^{\mu} = B\gamma^{\mu *}B^{\dagger} \qquad \gamma^{\mu *} = B^{*}\gamma^{\mu}\widetilde{B}. \tag{XX.76}$$

It can be shown (Problem XX.4) that B is antisymmetrical:

$$B = -\widetilde{B}$$

or, what amounts to the same, that

$$BB^{*} = B^{*}B = -I. \tag{XX.77}$$

If the γ matrices are those of the Dirac representation one has

$$\begin{aligned} B \equiv B_D &= \gamma^2\gamma^5 \\ &= -\mathrm{i}\varrho_3\sigma_y. \end{aligned}$$

Property (XX.77) is easily verified for this particular case.

11. Invariance of the Form of the Dirac Equation in an Orthochronous Change of Referential

In order to satisfy the relativity principle, the Dirac equation and the equation of continuity must have the same form in different Lorentz referentials. In actual fact, the principle requires this invariance of form only with respect to proper Lorentz transformations [1]), but it happens that the theory is formally invariant with respect to the complete group. We shall begin with a detailed study of invariance with respect to the orthochronous group. Time-reversal invariance will be examined at the end of this section along with other invariance properties characteristic of the Dirac equation but not directly related to Lorentz transformations.

---

[1]) And also with respect to space and time translations. This invariance can easily be demonstrated by an argument analogous to the one given in this paragraph. If the origin of axes is displaced by a four-vector a^{μ}, that is, if $x'^{\mu} = x^{\mu} + a^{\mu}$, then $A_{\mu}'(x') = A_{\mu}(x)$ and the law for the transformation of wave functions [that is, the analogue of (XX.85)] becomes simply:

$$\Psi'(x') = \Psi(x).$$

Let us therefore suppose that the dynamical state of the electron is represented in a referential (R) by a four-component wave function satisfying the Dirac equation:

$$[\gamma^{\mu}(i\partial_{\mu} - eA_{\mu}(x)) - m]\,\Psi(x) = 0. \tag{XX.78}$$

We suppose that a representation has been chosen once and for all for the operators of the space $\mathscr{E}^{(s)}$; the symbols γ^{μ} therefore denote well-defined matrices and relation (XX.78) represents 4 equations $(s = 1, 2, 3, 4)$:

$$\sum_{t=1,2,3,4} \sum_{\mu} (\gamma^{\mu})_{st} \left(i \frac{\partial}{\partial x^{\mu}} - eA_{\mu}(x^0x^1x^2x^3) \right) \psi_t(x^0x^1x^2x^3) - m\psi_s(x^0x^1x^2x^3) = 0$$

satisfied by the 4 components $\psi_s(x)$ of the wave function.

Consider this same system in a new referential (R') obtained from the preceding one by a certain orthochronous Lorentz transformation $\mathscr{L}$:

$$(R') = \mathscr{L}(R).$$

$\mathscr{L}$ is characterized by a certain matrix $\Omega^{\mu}{}_{\nu}$ having properties (XX.12) and (XX.13) (in addition $\Omega^0{}_0 > 0$) and defining the linear correspondence between the coordinates x^{μ} of a given point in the referential (R) and the coordinates x'^{μ} *of the same point* in the referential (R') that is, the law for the transformation of contravariant vectors [eq. (XX.11) and (XX.15)]. We write symbolically

$$x' = \mathscr{L}x, \qquad x = \mathscr{L}^{-1}x'. \tag{XX.79}$$

The partial-differentiation operators transform like covariant vectors:

$$\partial_{\mu} = \partial_{\nu}'\Omega^{\nu}{}_{\mu}. \tag{XX.80}$$

If we denote the covariant components of the potential in the new referential by $A_{\mu}'(x')$, they too are related to the $A_{\mu}(x)$ by the law for the transformation of covariant vectors:

$$A_{\mu}(x) \equiv A_{\mu}(\mathscr{L}^{-1}x') = A_{\nu}'(x')\Omega^{\nu}{}_{\mu}. \tag{XX.81}$$

Considered as a function of the new coordinates, $\Psi(x)$ obeys the equation obtained from (XX.78) by substituting into it from (XX.80) and (XX.81):

$$[\hat{\gamma}^{\mu}(i\partial_{\mu}' - eA_{\mu}'(x')) - m]\,\Psi(\mathscr{L}^{-1}x') = 0, \tag{XX.82}$$

where

$$\hat{\gamma}^{\mu} \equiv \Omega^{\mu}{}_{\varrho}\gamma^{\varrho}. \tag{XX.83}$$

The four γ^μ matrices are unitary and satisfy relations (XX.66). The four $\hat{\gamma}^\mu$ matrices are not necessarily unitary, but, because of the orthogonality of the $\Omega^\mu{}_\nu$ [relations (XX.13)], they also verify relations (XX.66), i.e.

$$\hat{\gamma}^\mu\hat{\gamma}^\nu+\hat{\gamma}^\nu\hat{\gamma}^\mu = \Omega^\mu{}_\varrho\Omega^\nu{}_\varrho(\gamma^\varrho\gamma^\sigma+\gamma^\sigma\gamma^\varrho)$$
$$= 2\Omega^\mu{}_\varrho g^{\varrho\sigma}\Omega^\nu{}_\sigma = 2g^{\mu\nu}.$$

Therefore in virtue of the fundamental theorem of § 10, there exists a non-singular matrix Λ defined to within a constant which transforms the $\hat{\gamma}$ into the γ:

$$\hat{\gamma}^\mu \equiv \Omega^\mu{}_\varrho\gamma^\varrho = \Lambda^{-1}\gamma^\mu\Lambda \qquad (\mu = 0, 1, 2, 3). \tag{XX.84}$$

Substituting this relationship into (XX.82) and putting

$$\Psi'(x') = \Lambda\Psi(x) \equiv \Lambda\Psi(\mathscr{L}^{-1}x') \tag{XX.85}$$

one finds, after multiplying on the left by Λ:

$$[\gamma^\mu(i\partial_\mu' - eA_\mu'(x'))]\,\Psi'(x') = 0.$$

This wave equation describes the evolution of the system in the new referential. It is seen to be formally identical with (XX.78). The Dirac equation is therefore formally invariant in an orthochronous change of referential and the law for the transformation of the wave function is given by equation (XX.85).

The matrix Λ, which is defined to within a constant, cannot in general be chosen to be unitary; however it will now be shown that the constant can always be chosen so as to have

$$\Lambda^\dagger = \gamma^0\Lambda^{-1}\gamma^0 \tag{XX.86}$$

Λ is then defined up to an arbitrary phase.

Since the $\Omega^\mu{}_\varrho$ are real and since the γ^μ are unitary and therefore verify (XX.75), the comparison of (XX.83) and its Hermitean conjugate gives

$$\hat{\gamma}^{\mu\dagger} = \gamma^0\hat{\gamma}^\mu\gamma^0.$$

Taking the Hermitean conjugate of (XX.84) and substituting the preceding relation one easily obtains

$$\hat{\gamma}^\mu = (\gamma^0\Lambda^\dagger\gamma^0)\gamma^\mu(\gamma^0\Lambda^\dagger\gamma^0)^{-1}.$$

Comparing with (XX.84), we see that the matrix $\Lambda\gamma^0\Lambda^\dagger\gamma^0$ commutes

with the four γ^μ and is therefore a multiple of the unit matrix:

$$\Lambda^\dagger = c\gamma^0\Lambda^{-1}\gamma^0. \tag{XX.87}$$

Next we show that the constant c is necessarily real and positive. From (XX.87) and (XX.84),

$$\Lambda^\dagger\Lambda = c\gamma^0(\Lambda^{-1}\gamma^0\Lambda) = c(\Omega^0{}_0 + \sum_k \Omega^0{}_k\gamma^0\gamma^k)$$

whence, taking into account (XX.72), $\mathrm{Tr}\Lambda^\dagger\Lambda = 4c\Omega^0{}_0$. Now since the trace of the Hermitean-definite matrix $\Lambda^\dagger\Lambda$ is necessarily real and positive, and since $\Omega^0{}_0$ is also real and positive, then so also is c. Thus if one multiplies Λ by $\sqrt{c}$, the resulting matrix is also a Λ matrix and verifies equation (XX.86). Q.E.D.

From the law (XX.85) for the transformation of wave functions, we obtain the law for the transformation of the adjoint functions:

$$\overline{\Psi}' \equiv \Psi'^\dagger\gamma^0 = \Psi^\dagger\Lambda^\dagger\gamma^0 = \overline{\Psi}\gamma^0\Lambda^\dagger\gamma^0$$

whence, taking into account (XX.86)

$$\overline{\Psi}'(x') = \overline{\Psi}(x)\Lambda^{-1}. \tag{XX.88}$$

Using this transformation law, the reader may easily verify that the adjoint equation (XX.64) is also formally invariant in an orthochronous change of referential.

It remains to show that the equation of continuity is formally invariant or, better, that j^μ [definition (XX.65)] transforms like a contravariant four-vector [1]).

This is easily verified; from (XX.85), (XX.88) and (XX.84),

$$\begin{aligned} j'^\mu(x') &\equiv (\overline{\Psi}'\gamma^\mu\Psi') = (\overline{\Psi}\Lambda^{-1}\gamma^\mu\Lambda\Psi) = \Omega^\mu{}_\varrho\,(\overline{\Psi}\gamma^\varrho\Psi) \\ &= \Omega^\mu{}_\varrho\, j^\varrho(x). \end{aligned}$$

For each Lorentz transformation, Λ is defined up to a phase by conditions (XX.84) and (XX.86). *In the present case, this phase has no physical significance.* In so far as possible it is desirable to remove the arbitrary in the phase, while preserving the property of the Λ to form a group homomorphic to the orthochronous Lorentz group (see the discussion of § XV.8).

[1]) Otherwise, the normalization of the wave function would depend on the reference system and the interpretation of j^0 as a position probability density would not be justified.

Now since the $\Omega^{\mu}{}_{\nu}$ are real, condition (XX.84) gives

$$\Omega^{\mu}{}_{\nu}\gamma^{\nu *} = (\Lambda^*)^{-1}\gamma^{\mu *}\Lambda^*,$$

whence, introducing the unitary matrix B [definition (XX.76)],

$$\Omega^{\mu}{}_{\nu}\gamma^{\nu} = (B\Lambda^*B^{\dagger})^{-1}\gamma^{\mu}(B\Lambda^*B^{\dagger}).$$

Comparing this equation with (XX.84), we see that $B\Lambda^*B^{\dagger}\Lambda^{-1}$ commutes with the four γ^{μ} matrices and is therefore a multiple of the unit matrix; it is easy to see [for example by calculating $\det(B\Lambda^*B^{\dagger}\Lambda^{-1})$] that this multiple is a phase factor; in other words

$$\Lambda^* = e^{i\lambda}B^{\dagger}\Lambda B.$$

since Λ is determined up to a phase factor, one can always choose this phase so as to have $e^{i\lambda} = 1$. In what follows this will always be done; Λ is then defined up to a sign.

In conclusion, to each orthochronous Lorentz transformation there correspond *two* Λ matrices, differing by a sign, defined by the three conditions

$$\Omega^{\mu}{}_{\nu}\gamma^{\nu} = \Lambda^{-1}\gamma^{\mu}\Lambda \qquad \text{(XX.89}a\text{)}$$

$$\Lambda^{\dagger} = \gamma^0\Lambda^{-1}\gamma^0 \qquad \text{(XX.89}b\text{)}$$

$$\Lambda^* = B^{\dagger}\Lambda B. \qquad \text{(XX.89}c\text{)}$$

It is clear that the set of Λ matrices thus defined forms a group and that this group is homomorphic to the orthochronous Lorentz group. We shall see in § 12 that the arbitrary in the sign of the Λ cannot be removed without violating this group property [1]).

[1]) Rather than condition (XX.89c), one could just as well take the more general condition: $\Lambda^* = \eta B^{\dagger}\Lambda B$, where η is a constant depending on the particular Lorentz transformation considered. The group property of the Λ is preserved if the η form an Abelian representation of the Lorentz group. Consequently, one necessarily has $\eta = 1$ for the transformations of the proper group $\mathscr{L}_0$, which gives back condition (XX.89c). For transformations of the reflection type that is for those of the sheet $s\mathscr{L}_0$, one may choose between the two following possibilities:

(a) $\eta = 1$ for any $s\mathscr{L}_0$; we obtain (XX.89c);

(b) $\eta = -1$ for any $s\mathscr{L}_0$, that is, $\Lambda^* = -B\Lambda B^{\dagger}$.

The physical content of the theory is obviously independent of this choice. The two groups, $G^{(a)}$ and $G^{(b)}$, which correspond respectively to (a) and (b) above, are both homomorphic to the orthochronous Lorentz group, but they are *not* isomorphic to each other. In particular, the two matrices which correspond to s have their square equal to I in $G^{(a)}$, and to $-I$ in $G^{(b)}$ [Cf. note 1, p. 908 below].

12. Transformation of the Proper Group

We shall now find explicit expressions for the Λ matrices defined by equation (XX.89). In this paragraph we limit ourselves to the transformations of the proper group.

Let us first consider the infinitesimal transformations. To each of the 6 infinitesimal "rotations" $g_{\mu\nu}-\varepsilon Z_{\mu\nu}^{(\alpha\beta)}$ there corresponds a matrix $\Lambda^{(\alpha\beta)}(\varepsilon)$ differing by an infinitesimal from the unit matrix and which may therefore be written in the form

$$\Lambda^{(\alpha\beta)}(\varepsilon) \simeq I+\mathrm{i}\varepsilon S_{\alpha\beta}, \tag{XX.90}$$

where $S_{\alpha\beta}$ is a finite matrix to be determined. One has

$$[\Lambda^{(\alpha\beta)}(\varepsilon)]^{-1} \simeq \Lambda^{(\alpha\beta)}(-\varepsilon) \simeq I-\mathrm{i}\varepsilon S_{\alpha\beta}.$$

Property (XX.89a) therefore gives

$$-\varepsilon g^{\mu\nu} Z_{\nu\varrho}^{(\alpha\beta)} \gamma^{\varrho} = -\mathrm{i}\varepsilon[S_{\alpha\beta}, \gamma^{\mu}],$$

or, using (XX.17),

$$[S_{\alpha\beta}, \gamma^{\mu}] = \mathrm{i}(\delta^{\mu}{}_{\beta}\gamma_{\alpha}-\delta^{\mu}{}_{\alpha}\gamma_{\beta}).$$

$S_{\alpha\beta}$ satisfies the same commutation relations with the γ^{μ} as the matrix $\frac{1}{2}\mathrm{i}\gamma_{\alpha}\gamma_{\beta}$. The difference therefore commutes with the four γ^{μ} matrices and is thus a constant. One easily sees that conditions (XX.89b) and (XX.89c) are satisfied if, and only if, this constant vanishes. It is convenient to put

$$\begin{aligned} \sigma_{\mu\nu} &\equiv \tfrac{1}{2}\mathrm{i}[\gamma_{\mu}, \gamma_{\nu}] \\ &\equiv \mathrm{i}\gamma_{\mu}\gamma_{\nu} \qquad (\mu \neq \nu). \end{aligned} \tag{XX.91}$$

One finds therefore:

$$\boxed{S_{\alpha\beta} = \tfrac{1}{2}\sigma_{\alpha\beta}} \tag{XX.92}$$

The symbols $S_{\alpha\beta}$ and $\sigma_{\alpha\beta}$ will also be used for the operators represented by the matrices $S_{\alpha\beta}$ and $\sigma_{\alpha\beta}$ respectively. We shall see further on that $S_{\alpha\beta}$ is an antisymmetrical tensor operator (6 components) representing the intrinsic angular momentum or spin of the particle. To be exact, the spin is the spatial part (3 components) of this operator. $S_{\alpha\beta}$ is related to the operators $\boldsymbol{\sigma}$ and $\boldsymbol{\alpha}$ of § 6 and 7 by the relations

$$S_{10} = \tfrac{1}{2}\mathrm{i}\alpha_x \qquad S_{20} = \tfrac{1}{2}\mathrm{i}\alpha_y \qquad S_{30} = \tfrac{1}{2}\mathrm{i}\alpha_z \tag{XX.93a}$$

$$S_{23} = \tfrac{1}{2}\sigma_x \qquad S_{31} = \tfrac{1}{2}\sigma_y \qquad S_{12} = \tfrac{1}{2}\sigma_z. \tag{XX.93b}$$

Any finite transformation of the proper Lorentz group may be considered as the product of successive infinitesimal transformations. We can therefore construct the Λ matrices corresponding to a finite change of referential by taking the product of matrices for infinitesimal transformations as defined above. If we proceed in this way conditions (XX.89*b*) and (XX.89*c*) are automatically fulfilled and we obtain one of the two possible Λ matrices.

In particular, the "rotation" of angle φ in the plane $x^\alpha x^\beta$ is a product of infinitesimal rotation matrices in this plane, and the matrix $\Lambda^{(\alpha\beta)}(\varphi)$ representing the transformation is therefore

$$\Lambda^{(\alpha\beta)}(\varphi) = \mathrm{e}^{\mathrm{i}\varphi S_{\alpha\beta}} \tag{XX.94}$$

Thus (cf. note, p. 882), if it is a special Lorentz transformation of velocity $v = \tanh\varphi$ directed along the x axis, one finds, taking into account relations (XX.93*a*) and the properties of α_x:

$$\Lambda^{xt}(\varphi) = \mathrm{e}^{-\frac{1}{2}\alpha_x\varphi} = \cosh\tfrac{1}{2}\varphi - \alpha_x \sinh\tfrac{1}{2}\varphi. \tag{XX.95}$$

More generally let $\Lambda_{\mathrm{sp}}(v)$ be the matrix associated with the special Lorentz transformation of velocity $\boldsymbol{v}$; then

$$\Lambda_{\mathrm{sp}}(\boldsymbol{v}) = \cosh\tfrac{1}{2}\varphi - (\boldsymbol{\alpha}\cdot\boldsymbol{u}) \sinh\tfrac{1}{2}\varphi,$$

where:

$$\boldsymbol{u} \equiv \boldsymbol{v}/v, \qquad \varphi = \tanh^{-1}\boldsymbol{v}.$$

Let us put

$$b \equiv 1/\sqrt{1-v^2} = \cosh\varphi. \tag{XX.96}$$

After an elementary calculation, the preceding expression takes the form:

$$\Lambda_{\mathrm{sp}}(\boldsymbol{v}) = \frac{1}{\sqrt{2(1+b)}}\,[1 + b - (\boldsymbol{\alpha}\cdot\boldsymbol{v})\,b]. \tag{XX.97}$$

Consider now rotations in the ordinary sense of the word. Expression (XX.94) gives for rotations about Oz:

$$\Lambda^{(xy)}(\varphi) = \mathrm{e}^{\mathrm{i}S_{12}\varphi} = \mathrm{e}^{\frac{1}{2}\mathrm{i}\sigma_z\varphi} = \cos\tfrac{1}{2}\varphi + \mathrm{i}\sigma_z \sin\tfrac{1}{2}\varphi. \tag{XX.98}$$

More generally the matrix $\Lambda_u(\varphi)$ associated with a rotation through an angle φ about the axis parallel to the unit vector $\boldsymbol{u}$ is

$$\Lambda_u(\varphi) = \cos\tfrac{1}{2}\varphi + \mathrm{i}\sigma_u \sin\tfrac{1}{2}\varphi \tag{XX.99}$$

with:

$$\sigma_u \equiv (\boldsymbol{\sigma}\cdot\boldsymbol{u}).$$

We are now ready to discuss the spin of the particle described by the Dirac equation. The spin is defined by the transformation properties of the internal variables with respect to spatial rotations. Equation (XX.99) gives the general expression for the transformation matrices of the internal variables in a rotation. The only difference between this expression and the one given by equation (XIII.84) is the sign in front of σ_u; one passes from one to the other by changing φ into $-\varphi$; it follows that the two matrices are inverses one of the other. This difference is due to the fact that in Chapter XIII, we considered the rotation of the variables and of the states while keeping the axes fixed; here we have taken the opposite point of view. We now see that the Dirac wave function transforms under rotation like the wave function of *a particle of spin* $\frac{1}{2}$.

Note in particular that a rotation through 2π about any axis does not give back the unit matrix; indeed one finds that

$$\Lambda_u(2n\pi) = (-)^n I, \tag{XX.100}$$

property characteristic of half-integral spins. From this it is clear that the arbitrary in the sign of the Λ cannot be removed without violating their property of forming a group.

We shall henceforth call the wave functions of the Dirac theory *spinors*.

13. Spatial Reflection and the Orthochronous Group

Once the law of transformation of spinors in proper changes of referential has been determined, we need only to know the law of transformation in the reflection s to be able to determine the law of transformation in any orthochronous change of referential.

Denote by Λ_s the matrix corresponding to the reflection s of the referential. In this case (XX.85) becomes

$$\Psi'(t, \mathbf{r}) = \Lambda_s \Psi(t, -\mathbf{r}). \tag{XX.101}$$

Condition (XX.89a) gives

$$\Lambda_s^{-1} \gamma^0 \Lambda_s = \gamma^0 \qquad \Lambda_s^{-1} \gamma^k \Lambda_s = -\gamma^k,$$

whence

$$\Lambda_s = c_s \gamma^0.$$

The constant c_s is fixed by (XX.89b) and (XX.89c); one finds

$$c_s = \pm 1.$$

Therefore:

$$\Lambda_s = \pm \gamma^0 \qquad \text{(XX.102)}$$

Λ_s is defined up to a sign in accordance with what was said above [1]).

14. Construction of Covariant Quantities

From the components of the spinor $\Psi(x)$ and those of its adjoint $\overline{\Psi}(x)$, one can form in all 16 linearly independent functions of x^0, x^1, x^2, x^3 that are bilinear in Ψ and $\overline{\Psi}$. They can be grouped into 5 fields of well-defined tensorial character, namely: a scalar S, a vector V^μ, an antisymmetrical tensor with two indices $T^{[\mu\nu]}$, an antisymmetrical tensor with three indices, or pseudovector $A^{[\lambda\mu\nu]}$ and an antisymmetrical tensor with four indices, or pseudoscalar P. These are given in Table XX.2.

The indicated tensorial characters can easily be demonstrated using the law for the transformation of the spinors Ψ and $\overline{\Psi}$ [eq. (XX.85) and (XX.88)] and relation (XX.89a) between the matrix Λ and the coefficients $\Omega^\mu{}_\nu$ of the corresponding Lorentz transformation.

Recall that the law for the transformation of a pseudoscalar differs from that of a scalar only by the presence of the additional factor $\det |\Omega^\mu{}_\nu|$:

$$P(x') = \det |\Omega^\mu{}_\nu| \, P(x).$$

TABLE XX.2

Tensors bilinear in Ψ and $\overline{\Psi}$

Tensor	Number of Components	Nature
$S(x) \equiv (\overline{\Psi}\Psi)$	1	Scalar
$V^\mu(x) \equiv (\overline{\Psi}\gamma^\mu\Psi)$	4	Vector
$T^{[\mu\nu]}(x) \equiv (\overline{\Psi}\gamma^\mu\gamma^\nu\Psi) \quad (\mu \neq \nu)$	6	Tensor of rank 2
$A^{[\lambda\mu\nu]}(x) \equiv (\overline{\Psi}\gamma^\lambda\gamma^\mu\gamma^\nu\Psi) \quad (\lambda \neq \mu, \mu \neq \nu, \nu \neq \lambda)$	4	Pseudovector
$P(x) \equiv (\overline{\Psi}\gamma^5\Psi)$	1	Pseudoscalar

[1]) Expression (XX.102) corresponds to choice (a) defined in note, p. 904. Choice (b) leads to:

$$\Lambda_s = \pm i\gamma^0.$$

Thus a pseudoscalar field transforms like a scalar field in a proper Lorentz transformation, but changes sign in a reflection s. Similarly, the law for the transformation of a pseudovector differs from that of a vector only by the presence of the additional factor $\det |\Omega^{\mu}_{\nu}|$.

The vector $V^{\mu}(x)$ has already been interpreted as the current-density four-vector:

$$V^{\mu}(x) \equiv j^{\mu}(x).$$

The other covariants are also capable of interpretation. Thus $T^{[\mu\nu]}$ is, to within a constant, equal to a tensor $S^{\mu\nu}$, which can be interpreted as a spin density:

$$T^{[\mu\nu]} = -2\mathrm{i}S^{\mu\nu}(x) \equiv -2\mathrm{i}(\overline{\Psi}S^{\mu\nu}\Psi).$$

15. A Second Formulation of the Invariance of Form: Transformation of States

In the preceding paragraph, each transformation is considered as an operation on the reference axes, and the physical system is not modified. Inversely, one may effect the transformation on the physical system and leave the axes fixed; this second point of view was systematically adopted in the third part (cf. in particular the remarks of § XIII.11). Although the results are expressed in different language, it is clear that the two points of view are equivalent.

In order to clarify this equivalence, let (S) be the state of the physical system represented in the referential (R) by the spinor $\Psi(x)$. Let (S') be the state obtained by effecting the transformation $\mathscr{L}$ on (S), and $(\hat{R})$ the referential which is taken into (R) by the same transformation (cf. Fig. (XX.1):

$$(S') = \mathscr{L}(S), \qquad (\hat{R}) = \mathscr{L}^{-1}(R).$$

We consider the following three spinors:

$\Psi(x)$ representing (S) in referential (R)

$\hat{\Psi}(\hat{x})$,, (S) ,, ,, $(\hat{R})$

$\Psi'(x)$,, (S') ,, ,, (R)

It is clear that $\hat{\Psi}$ and Ψ' are equal for equal values of their arguments:

$$\Psi'(x) = \hat{\Psi}(x). \tag{XX.103}$$

The correspondence between $\hat{\Psi}$ and Ψ was established in § 11. Since one passes from $(\hat{R})$ to (R) by the transformation $\mathscr{L}$, one has, applying

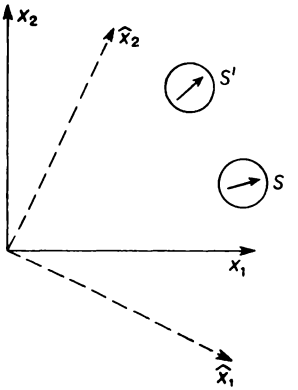


Fig. XX.1. The two ways of looking at a Lorentz transformation: change of referential ($\hat{x} \to x$) and transformation of the system ($S \to S'$).

(XX.85) and denoting the matrix associated with $\mathscr{L}$ by Λ:

$$\Psi(\hat{x}) = \Lambda\hat{\Psi}(\mathscr{L}^{-1}\hat{x}).$$

Therefore:

$$\Psi'(x) = \Lambda^{-1}\Psi(\mathscr{L}x). \qquad \text{(XX.104)}$$

Comparing with equation (XX.85), one sees that the transformation of states is realized by the inverse of the operator corresponding to the change of referential.

These remarks also apply to the electromagnetic field in which the Dirac particle moves. Denote this field by (A) and the field obtained by the transformation $\mathscr{L}$ by (A'):

$$(A') = \mathscr{L}(A).$$

We consider the following three (covariant) four-vectors:

$$\begin{array}{llllll} A_\mu(x) & \text{representing} & (A) & \text{in} & \text{referential} & (R), \\ \hat{A}_\mu(\hat{x}) & ,, & (A) & ,, & ,, & (\hat{R}), \\ A_\mu'(x) & ,, & (A') & ,, & ,, & (R). \end{array}$$

We can repeat the argument given above for the spinors. One obviously has:

$$A_\mu'(x) = \hat{A}_\mu(x). \qquad \text{(XX.105)}$$

But, according to (XX.81):

$$\hat{A}_\mu(\mathscr{L}^{-1}x) = A_\nu(x)\Omega^\nu{}_\mu$$

whence

$$A_\mu'(x) = A_\nu(\mathscr{L}x)\Omega^\nu{}_\mu. \tag{XX.106}$$

Suppose now that $\Psi(x)$ satisfies the Dirac equation in the potential $A_\mu(x)$:

$$[\gamma^\mu(i\partial_\mu - eA_\mu) - m]\Psi = 0. \tag{XX.107}$$

Taking into account equalities (XX.103) and (XX.105), the invariance of the form of the Dirac equation in the change of referential $(R) \to (\hat{R})$ gives

$$[\gamma^\mu(i\partial_\mu - eA_\mu') - m]\Psi' = 0. \tag{XX.108}$$

The invariance of form can therefore be expressed in the following way:

If $\Psi(x)$ satisfies the Dirac equation in a potential $A_\mu(x)$, the state $\Psi'(x)$ obtained by the transformation $\mathscr{L}$ satisfies the Dirac equation in the transformed potential $A_\mu'(x)$.

16. Invariance of the Law of Motion

Equations (XX.107) and (XX.108) are in general different. They are identical when the external potential (A) is invariant in the transformation $\mathscr{L}$, that is, when

$$A_\mu'(x) = A_\mu(x).$$

In this case, Ψ and Ψ' obey the *same* wave equation. Thus, *the law of motion of the dynamical states is invariant in any transformation $\mathscr{L}$ that conserves the external potential.*

In all of the preceding work $\mathscr{L}$ was any orthochronous Lorentz transformation. However, all that has been said can be repeated for space-time *translations* [cf. note, p. 900].

The two properties mentioned above — invariance of form and invariance of the law of motion — remain valid when $\mathscr{L}$ represents a space or time translation.

17. Transformation Operators. Momentum, Angular Momentum, Parity

To continue this analysis in accordance with the general scheme set forth in Chapter XV we write the transformation law (XX.104) in the form

$$\Psi' = T\Psi, \tag{XX.109}$$

where T is an appropriate linear operator. The invariance of the Dirac equation in the transformation can then be expressed by the operator relation

$$T\,\mathscr{D}(A)\,T^{-1}=\mathscr{D}(A'), \tag{XX.110}$$

in which $\mathscr{D}(A)$ and $\mathscr{D}(A')$ denote the Dirac operators in the potentials A and A' respectively:

$$\mathscr{D}(A)\equiv\gamma^{\mu}(\mathrm{i}\partial_{\mu}-eA_{\mu}). \tag{XX.111}$$

The condition for the invariance of the law of motion under the transformation $\mathscr{L}$ is then expressed by the commutation relation

$$[T,\mathscr{D}(A)]=0. \tag{XX.112}$$

The operator T is easily constructed. It is the product of an operator $T^{(s)}$ acting on the spin variables alone, and an operator $T^{(0)}$ acting on the orbital variables alone:

$$T=T^{(s)}\otimes T^{(0)}.$$

Comparing (XX.109) and (XX.104), one sees that

$$T^{(s)}=\Lambda^{-1}, \tag{XX.113}$$

where Λ denotes the *operator* represented by the matrix Λ defined in § 11.

Let us derive the explicit form of T for the infinitesimal Lorentz translations and "rotations", and for the reflection s.

For the translations one has $T^{(s)}=1$. Let us introduce the differential operator

$$\boxed{p_{\mu}\equiv\mathrm{i}\partial_{\mu}} \tag{XX.114}$$

p_{μ} represents the *energy-momentum four-vector* (more precisely, the covariant components of the energy-momentum four-vector). For an infinitesimal translation ε along the direction of the axis x^{α}, one finds

$$T=1+\mathrm{i}\varepsilon p_{\alpha}.$$

Consider now an "infinitesimal rotation" of angle ε in the plane $x^{\alpha}x^{\beta}$. In this case one has

$$\begin{aligned}(\mathscr{L}x)^{\mu} &= x^{\mu}-\varepsilon Z^{(\alpha\beta)\mu}{}_{\nu}x^{\nu}\\ &= x^{\mu}-\varepsilon(\delta^{\mu}{}_{\alpha}x_{\beta}-\delta^{\mu}{}_{\beta}x_{\alpha}).\end{aligned}$$

If $\psi_s(x)$ is a particular component of the spinor $\Psi(x)$, then to the first order in ε:

$$\psi_s(\mathscr{L}x) \simeq \psi_s(x) + \varepsilon\left(x_\alpha \frac{\partial \psi_s}{\partial x^\beta} - x_\beta \frac{\partial \psi_s}{\partial x^\alpha}\right).$$

If we introduce the differential operator

$$L_{\alpha\beta} \equiv x_\alpha p_\beta - x_\beta p_\alpha \tag{XX.115}$$

the preceding equation takes the form

$$\psi_s(\mathscr{L}x) \simeq (1 - i\varepsilon L_{\alpha\beta})\psi_s(x).$$

On the other hand, from (XX.90) and (XX.113),

$$T^{(s)} \simeq (1 - i\varepsilon S_{\alpha\beta}),$$

where $S_{\alpha\beta}$ is defined by equation (XX.92). Finally, formula (XX.109) giving the law for the transformation of a spinor becomes in the "infinitesimal rotation" case

$$\begin{aligned}\Psi'(x) &\simeq (1 - i\varepsilon S_{\alpha\beta})(1 - i\varepsilon L_{\alpha\beta})\Psi(x) \\ &\simeq (1 - i\varepsilon J_{\alpha\beta})\Psi(x),\end{aligned}$$

where

$$\boxed{J_{\alpha\beta} \equiv L_{\alpha\beta} + S_{\alpha\beta} \equiv x_\alpha p_\beta - x_\beta p_\alpha + \tfrac{1}{2}\sigma_{\alpha\beta}} \tag{XX.116}$$

The three spatial components J_{23}, J_{31} and J_{12} of $J_{\alpha\beta}$ are associated with infinitesimal rotations about the axes Ox, Oy and Oz respectively; they are the components of the total angular momentum $\boldsymbol{J}$ and one has

$$\begin{aligned}\boldsymbol{J} &= \boldsymbol{L} + \boldsymbol{S} \\ \boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p}, &\qquad \boldsymbol{S} = \tfrac{1}{2}\boldsymbol{\sigma}.\end{aligned} \tag{XX.117}$$

The components of $\boldsymbol{L}$ act on the orbital variables alone: $\boldsymbol{L}$ is the orbital angular momentum. The components of $\boldsymbol{S}$ act on the internal variables alone: $\boldsymbol{S}$ is the spin vector of the particle.

The reader will verify that $\boldsymbol{J}$, $\boldsymbol{L}$ and $\boldsymbol{S}$ satisfy the commutation relations characteristic of angular momenta and that $\boldsymbol{S}$ has the property

$$\boldsymbol{S}^2 = \tfrac{3}{4}$$

characteristic of a particle of spin $\frac{1}{2}$.

The operator associated with spatial reflection [1]) will be called the

[1]) See *Note added in proof*, p. 956.

parity operator and denoted by P. Let us denote by $P^{(0)}$ the "orbital parity" operator

$$P^{(0)}\Psi(t, \mathbf{r}) = \Psi(t, -\mathbf{r}).$$

According to (XX.113) we have for P the choice between two expressions differing by a sign, expressions easily obtained from the study of § 13 [cf. relation (XX.102)]. We adopt the most convenient one:

$$\boxed{P = \gamma^0 P^{(0)}} \qquad \text{(XX.118)}$$

Note that P is Hermitean and that $P^2 = 1$.

18. Conservation Laws and Constants of the Motion

If the transformation depends on the time, the associated operator T explicitly brings in the time dependence of Ψ. This occurs notably in the case of the time translations and the special Lorentz transformations.

On the other hand, *for any transformation independent of the time, the action of T is defined independently of the law of motion of the state vector to which it is applied.* T can then be defined as a transformation operator acting on the state-vectors and observables of the system as was done in Chapter XV (Section II); the invariance properties of the law of motion of the states may then be expressed in the form of conservation laws.

For example, if $\mathscr{L}$ is a spatial transformation, T is a certain function of the operators of reflection, infinitesimal translation and infinitesimal rotation, that is, a certain function of P, $\boldsymbol{J}$ and $\boldsymbol{p}$; T therefore commutes with γ^0. And since

$$\gamma^0\left(\mathrm{i}\frac{\partial}{\partial t} - H_D\right) \equiv \mathscr{D}(A) - m,$$

the commutation relation (XX.112) is equivalent, in this case, to

$$[T, H_D] = 0.$$

This condition is the same as the one studied in § XV.12, and what was said there concerning the connection between the symmetries of the Hamiltonian and the laws of conservation may be applied here.

Thus, if $A_\mu(x)$ is invariant under translation, one obtains the commutation relations

$$[\boldsymbol{p}, H_D] = 0$$

and the conservation of momentum. If $A_\mu(x)$ is spherically symmetrical,

$$[\boldsymbol{J}, H_D] = 0$$

and the total angular momentum is conserved. If $A_\mu(x)$ is invariant under reflection in the origin,

$$[P, H_D] = 0$$

and the parity is conserved.

19. Time Reversal and Charge Conjugation

In this paragraph, we shall demonstrate the invariance of the form of the Dirac equation under two antilinear operations, time reversal and charge conjugation. It is convenient, for this, to introduce an antiunitary operator K [1]) of state-vector space having certain particularly simple general properties.

THE ANTIUNITARY OPERATOR K. K is by definition the antiunitary operator which transforms $\boldsymbol{p}$ into $-\boldsymbol{p}$ and conserves $\boldsymbol{r}$ and γ^μ:

$$K\boldsymbol{r}K^\dagger = \boldsymbol{r}, \qquad K\boldsymbol{p}K^\dagger = -\boldsymbol{p}. \tag{XX.119}$$

$$K\gamma^\mu K^\dagger = \gamma^\mu \qquad (\mu = 0, 1, 2, 3). \tag{XX.120}$$

It will be shown that such an operator exists, is defined to within a phase, and has the property

$$K^2 = -1. \tag{XX.121}$$

That K, if it exists, is defined to within a phase follows from relations (XX.119–120) and the fact that state-vector space is irreducible with respect to the basic operators $\boldsymbol{r}$, $\boldsymbol{p}$ and γ^μ. Let us now choose a particular representation, the Dirac representation for example. Each *operator* γ^μ is then represented by a certain matrix $\gamma_D{}^\mu$. Denote the operator represented by the "B matrix" which transforms the $\gamma_D{}^\mu$ into their respective complex conjugates by B_D. Here, B_D will be considered as a (unitary) operator of the total space and not as an operator of spin space alone; it is a unitary operator commuting with $\boldsymbol{r}$ and $\boldsymbol{p}$. Let K_D be the complex-conjugation operator associated with the representation (definition of § XV.5). Relations (XX.76) give:

$$\gamma^\mu = B_D(K_D\gamma^\mu K_D{}^\dagger)B_D{}^\dagger.$$

[1]) This operator *is not* the time reversal operator. The latter is denoted below by K_T.

Therefore, the antiunitary operator

$$K \equiv B_D K_D$$

satisfies relations (XX.120). Since B_D commutes with $\boldsymbol{r}$ and $\boldsymbol{p}$, and since from the definition of K_D,

$$K_D \boldsymbol{r} K_D = \boldsymbol{r}, \qquad K_D \boldsymbol{p} K_D = -\boldsymbol{p},$$

K also satisfies relations (XX.119). Finally, since $K_D = K_D^{\dagger}$, (XX.77) gives:

$$B_D(K_D B_D K_D) \equiv K^2 = -1.$$

This is just (XX.121); it obviously remains true if K is multiplied by any phase factor. Q.E.D.

CHARGE-CONJUGATION. Multiplying both sides of equation (XX.107) on the left by K and using the fact that K is antilinear and commutes with γ^μ, ∂_μ and $A_\mu(x)$, one finds:

$$[\gamma^\mu(-\mathrm{i}\partial_\mu - eA_\mu(x)) - m]K\Psi(x) = 0. \tag{XX.122}$$

$K\Psi$ therefore satisfies a wave equation differing from the Dirac equation only by the substitution of $-\mathrm{i}$ for $+\mathrm{i}$. Let us multiply both sides by γ^5. Since γ^5 anticommutes with γ^μ and commutes with all of the other operators in the bracket on the left-hand side, we obtain

$$[\gamma^\mu(\mathrm{i}\partial_\mu + eA_\mu(x)) - m]\gamma^5 K\Psi(x) = 0. \tag{XX.123}$$

Put:

$$\boxed{K_C \equiv \gamma^5 K} \tag{XX.124}$$

$$\Psi^C(x) \equiv K_C \Psi(x). \tag{XX.125}$$

Equation (XX.123) now becomes:

$$[\gamma^\mu(\mathrm{i}\partial_\mu + eA_\mu(x)) - m]\Psi^C(x) = 0. \tag{XX.126}$$

The equation satisfied by $\Psi^C(x)$ differs from the one satisfied by $\Psi(x)$ only in the sign of the charge. Thus, *if* $\Psi(x)$ *represents the motion of a Dirac particle of mass* m *and charge* e *in the potential* $A_\mu(x)$, $\Psi^C(x)$ *represents the motion of a Dirac particle of the same mass* m *and of opposite charge* $(-e)$ *in the same potential* $A_\mu(x)$.

The spinors Ψ and Ψ^C are called charge conjugates one of the other, and the transformation K_C is called *charge conjugation.*

It follows from the properties of K and γ^5 that

$$K_C{}^2 = 1. \tag{XX.127}$$

Thus the correspondence between Ψ and Ψ^C is reciprocal. It is easy to show that charge conjugation commutes with translations and orthochronous Lorentz transformations. More precisely, if $L\Psi$ is the transform of Ψ in one of these transformations, its charge-conjugate is $L\Psi^C$ in the case of a translation or a proper Lorentz transformation, and $-L\Psi^C$ in the case of a reflection (cf. Problem XX.5).

TIME REVERSAL. The time reversal invariance of the Dirac equation can be demonstrated directly, but it is just as simple to start from the preceding results on charge conjugation.

A given potential $A_\mu(t, \mathbf{r})$ is created by a certain number of charges in motion. The time-reversed potential $A_\mu'(t, \mathbf{r})$, is obtained by reversing the motion of these charges. In this operation the currents, and therefore the magnetic field, change their sign while electric charges, and therefore the electric field, remain unchanged

$$\mathscr{H}'(t, \mathbf{r}) = -\mathscr{H}(-t, \mathbf{r}), \qquad \mathscr{E}'(t, \mathbf{r}) = \mathscr{E}(-t, \mathbf{r}).$$

It follows that A_μ "transforms like a pseudovector":

$$\mathbf{A}'(t, \mathbf{r}) = -\mathbf{A}(-t, \mathbf{r}) \qquad A_0'(t, \mathbf{r}) = A_0(-t, \mathbf{r}).$$

If we change t into $-t$ in equation (XX.126), we therefore have

$$[-\gamma^0(i\partial_0 - eA_0'(t, \mathbf{r})) + \sum_k \gamma^k(i\partial_k - eA_k'(t, \mathbf{r})) - m]\,\Psi^C(-t, \mathbf{r}) = 0.$$

Let us multiply both sides by $\gamma^5\gamma^0$. Since this operator anticommutes with γ^0 and commutes with γ^k, we obtain:

$$[\gamma^\mu(i\partial_\mu - eA_\mu'(t, \mathbf{r})) - m]\,\Psi'(t, \mathbf{r}) = 0 \tag{XX.128}$$

if we put:

$$\Psi'(t, \mathbf{r}) \equiv \gamma^5\gamma^0\Psi^C(-t, \mathbf{r}) \tag{XX.129}$$
$$= \gamma^0 K\Psi(-t, \mathbf{r}).$$

Let us introduce the (antiunitary) time-reversal operator:

$$\boxed{K_T \equiv \gamma^0 K} \tag{XX.130}$$

$\Psi'(t, \mathbf{r})$ is by definition the time-reversal transform of $\Psi(-t, \mathbf{r})$. It satisfies equation (XX.128). Therefore, if Ψ satisfies the Dirac equation

in the potential A_μ, its time-reversal transform Ψ' satisfies the Dirac equation in the potential A_μ' obtained from A_μ by time reversal.

In particular, if A_μ is invariant under time reversal (for example, if the particle is in a stationary electric field: $\boldsymbol{A}=0, \partial A_0/\partial t=0$), Ψ and Ψ' both obey the same Dirac equation.

From the properties of γ^0 and K, one obtains:

$$K_T{}^2 = -1. \tag{XX.131}$$

This result, characteristic of systems of half-integral angular momenta, has already been obtained in the non-relativistic case [eq. (XV.88)]. The consequences, in particular Kramers degeneracy, are also valid here.

Expressing B_D in terms of the ϱ and σ (cf. end of § 10), we easily obtain from definitions (XX.124) and (XX.130) the equivalent relations

$$K_C = \mathrm{i}\varrho_2 \sigma_y K_D$$
$$K_T = \mathrm{i}\sigma_y K_D.$$

These are useful when manipulating the operators K_C and K_T in the Dirac representation.

20. Gauge Invariance

For completeness, we shall mention here the property of gauge invariance (cf. § XXI.20).

Changing the gauge of the electromagnetic potential means replacing the components $A_\mu(x)$ by

$$A_\mu'(x) \equiv A_\mu(x) - \partial_\mu G(x), \tag{XX.132}$$

where $G(x)$ is an arbitrary function of the space-time coordinates. This gives

$$A_0' = A_0 - \frac{\partial G}{\partial t}, \qquad \boldsymbol{A}' = \boldsymbol{A} + \nabla G.$$

The electric and magnetic fields are invariant in such a transformation.

If $\Psi(x)$ is a solution of the Dirac equation in the potential A_μ, then clearly

$$\Psi'(x) \equiv \mathrm{e}^{\mathrm{i}eG(x)}\Psi(x) \tag{XX.133}$$

is a solution of the Dirac equation in the potential A_μ'. This is called the gauge-invariance property of the Dirac equation.

IV. INTERPRETATION OF THE OPERATORS AND SIMPLE SOLUTIONS

21. The Dirac Equation and the Correspondence Principle

When the electromagnetic field is not zero, the solutions of the Dirac equation satisfy a second-order equation different to the Klein–Gordon equation but conforming with the correspondence principle.

To obtain this equation, one can start from the covariant form (XX.50) and write that the action of the operator $(-i\gamma^\lambda D_\lambda - m)$ on the left-hand side gives zero:

$$[\gamma^\lambda\gamma^\mu D_\lambda D_\mu + m^2]\Psi = 0. \tag{XX.134}$$

From the algebraic properties of the γ^μ operators, one obtains

$$\gamma^\lambda\gamma^\mu \equiv g^{\lambda\mu} + \tfrac{1}{2}[\gamma^\lambda, \gamma^\mu]. \tag{XX.135}$$

By renaming the dummy indices, we have

$$[\gamma^\lambda, \gamma^\mu] D_\lambda D_\mu \equiv -[\gamma^\lambda, \gamma^\mu] D_\mu D_\lambda = \tfrac{1}{2}[\gamma^\lambda, \gamma^\mu]\,[D_\lambda, D_\mu]. \tag{XX.136}$$

and, from the definition of the operator D_μ [eq. (XX.10)]

$$\begin{aligned}[D_\lambda, D_\mu] &\equiv ie[\partial_\lambda, A_\mu] + ie[A_\lambda, \partial_\mu] \\ &\equiv ie\left(\frac{\partial A_\mu}{\partial x^\lambda} - \frac{\partial A_\lambda}{\partial x^\mu}\right) \equiv ieF_{\lambda\mu}.\end{aligned} \tag{XX.137}$$

Equations (XX.135–137) give

$$\gamma^\lambda\gamma^\mu D_\lambda D_\mu \equiv D_\mu D^\mu + eS^{\lambda\mu}F_{\lambda\mu}, \tag{XX.138}$$

where $S^{\lambda\mu}$ represents the spin of the particle [definition (XX.92)]. Equation (XX.134) can therefore be put in the form

$$[D_\mu D^\mu + eS^{\lambda\mu}F_{\lambda\mu} + m^2]\Psi = 0. \tag{XX.139}$$

Comparing (XX.139) with the form (XX.30′) of the Klein–Gordon equation, it is seen that it differs by the presence of the term

$$eS^{\lambda\mu}F_{\lambda\mu}, \tag{XX.140}$$

which is a term coupling the spin of the particle to the electromagnetic field. This term has no classical analogue and its contribution is negligible when the classical approximation is valid. The motion of a Dirac wave packet is then the same as the motion of a Klein–Gordon wave packet.

22. Dynamical Variables of a Dirac Particle

From time to time we have given the physical interpretation of a certain number of dynamical variables of the Dirac theory. We shall now take up this question in a more systematic way, and indicate, in particular, the variables of the Quantum Theory which correspond to the different classical quantities of § 4.

The relativistic invariance of the theory plays no essential role in this discussion. We adopt the same point of view as in non-relativistic Quantum Mechanics: the system is defined by giving a certain number of dynamical variables obeying a well-defined algebra, and the Dirac equation – in the Dirac form [eqs. (XX.36) and (XX.44)] – describes the evolution of the dynamical states in the Schrödinger "representation".

In what follows time is therefore treated as a simple parameter, while the spatial coordinates are included among the dynamical variables of the system. The fundamental variables are $\boldsymbol{r}$ and $\boldsymbol{p}$ together with the internal variables α and β. The whole of representation theory applies here without change. In particular, in the Dirac representation, the state-vectors $|\Psi\rangle$, $|\Phi\rangle$, ..., are represented by four-component wave functions, $\Psi(\boldsymbol{r})$, $\Phi(\boldsymbol{r})$, ... of the coordinates x, y, z. In this representation the scalar product $\langle\Phi|\Psi\rangle$ is defined as a summation over the 4 possible values of the index representing the internal degree of freedom and an integration over the coordinates x, y, z:

$$\langle\Phi|\Psi\rangle = \sum_{s=1}^{4} \int \varphi_s^*(\boldsymbol{r})\, \psi_s(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r}.$$

This definition of the scalar product is consistent with the definition of the position probability density given in § 6 [formula (XX.35)]. More generally, we shall adopt here without change the statistical interpretation of the theory as set forth in the first part of this book; in particular, the average value of an operator Q for a given state of the system is given by

$$\langle Q\rangle = \langle u|Q|u\rangle,$$

where $|u\rangle$ is the normalized ket representing that state.

The observables of the theory that do not act on the internal degree of freedom have an obvious interpretation; in particular we have:

$\boldsymbol{r}$, the position vector,

$\boldsymbol{p}$, the (Lagrange canonical) momentum, called the momentum in this book.

$\boldsymbol{\pi} \equiv \boldsymbol{p} - e\boldsymbol{A}(\boldsymbol{r}, t)$, the mechanical momentum.

Among the functions of $\boldsymbol{r}$ we have the operator

$$\delta(\boldsymbol{r} - \boldsymbol{r}_0),$$

which is the projector onto the subspace of the eigenvalue $\boldsymbol{r}_0$; it represents the position probability density at the point $\boldsymbol{r}_0$.

Among the observables depending on the internal degree of freedom we define [1]:

the energy:

$$H \equiv e\varphi + \boldsymbol{\alpha}\cdot\boldsymbol{\pi} + \beta m; \tag{XX.141}$$

the relativistic mass:

$$M \equiv H - e\varphi \equiv \boldsymbol{\alpha}\cdot\boldsymbol{\pi} + \beta m; \tag{XX.142}$$

the current density at $\boldsymbol{r}$:

$$\boldsymbol{j}(\boldsymbol{r}_0) \equiv \boldsymbol{\alpha}\delta(\boldsymbol{r} - \boldsymbol{r}_0); \tag{XX.143}$$

the total angular momentum:

$$\boldsymbol{J} \equiv (\boldsymbol{r}\times\boldsymbol{p}) + \tfrac{1}{2}\boldsymbol{\sigma}; \tag{XX.144}$$

the spin:

$$\boldsymbol{S} \equiv \tfrac{1}{2}\boldsymbol{\sigma}; \tag{XX.145}$$

the parity:

$$P \equiv \beta P^{(0)}. \tag{XX.146}$$

The definitions of H and M are based on the correspondence with classical mechanics; that of $\boldsymbol{j}(\boldsymbol{r}_0)$ follows from the equation of continuity; those of $\boldsymbol{J}$ and $\boldsymbol{S}$ are related to the transformations of the states under rotation and that of P to the transformations under reflection.

Finally, *the correspondence principle leads to the interpretation of* $\boldsymbol{\alpha}$ *as the velocity of the particle*. This interpretation is also suggested by the expression for the current density. To establish it, we compare the classical equations (XX.18), (XX.19) and (XX.21) with the corresponding quantum equations, and to do this we must obviously pass

[1] Note that $\boldsymbol{p}$ depends on the choice of gauge; only the momentum of the total system (particle + electromagnetic field) is independent of this choice. The same remarks apply to the energy H and to the angular momentum $\boldsymbol{J}$ (cf. § XXI.23).

over to the Heisenberg "representation". The Heisenberg equations for the variables $\boldsymbol{r}$ and $\boldsymbol{\pi}$ (in the Heisenberg "representation") are written:

$$\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} = -\,\mathrm{i}[\boldsymbol{r}, H]$$

$$\frac{\mathrm{d}\boldsymbol{\pi}}{\mathrm{d}t} = -\,\mathrm{i}[\boldsymbol{\pi}, H] + \frac{\partial \boldsymbol{\pi}}{\partial t}.$$

Replacing H and $\boldsymbol{\pi}$ on the right-hand side by the expressions given above and using the commutation or anti-commutation relations for the operators $\boldsymbol{r}$, $\boldsymbol{p}$, $\boldsymbol{\alpha}$ and β, one finds, after a rather long but straightforward calculation (Problem XX.6),

$$\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} = \boldsymbol{\alpha} \tag{XX.147}$$

$$\frac{\mathrm{d}\boldsymbol{\pi}}{\mathrm{d}t} = e(\mathscr{E} + \boldsymbol{\alpha} \times \mathscr{H}). \tag{XX.148}$$

Also, from (XX.142) and the properties of $\boldsymbol{\alpha}$, one has the identity

$$\boldsymbol{\pi} = \tfrac{1}{2}(M\boldsymbol{\alpha} + \boldsymbol{\alpha}M). \tag{XX.149}$$

Equations (XX.147–149) between dynamical variables in the Heisenberg "representation" may be identified respectively with equations (XX.18), (XX.21) and (XX.19) between the classical dynamical variables if one treats $\boldsymbol{\alpha}$ as the classical velocity $\boldsymbol{v}$.

Note that the components of the velocity $\boldsymbol{\alpha}$ do not commute, and that each of them has in all two eigenvalues, $+c$ and $-c$ ($+1$ and -1 in the units used here). Here we have no difficulty of principle, but simply an additional indication that the classical picture of the phenomena should not be taken too seriously. We shall return to this question in § 37.

23. The Free Electron. Plane Waves

In the rest of this section we examine the solutions of the Dirac equation, first in the absence of a field, then in a static central potential. Solving the Dirac equation is then equivalent to finding the eigensolutions of the Hamiltonian H_D. Unless otherwise indicated, the calculations will be made in the Dirac representation, and we shall frequently make use of the operators ϱ_1, ϱ_2, ϱ_3 and σ_x, σ_y, σ_z introduced in § 7.

We first suppose the field null. The Hamiltonian H_D then commutes with the three components of the momentum. We therefore wish to find the eigensolutions of H_D corresponding to a well-defined value $\boldsymbol{p}$ for the momentum. Such solutions are plane waves, that is, functions of the form

$$u(\boldsymbol{p})\,e^{i\boldsymbol{p}\cdot\boldsymbol{r}},$$

where $u(\boldsymbol{p})$ is a four-component spinor independent of $\boldsymbol{r}$. It is determined by the eigenvalue equation

$$Hu(\boldsymbol{p}) = Eu(\boldsymbol{p}), \tag{XX.150}$$

where H is the following operator of $\mathscr{E}^{(s)}$ space:

$$H \equiv \boldsymbol{\alpha}\cdot\boldsymbol{p}+\beta m \equiv \varrho_1(\boldsymbol{\sigma}\cdot\boldsymbol{p})+\varrho_3 m. \tag{XX.151}$$

A simple calculation gives:

$$H^2 = p^2+m^2.$$

The only possible eigenvalues of H are therefore the two values $\pm\sqrt{p^2+m^2}$, i.e.:

$$\begin{aligned} &E = \varepsilon E_p \qquad (\varepsilon = \pm 1)\\ &E_p = \sqrt{p^2+m^2}. \end{aligned} \tag{XX.152}$$

It is easy to show — for example by using the fact that ϱ_2 anticommutes with H — that these two values are doubly degenerate.

The *component* $\boldsymbol{\sigma}\cdot\boldsymbol{p}/2p$ *of the spin in the direction of* $\boldsymbol{p}$ *commutes with* H. This can be seen from the last of expressions (XX.151); (the other spin components do not commute with H). We are therefore led to look for the common eigensolutions of H and $\boldsymbol{\sigma}\cdot\boldsymbol{p}/2p$. We obtain the following 4 pairs of eigenvalues:

$$(+E_p,+\tfrac{1}{2}) \quad (+E_p,-\tfrac{1}{2}) \quad (-E_p,+\tfrac{1}{2}) \quad (-E_p,-\tfrac{1}{2}).$$

To each of these pairs there corresponds a single eigenstate. The corresponding eigenspinor is easily determined from the two eigenvalue equations. An alternative method for finding this spinor will be given in the following paragraph.

The components of the 4 eigenspinors (normalized to unity) are given in Table XX.3 for the particularly simple case when $\boldsymbol{p}$ is directed along the z axis. Recall that, in the Dirac representation, β and σ_z are represented by diagonal matrices.

TABLE XX.3

Components of the spinors corresponding to the wave of momentum $\mathbf{p} \equiv (0, 0, p)$ *in the Dirac representation* $(E_p = \sqrt{m^2 + p^2})$.

Energy $E =$		Positive $+ E_p$		Negative $- E_p$	
Spin $\boldsymbol{\sigma}\cdot\mathbf{p}/2p \equiv \frac{1}{2}\sigma_z$		$+\frac{1}{2}$	$-\frac{1}{2}$	$+\frac{1}{2}$	$-\frac{1}{2}$
$\left(\frac{2E_p}{E_p+m}\right)^{\frac{1}{2}} \times$	$u_1 =$	1	0	$-\frac{p}{E_p+m}$	0
	$u_2 =$	0	1	0	$\frac{p}{E_p+m}$
	$u_3 =$	$\frac{p}{E_p+m}$	0	1	0
	$u_4 =$	0	$-\frac{p}{E_p+m}$	0	1

The 4 spinors are normalized to unity: $u^\dagger u = 1$.

24. Construction of the Plane Waves by a Lorentz Transformation

When $A_\mu = 0$, any Lorentz transform of a solution of the Dirac equation is another solution of the Dirac equation. In particular, any plane wave of momentum $\mathbf{p}$ can be obtained by a Lorentz transformation from a plane wave of momentum zero. We shall now describe a method based on this remark for constructing the spinors $u(\mathbf{p})$ of the preceding paragraph.

For a zero momentum, equation (XX.150) becomes

$$\beta m u(0) = E u(0).$$

The two possible eigenvalues are $+m$ and $-m$. To the eigenvalue $\varepsilon m (\varepsilon = \pm 1)$ there corresponds the spinor $u^{(\varepsilon)}(0)$; it is an eigenvector of the operator β. We shall suppose it normalized to unity, and we fix the direction of the spin in an arbitrary way; $u^{(\varepsilon)}(0)$ is then defined up to a phase.

The plane wave

$$\Psi_0{}^{(\varepsilon)} = u^{(\varepsilon)}(0) e^{-i\varepsilon m t}$$

is a solution of the Dirac equation corresponding to a momentum zero and an energy εm, that is, to the energy-momentum four-vector $(\varepsilon m, 0)$.

Consider the same solution in a new referential having a velocity

$\mathbf{v} = -\mathbf{p}/\varepsilon\sqrt{m^2+p^2} = -\mathbf{p}/\varepsilon E_p$ with respect to the preceding one. In this new referential, the energy-momentum of the particle is

$$p^\mu \equiv (\varepsilon E_p, \mathbf{p}). \tag{XX.153}$$

The solution is there represented by the plane wave:

$$\begin{aligned}\Psi_{\mathbf{p}}^{(\varepsilon)} &= \Lambda_{\mathrm{sp}}(\mathbf{v})\, u^{(\varepsilon)}(0) \exp(-ip^\mu x_\mu) \\ &= [\Lambda_{\mathrm{sp}}(\mathbf{v})\, u^{(\varepsilon)}(0)] \exp[-i(\varepsilon E_p t - \mathbf{p}\cdot\mathbf{r})].\end{aligned}$$

The term in brackets is therefore proportional to one of the sought-for spinors $u(\mathbf{p})$ which we shall henceforth denote by $u^{(\varepsilon)}(\mathbf{p})$. Its norm is the time component of the associated current four-vector; it may be obtained from the current four-vector associated with $u^{(\varepsilon)}(0)$ by a Lorentz transformation; it is therefore equal to:

$$b \equiv (1-v^2)^{-\frac{1}{2}} = E_p/m.$$

We therefore adopt the definition:

$$u^{(\varepsilon)}(\mathbf{p}) \equiv b^{-\frac{1}{2}} \Lambda_{\mathrm{sp}}(\mathbf{v})\, u^{(\varepsilon)}(0).$$

Substituting (XX.97) into this definition with the values of $\mathbf{v}$ and b given above, one finds

$$u^{(\varepsilon)}(\mathbf{p}) = [2E_p(m+E_p)]^{-\frac{1}{2}}[m+E_p+\varepsilon\boldsymbol{\alpha}\cdot\mathbf{p}]u^{(\varepsilon)}(0), \tag{XX.154}$$

which is the required expression. In particular if $u^{(\varepsilon)}(0)$ is an eigenstate of $(\boldsymbol{\sigma}\cdot\mathbf{p})$, then so is $u^{(\varepsilon)}(\mathbf{p})$, which is therefore in this case one of the spinors defined in the preceding paragraph. In particular one obtains the results of Table 3 when $\mathbf{p}$ is directed along the z axis.

Expression (XX.154) may also be put in the form

$$u^{(\varepsilon)}(\mathbf{p}) = [2E_p(m+E_p)]^{-\frac{1}{2}}[m+\gamma^\mu p_\mu]u^{(\varepsilon)}(0), \tag{XX.155}$$

in which p_μ represents the energy-momentum four-vector defined by equation (XX.153).

25. Central Potential

We now look for the eigensolutions of a Dirac particle in a static central potential $V(r)$. The Dirac Hamiltonian is then

$$H_D \equiv \boldsymbol{\alpha}\cdot\mathbf{p}+\beta m+V(r). \tag{XX.156}$$

It is invariant under rotation and reflection:

$$[H_D, \mathbf{J}] = 0, \qquad [H_D, P] = 0.$$

We therefore look for eigensolutions of well-defined angular momentum and parity.

It is convenient to write the solutions Ψ in the form

$$\Psi = \begin{pmatrix} \Phi \\ \chi \end{pmatrix}, \tag{XX.157}$$

where

$$\Phi \equiv \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}, \qquad \chi \equiv \begin{pmatrix} \psi_3 \\ \psi_4 \end{pmatrix}. \tag{XX.158}$$

Projecting Ψ onto the subspaces $\beta = +1$ and $\beta = -1$, one finds

$$\tfrac{1}{2}(1+\beta)\Psi = \begin{pmatrix} \Phi \\ 0 \end{pmatrix}, \qquad \tfrac{1}{2}(1-\beta)\Psi = \begin{pmatrix} 0 \\ \chi \end{pmatrix}. \tag{XX.159}$$

Φ and χ are functions of $\boldsymbol{r}$ and of the spin component μ along the z axis; they may also be considered as functions of the radial variable r and of the "angular variables" (θ, φ, μ): in this they are entirely analogous to the wave functions of the Pauli theory.

Let us now suppose that Ψ is a simultaneous eigenfunction of $\boldsymbol{J}^2$, J_z and P. We denote the angular momentum quantum numbers by (JM). For convenience, we indicate the parity with the aid of the quantum number ϖ such that:

$$\varpi = \begin{cases} +1 \text{ for states of parity } (-)^{J+\frac{1}{2}} \\ -1 \text{ for states of parity } (-)^{J-\frac{1}{2}} \end{cases} \tag{XX.160}$$

Thus, by hypothesis:

$$\begin{gathered} \boldsymbol{J}^2 \begin{pmatrix} \Phi \\ \chi \end{pmatrix} = J(J+1) \begin{pmatrix} \Phi \\ \chi \end{pmatrix}, \qquad J_z \begin{pmatrix} \Phi \\ \chi \end{pmatrix} = M \begin{pmatrix} \Phi \\ \chi \end{pmatrix} \\ P^{(0)} \begin{pmatrix} \Phi \\ \chi \end{pmatrix} = (-)^{J+\frac{1}{2}\varpi} \begin{pmatrix} \Phi \\ -\chi \end{pmatrix}. \end{gathered} \tag{XX.161}$$

Let $\mathscr{Y}_{LJ}^{M}(\theta, \varphi, \mu)$ be the function of total angular momentum (JM) formed by the composition of a spin $\frac{1}{2}$ with the spherical harmonics of order L. The parity of this function is $(-)^L$. Also, according to the rules for the composition of angular momenta, L can take only the two values

$$L = l \equiv J + \tfrac{1}{2}\varpi \qquad L = l' \equiv J - \tfrac{1}{2}\varpi \tag{XX.162}$$

and the two functions $\mathscr{Y}_{lJ}^{M}$ and $\mathscr{Y}_{l'J}^{M}$ are of opposite parity, that of the first being $(-)^{J+\frac{1}{2}\varpi}$ and that of the second $(-)^{J-\frac{1}{2}\varpi}$. According to equations (XX.161), Φ is a function of $(r, \theta, \varphi, \mu)$ of angular momentum (JM) and of parity $(-)^{J+\frac{1}{2}\varpi}$; it is therefore necessarily equal to the product of a function of r by $\mathscr{Y}_{lJ}^{M}$. A similar argument shows that χ is equal to the product of a function of r by $\mathscr{Y}_{l'J}^{M}$.

In conclusion, if $\Psi_{\varpi J}^{M}$ represents a state of angular momentum (JM) and of parity $(-)^{J+\frac{1}{2}\varpi}$, it can be written in the form

$$\Psi_{\varpi J}^{M} = \frac{1}{r}\begin{pmatrix} F\,\mathscr{Y}_{lJ}^{M} \\ \mathrm{i}G\,\mathscr{Y}_{l'J}^{M} \end{pmatrix}, \tag{XX.163}$$

where l and l' are given by equations (XX.162), and F and G are arbitrary functions of r.

Consider now the eigenvalue problem

$$H_D \Psi_{\varpi J}^{M} = E \Psi_{\varpi J}^{M}. \tag{XX.164}$$

To solve this equation we separate the "angular" variables from the radial variables in the operator H_D. The method to be followed is wholly analogous to the one of Chapter IX.

We introduce the radial momentum p_r and the "radial velocity" α_r:

$$p_r \equiv -\,\mathrm{i}\,\frac{1}{r}\frac{\partial}{\partial r}\,r \tag{XX.165}$$

$$\alpha_r \equiv \boldsymbol{\alpha}\cdot\hat{\boldsymbol{r}} = \varrho_1\,(\boldsymbol{\sigma}\cdot\boldsymbol{r})/r. \tag{XX.166}$$

From identity (XIII.83), one obtains

$$\begin{aligned}(\boldsymbol{\alpha}\cdot\boldsymbol{r})(\boldsymbol{\alpha}\cdot\boldsymbol{p}) = (\boldsymbol{\sigma}\cdot\boldsymbol{r})(\boldsymbol{\sigma}\cdot\boldsymbol{p}) &= \boldsymbol{r}\cdot\boldsymbol{p}+\mathrm{i}\boldsymbol{\sigma}\cdot\boldsymbol{L} \\ &= rp_r+\mathrm{i}(1+\boldsymbol{\sigma}\cdot\boldsymbol{L}).\end{aligned}$$

Whence, multiplying on the left by α_r/r and using the obvious property $\alpha_r^2 = 1$, the identity:

$$\boldsymbol{\alpha}\cdot\boldsymbol{p} \equiv \alpha_r\left(p_r+\frac{\mathrm{i}}{r}\,(1+\boldsymbol{\sigma}\cdot\boldsymbol{L})\right). \tag{XX.167}$$

Let us examine the operator $1+\boldsymbol{\sigma}\cdot\boldsymbol{L}$. One easily shows that

$$1+\boldsymbol{\sigma}\cdot\boldsymbol{L} = \boldsymbol{J}^2+\tfrac{1}{4}-\boldsymbol{L}^2.$$

Further, from (XX.163), it is clear that the action of $\boldsymbol{L}^2$ on $\Psi_{\varpi J}^{M}$ is equivalent to that of the operator

$$(J+\tfrac{1}{2}\varpi\beta)(J+\tfrac{1}{2}\varpi\beta+1) \equiv J(J+1)+\tfrac{1}{4}+\tfrac{1}{2}\varpi\beta(2J+1).$$

Therefore:

$$(1+\boldsymbol{\sigma}\cdot\boldsymbol{L})\Psi_{\varpi J}^{M} = -\tfrac{1}{2}\varpi(2J+1)\beta\Psi_{\varpi J}^{M}. \qquad \text{(XX.168)}$$

Substituting relations (XX.167) and (XX.168) into equation (XX.164), one obtains

$$\left[\alpha_r\left(p_r - \frac{\mathrm{i}\varpi(J+\frac{1}{2})}{r}\,\beta\right) + m\beta + V(r)\right]\Psi_{\varpi J}^{M} = E\Psi_{\varpi J}^{M}.$$

By replacing the eigenfunction by expression (XX.163), the operators p_r and α_r by their definitions (XX.165) and (XX.166) and using the identites (cf. Problem XX.8):

$$\begin{aligned}(\boldsymbol{\sigma}\cdot\hat{\boldsymbol{r}})\,\mathscr{Y}_{lJ}^{M} &= -\mathscr{Y}_{l'J}^{M}\\ (\boldsymbol{\sigma}\cdot\hat{\boldsymbol{r}})\,\mathscr{Y}_{l'J}^{M} &= -\mathscr{Y}_{lJ}^{M}\end{aligned} \qquad \text{(XX.169)}$$

this equation leads to two coupled differential equations for the radial functions $F(r)$ and $G(r)$, namely

$$\left[-\frac{\mathrm{d}}{\mathrm{d}r}+\frac{\varpi(J+\frac{1}{2})}{r}\right]G = (E-m-V)\,F \qquad \text{(XX.170}a\text{)}$$

$$\left[\frac{\mathrm{d}}{\mathrm{d}r}+\frac{\varpi(J+\frac{1}{2})}{r}\right]F = (E+m-V)\,G. \qquad \text{(XX.170}b\text{)}$$

These equations here play the role of equation (IX.20) in the non-relativistic theory.

After integration over the angles, the norm of $\Psi_{\varpi J}^{M}$ is given by the expression:

$$\langle\Psi_{\varpi J}^{M}|\Psi_{\varpi J}^{M}\rangle = \int_0^\infty (|F|^2+|G|^2)\,\mathrm{d}r \qquad \text{(XX.171)}$$

to be compared with expression (IX.21).

The discussion of the regularity of F and G is in all ways analogous to the discussion of the regularity of $y_l(r)$ in the non-relativistic theory. We shall not go into the details here.

26. Free Spherical Waves

For $V=0$, the method of the preceding paragraph gives the stationary solutions of the Dirac equation of the free electron which

correspond to well-defined angular momentum and parity; these are the Dirac free spherical waves.

In this case, eq. (XX.170*b*) gives,

$$G = \frac{1}{E+m}\left[\frac{\mathrm{d}}{\mathrm{d}r} + \frac{\varpi(J+\frac{1}{2})}{r}\right] F. \qquad \text{(XX.172)}$$

Substituting this expression into (XX.170*a*), one finds:

$$\begin{aligned}(E^2-m^2)\,F &= \left[-\frac{\mathrm{d}}{\mathrm{d}r} + \frac{\varpi(J+\frac{1}{2})}{r}\right]\left[\frac{\mathrm{d}}{\mathrm{d}r} + \varpi\frac{(J+\frac{1}{2})}{r}\right] F \\ &= \left[-\frac{\mathrm{d}^2}{\mathrm{d}r^2} + \frac{(J+\frac{1}{2})\,(J+\varpi+\frac{1}{2})}{r^2}\right] F.\end{aligned}$$

It can easily be shown that

$$(J+\tfrac{1}{2})(J+\varpi+\tfrac{1}{2}) = l(l+1),$$

where l is the integer defined by equation (XX.162). The preceding equation is therefore identical with the free wave radial equation of the non-relativistic theory if one substitutes for (E^2-m^2) the product of $2m$ by the non-relativistic energy. It has one and only one regular solution for any positive value of (E^2-m^2). If one puts

$$k = \sqrt{E^2-m^2} \qquad (|E| \geqslant m)$$

it becomes

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}r^2} - \frac{l(l+1)}{r^2} + k^2\right] F = 0.$$

Its regular solution (defined up to a constant) is given by

$$F = r j_l(kr).$$

The corresponding G function is obtained by applying relation (XX.172). Using the recursion relations (B.42) and (B.43) [the first for $\varpi = 1$, the second for $\varpi = -1$, both being written with $\gamma = 0$], one finds

$$G = \frac{\varpi k}{E+m}\, r j_{l'}(kr).$$

In conclusion, for any value of the energy E situated outside the interval $(-m, +m)$, there exists a free spherical wave of angular

momentum (JM) and parity $(-)^{J+\frac{1}{2}\varpi}$. In the form (XX.163), it is written

$$Cst. \times \begin{pmatrix} |E+m|^{\frac{1}{2}}\, j_l(\sqrt{E^2-m^2}\, r)\, \mathcal{Y}_{lJ}^{M} \\ i\varpi\varepsilon |E-m|^{\frac{1}{2}}\, j_{l'}(\sqrt{E^2-m^2}\, r)\, \mathcal{Y}_{l'J}^{M} \end{pmatrix} \qquad \text{(XX.173)}$$

where: $\varepsilon = E/|E|$.

27. The Hydrogen Atom

As a second example, let us look for the bound states of an electron in the Coulomb field of an atomic nucleus. The nucleus will be treated as a point charge equal to $(-Z)$ times the charge of the electron, and fixed at the origin [1]). We must therefore find *the bound states* of a Dirac particle in the central potential

$$V = -\frac{Ze^2}{r}.$$

This eigenvalue problem can be exactly solved. Here we shall give only the broad outline of the method, which is a simple extension of the one set forth in § XI.4.

It is clear from an examination of the asymptotic behavior of the solutions of the set of radial equations (XX.170) that E must be contained in the interval $(-m, +m)$. The desired eigenvalues are those for which the solutions that are regular at the origin behave asymptotically like $\exp(-\sqrt{m^2-E^2}\, r)$.

Put:

$$\varkappa = \sqrt{m^2-E^2} \qquad \nu = \sqrt{\frac{m-E}{m+E}} \qquad \text{(XX.174)}$$

$$\zeta = Ze^2 \qquad \tau = \varpi(J+\tfrac{1}{2}) \qquad \text{(XX.175)}$$

and introduce the variable

$$\varrho \equiv \varkappa r. \qquad \text{(XX.176)}$$

The set (XX.170) is equivalent to

$$\left(-\frac{\mathrm{d}}{\mathrm{d}\varrho} + \frac{\tau}{\varrho}\right) G = \left(-\nu + \frac{\zeta}{\varrho}\right) F \qquad \text{(XX.177}a\text{)}$$

$$\left(\frac{\mathrm{d}}{\mathrm{d}\varrho} + \frac{\tau}{\varrho}\right) F = \left(\nu^{-1} + \frac{\zeta}{\varrho}\right) G. \qquad \text{(XX.177}b\text{)}$$

[1]) This supposes the nucleus infinitely heavy. The error thus made cannot be neglected, for it is of the order of magnitude of the relativistic effects. It is largely compensated for if the mass of the electron m is everywhere replaced by the reduced mass.

We look for solutions of the form

$$F(\varrho)=\varrho^s e^{-\varrho}(a_0+a_1\varrho+a_2\varrho^2+\ldots) \quad (a_0\neq 0) \qquad \text{(XX.178}a\text{)}$$

$$G(\varrho)=\varrho^s e^{-\varrho}(b_0+b_1\varrho+b_2\varrho^2+\ldots) \quad (b_0\neq 0). \qquad \text{(XX.178}b\text{)}$$

Substituting these expansions into equations (XX.177) and equating terms of successive orders one obtains a series of equations of which the first fixes s and the subsequent ones permit the determination of the coefficients $a_0, b_0, a_1, b_1, \ldots, a_n, b_n \ldots$ by recurrence. The equation in s has the two roots $\pm\sqrt{\tau^2-\zeta^2}$. A necessary and sufficient condition for F and G to fulfil the conditions of regularity at the origin $F(0)=G(0)=0$, is that $s>0$. Thus only the positive root is to be kept [1]):

$$s=\sqrt{\tau^2-\zeta^2}.$$

Thus for each value of E there is one solution that is regular at the origin. In general it behaves like $\varrho^s e^{\varrho}$ at infinity unless the two expansions (XX.178) have only a finite number of terms. This can only happen for certain particular values of E; these are the required energy levels. The calculation shows that they are given by the expression

$$m\left[1+\frac{\zeta^2}{(n'+s)^2}\right]^{-\frac{1}{2}},$$

where n', the radial quantum number, is the degree of the polynomials figuring in expressions (XX.178). For each positive value of n' there exists a regular solution for each of the two values of ϖ; for $n'=0$ there exists a regular solution for $\varpi=-1$, but no solution for $\varpi=+1$.

Let us introduce the principal quantum number

$$n=J+\tfrac{1}{2}+n'.$$

The preceding results may then be reformulated in the following way. The levels of the discrete spectrum depend on the two quantum numbers n and J according to the formula

$$E_{nJ}=m\left[1+\frac{Z^2e^4}{(n-\varepsilon_J)^2}\right]^{-\frac{1}{2}} \qquad \text{(XX.179)}$$

$$\varepsilon_J=J+\tfrac{1}{2}-\sqrt{(J+\tfrac{1}{2})^2-Z^2e^4}, \qquad \text{(XX.179')}$$

[1]) We suppose $\zeta<|\tau|$, i.e. $Ze^2<(J+\frac{1}{2})$. This condition is always fulfilled if $Z<137$, which is always the case in practice. If it were not, the discussion of the regularity conditions at the origin would be much more delicate.

where n can take all positive integral values and J all half-integral values in the interval $(0, n)$:

$$n = 1, 2, \ldots, \infty; \qquad J = \tfrac{1}{2}, \tfrac{3}{2}, \ldots, n - \tfrac{1}{2}.$$

To each value of J there correspond two series of $(2J+1)$ solutions of opposite parities, except for the value $J = n - \frac{1}{2}$ to which there corresponds a single series of $(2J+1)$ solutions of parity $(-)^{n-1}$. Rather than specify the parities, one may instead specify the values of l, the orbital angular momentum of the first two components of the spinor; recall that the parity of the spinor is $(-)^l$.

The spectroscopic notation nl_J is generally used to distinguish these different series of solutions one from another. The following table lists the first few levels in increasing order, with the corresponding spectroscopic terms [each term has a degeneracy of order $2J+1$]:

$n=1$	$J=\frac{1}{2}$	$1s_{1/2}$			$(n'=0)$
$n=2$	$J=\frac{1}{2}$	$2s_{1/2}$	$2p_{1/2}$		$(n'=1)$
	$J=\frac{3}{2}$		$2p_{3/2}$		$(n'=0)$
$n=3$	$J=\frac{1}{2}$	$3s_{1/2}$	$3p_{1/2}$		$(n'=2)$
	$J=\frac{3}{2}$		$3p_{3/2}$	$3d_{3/2}$	$(n'=1)$
	$J=\frac{5}{2}$			$3d_{5/2}$	$(n'=0)$.

If expression (XX.179) is expanded into a power series in Z^2e^4 one finds

$$E_{nJ} = m\left[1 - \frac{Z^2e^4}{2n^2} - \frac{(Z^2e^4)^2}{2n^4}\left(\frac{n}{J+\frac{1}{2}} - \frac{3}{4}\right) + \ldots\right].$$

The first term is the mass term. The second, $-Z^2e^4/2n^2$, is exactly equal to the quantity given by the non-relativistic theory. The third and following terms give the relativistic corrections. These corrections partially remove the "accidental degeneracy" of the non-relativistic levels: for n fixed, the binding energy $m - E$ of each term is slightly increased; the increase depends only on J, and is larger for smaller J.

The experimental results on the fine structure of the hydrogen atom and hydrogen-like atoms (notably He^+) are in broad agreement with these predictions.

However, the agreement is not perfect. The largest discrepancy is observed in the fine structure of the $n=2$ levels of the hydrogen atom [1]). In the non-relativistic approximation, the three levels $2s_{1/2}$,

[1]) W. E. Lamb and R. C. Retherford, Phys. Rev. 72 (1947) 241.

$2p_{1/2}$ and $2p_{3/2}$ are equal. In the Dirac theory, the levels $2s_{1/2}$ and $2p_{1/2}$ are still equal, while the $2p_{3/2}$ level is slightly lower (the separation is of the order of 10^{-4}eV). The level distance $2p_{3/2} - 2p_{1/2}$ agrees with the theory but the level $2s_{1/2}$ is lower than the level $2p_{1/2}$ and the distance $2s_{1/2} - 2p_{1/2}$ is equal to about a tenth of the distance $2p_{3/2} - 2p_{1/2}$. This effect is known as the Lamb shift. To explain it, we need a rigorous treatment of the interaction between the electron, the proton and the quantized electromagnetic field; in the Dirac theory one retains only the Coulomb potential which is the main term in that interaction; the Lamb shift represents "radiative corrections" to this approximation [1]).

V. NON-RELATIVISTIC LIMIT OF THE DIRAC EQUATION

28. Large and Small Components

Consider the positive energy plane waves whose components are given in Table XX.3. Let us suppose that the energy E_p differs little from the rest energy:

$$W \equiv E_p - m \ll m.$$

The non-relativistic approximation is then valid, for the kinetic energy W is nearly equal to $\frac{1}{2}mv^2$ and we have

$$\frac{W}{m} \simeq \tfrac{1}{2}v^2 \ll 1.$$

It will be seen that the non-vanishing component corresponding to $\beta = +1$ is then much larger than the one corresponding to $\beta = -1$:

$$\sigma_z = +1 \qquad \frac{u_3}{u_1} = \left(\frac{W}{W+2m}\right)^{\frac{1}{2}} \simeq O\left(\frac{v}{c}\right) \ll 1$$

$$\sigma_z = -1 \qquad \frac{u_4}{u_2} = -\left(\frac{W}{W+2m}\right)^{\frac{1}{2}} \simeq O\left(\frac{v}{c}\right) \ll 1.$$

A similar observation applies for the free spherical waves [cf. expression (XX.173)] or for the eigenfunctions of the hydrogen atom (cf.

[1]) The most recent measurements give 1057.77 ± 0.70 Mc/s (Mc/s ≡ Megacycle per second) for the separation $2s_{1/2} - 2p_{1/2}$; the theoretical value obtained when the "radiative corrections" predicted by Quantum Electrodynamics are taken into account is 1057.99 ± 0.2 Mc/s [C. M. Sommerfield, Phys. Rev. 107 (1957) 328].

notably the particular eigensolutions defined in Problem XX.10). This suggests that in the non-relativistic approximation two of the components of the spinor Ψ, the components Ψ_3 and Ψ_4 corresponding to the eigenvalue -1 of β, are very small and may be neglected, and that the Dirac theory is then equivalent to a two-component theory.

In order to put this point properly in evidence, we write the Dirac spinor Ψ in the form (XX.157), Φ and χ being defined by equations (XX.158) or (XX.159). As has already been noted in § 25, Φ and χ may be regarded as vectors of the state-vector space of the two component non-relativistic theory.

With these notations the Dirac equation relative to a stationary state of energy E is written, in the Dirac form and representation:

$$(\boldsymbol{\sigma}\cdot(\boldsymbol{p}-e\boldsymbol{A}))\,\chi+(e\varphi+m)\,\Phi = E\Phi \qquad \text{(XX.180}a\text{)}$$

$$(\boldsymbol{\sigma}\cdot(\boldsymbol{p}-e\boldsymbol{A}))\,\Phi+(e\varphi-m)\,\chi = E\chi. \qquad \text{(XX.180}b\text{)}$$

Let us put:

$$\begin{gathered}\boldsymbol{\pi} = \boldsymbol{p}-e\boldsymbol{A}, \qquad M = E-e\varphi \\ W = E-m, \qquad M' = \tfrac{1}{2}(m+M) = m+\tfrac{1}{2}(W-e\varphi).\end{gathered} \qquad \text{(XX.181)}$$

Solving equation (XX.180b) for χ and then substituting into equation (XX.180a), we obtain

$$\chi = \frac{1}{2M'}\,(\boldsymbol{\sigma}\cdot\boldsymbol{\pi})\,\Phi \qquad \text{(XX.182)}$$

$$\left[(\boldsymbol{\sigma}\cdot\boldsymbol{\pi})\,\frac{1}{2M'}\,(\boldsymbol{\sigma}\cdot\boldsymbol{\pi})+e\varphi\right]\Phi = W\Phi. \qquad \text{(XX.183)}$$

The set of equations (XX.182–183) is *exactly equivalent* to the Dirac equation.

In the non-relativistic limit

$$W,\ e\varphi,\ \boldsymbol{p},\ e\boldsymbol{A} \ll m, \qquad M' \simeq m. \qquad \text{(XX.184)}$$

It is clear from equation (XX.182) that

$$\chi \ll \Phi$$

and that the ratio of these two quantities is of the order of p/m, i.e. v/c. χ and Φ are known as the *small and large* components respectively.

In the rest of this section, we shall make use of the concept of "*even*" and "*odd*" operators. By definition:

(*i*) an operator $\mathscr{P}$ is "even" if it has no matrix element linking small and large components (examples: $\boldsymbol{p}$, $\boldsymbol{r}$, $\boldsymbol{L}$, $\boldsymbol{\sigma}$, $\boldsymbol{J}$, $P(r_0)$, β);

(*ii*) an operator $\mathscr{I}$ is "odd" if its non-vanishing matrix elements link small and large components (examples: $\boldsymbol{\alpha}$, $\beta\boldsymbol{\alpha}$, γ^5, $\boldsymbol{j}(r_0)$).

It is equivalent to say that $\mathscr{P}$ is an operator that commutes with β, $\mathscr{I}$ an operator that anticommutes with β:

$$\mathscr{P} = \beta\mathscr{P}\beta, \qquad \mathscr{I} = -\beta\mathscr{I}\beta. \tag{XX.185}$$

Any operator Q is the sum of an "even" and an "odd" operator, and furthermore this decomposition is unique:

$$Q = \tfrac{1}{2}[Q+\beta Q\beta]+\tfrac{1}{2}[Q-\beta Q\beta].$$

The product of two "even" or of two "odd" operators is an "even" operator; the product of an "even" operator by an "odd" operator is an "odd" operator.

29. The Pauli Theory as the Non-relativistic Limit of the Dirac Theory

We now return to the system of equations (XX.182–183). In neglecting the small components, one makes an error of order v^2/c^2 in the normalization of the wave function. An error of the same order is made in replacing the operator M' in equation (XX.183) by the mass m. In this approximation equation (XX.183) takes the form of an eigenvalue equation [1]):

$$H_{n.r.}\Phi = W\Phi \tag{XX.186}$$

of a certain Hamiltonian

$$H_{n.r.} \equiv \frac{1}{2m}(\boldsymbol{\sigma}\cdot\boldsymbol{\pi})(\boldsymbol{\sigma}\cdot\boldsymbol{\pi}) + e\varphi \tag{XX.187}$$

acting on the two-component wave function Φ. Equation (XX.186) defines the energy W to within v^2/c^2.

In order to put $H_{n.r.}$ in a more familiar form we apply identity (XIII.83), noting that the components of $\boldsymbol{\pi}$ do not commute, and that therefore:

$$\boldsymbol{\pi}\times\boldsymbol{\pi} = ie\ \mathrm{curl}\ \boldsymbol{A} = ie\mathscr{H}.$$

One then obtains

$$H_{n.r.} \equiv \frac{1}{2m}(\boldsymbol{p}-e\boldsymbol{A})^2 - \frac{e}{2m}(\boldsymbol{\sigma}\cdot\mathscr{H}) + e\varphi. \tag{XX.188}$$

---

[1]) Equation (XX.183) is not a true eigenvalue equation since the operator in brackets on the left-hand side depends on the "eigenvalue" W through M'.

This will be recognized as the Hamiltonian of the Pauli theory corresponding to a particle of mass m, charge e and intrinsic angular momentum:

$$\mu = \mu_B \sigma \qquad (\mu_B \equiv \text{Bohr magneton} \equiv e/2m).$$

Not only does the Dirac theory predict the existence of an intrinsic magnetic moment for the electron, but it gives its correct value (§ XIII.18). This is one of the major successes of the theory [1]).

In order to prove that the Dirac theory in the approximation considered here is equivalent to the two-component Pauli theory, we must be able to find operators corresponding to the operators in the Dirac theory but acting only on the large components.

This can actually be done provided that the said operators enter into the calculations only through their matrix elements between states Ψ', Ψ'' whose energy is positive and sufficiently close to that of the rest mass for the non-relativistic approximation to be valid.

If we are concerned with an even operator $\mathscr{P}$, the matrix element $\langle \Psi'' | \mathscr{P} | \Psi' \rangle$ takes the form

$$\langle \Phi'' | \mathscr{P} | \Phi' \rangle + \langle \chi'' | \mathscr{P} | \chi' \rangle.$$

The second term is $(v/c)^2$ times smaller than the first, and may be neglected in the approximation discussed here. $\mathscr{P}$ may then be replaced by its projection on the space of the large components: this projection represents in the non-relativistic Pauli theory the physical quantity represented by $\mathscr{P}$ in the Dirac theory.

---

[1]) In actual fact, the experimental value μ_{exp} differs slightly from this theoretical value for the magnetic moment of the electron [P. Kusch and H. M. Foley, Phys. Rev. 72 (1947) 1256]. More recent measurements give

$$\Delta\mu_{exp} \equiv \mu_{exp} - \mu_B = (1.165 \pm 0.011) \times 10^{-3}\, \mu_B.$$

The "radiative corrections" of Quantum Electrodynamics account for the existence of this anomalous magnetic moment; they give (Sommerfield, *loc. cit.*):

$$\Delta\mu_{th} = 1.163 \times 10^{-3}\, \mu_B.$$

Note in passing that the equation obtained by adding the term $\varkappa \mu_B \sigma_{\mu\nu} F^{\mu\nu}$ to the Dirac operator has all the invariance properties of the Dirac equation, and this for any given value of the numerical constant $\varkappa$; such an equation describes a particle of mass m, of charge e and of intrinsic magnetic moment $(1 + \varkappa)\, \mu_B$.

If we are concerned with an odd operator, $\mathscr{I}$, one has

$$\langle \Psi'' | \mathscr{I} | \Psi' \rangle = \langle \Phi'' | \mathscr{I} | \chi' \rangle + \langle \chi'' | \mathscr{I} | \Phi' \rangle.$$

Here the small components are involved explicitly in each term on the right-hand side. In the approximation in question however equation (XX.182) gives

$$|\chi\rangle = \varrho_1 \frac{\boldsymbol{\sigma}\cdot\boldsymbol{\pi}}{2m} |\Phi\rangle$$

from which

$$\langle \Psi'' | \mathscr{I} | \Psi' \rangle = \frac{1}{2m} \langle \Phi'' | [\mathscr{I} \varrho_1 (\boldsymbol{\sigma}\cdot\boldsymbol{\pi}) + (\boldsymbol{\sigma}\cdot\boldsymbol{\pi}) \varrho_1 \mathscr{I}] | \Phi' \rangle.$$

$\mathscr{I}$ may therefore be replaced by the projection on the space of the large components of the operator

$$\frac{1}{2m} [\mathscr{I} \varrho_1 (\boldsymbol{\sigma}\cdot\boldsymbol{\pi}) + (\boldsymbol{\sigma}\cdot\boldsymbol{\pi}) \varrho_1 \mathscr{I}].$$

Thus the "velocity" $\boldsymbol{\alpha} = \varrho_1 \boldsymbol{\sigma}$ may be replaced by the operator in the space of the large components

$$\frac{\boldsymbol{\pi}}{m} \equiv \frac{\boldsymbol{\sigma}(\boldsymbol{\sigma}\cdot\boldsymbol{\pi}) + (\boldsymbol{\sigma}\cdot\boldsymbol{\pi})\boldsymbol{\sigma}}{2m}.$$

Similarly, the current density at the point $\boldsymbol{r}_0$:

$$\boldsymbol{j}(\boldsymbol{r}_0) \equiv \varrho_1 \boldsymbol{\sigma} \delta(\boldsymbol{r} - \boldsymbol{r}_0)$$

may be replaced by the operator in the space of the large components:

$$(\boldsymbol{j}(\boldsymbol{r}_0))_{n.r.} \equiv \delta(\boldsymbol{r} - \boldsymbol{r}_0)\, \boldsymbol{\sigma} \left(\frac{\boldsymbol{\sigma}\cdot\boldsymbol{\pi}}{2m}\right) + \left(\frac{\boldsymbol{\sigma}\cdot\boldsymbol{\pi}}{2m}\right) \boldsymbol{\sigma}\, \delta(\boldsymbol{r} - \boldsymbol{r}_0)$$

or again, applying identity (XIII.83):

$$(\boldsymbol{j}(\boldsymbol{r}_0))_{n.r.} \equiv \boldsymbol{j}^{(\mathrm{I})} + \boldsymbol{j}^{(\mathrm{II})} \tag{XX.189}$$

$$\boldsymbol{j}^{(\mathrm{I})} \equiv \frac{\delta(\boldsymbol{r} - \boldsymbol{r}_0)\, \boldsymbol{\pi} + \boldsymbol{\pi}\, \delta(\boldsymbol{r} - \boldsymbol{r}_0)}{2m} \tag{XX.190a}$$

$$\boldsymbol{j}^{(\mathrm{II})} \equiv \mathrm{i} \frac{\delta(\boldsymbol{r} - \boldsymbol{r}_0)\, (\boldsymbol{p} \times \boldsymbol{\sigma}) - (\boldsymbol{p} \times \boldsymbol{\sigma})\, \delta(\boldsymbol{r} - \boldsymbol{r}_0)}{2m}. \tag{XX.190b}$$

In this limit, the Dirac electric current $e\boldsymbol{j}$ is therefore made up of two terms. The first, $e\boldsymbol{j}^{(\mathrm{I})}$, is identical with the current of the Schrödinger theory (cf. Problem IV.1). In order to interpret the second we consider

its matrix element between the large components Φ' and Φ''. The calculation shows that

$$\langle \Phi'' | e\boldsymbol{j}^{(II)} | \Phi' \rangle = \frac{e}{2m} \operatorname{curl} \langle \Phi'' | \delta(\boldsymbol{r} - \boldsymbol{r}_0)\, \boldsymbol{\sigma} | \Phi' \rangle .$$

This is a magnetic current term, and the quantity

$$\frac{e}{2m} \langle \Phi'' | \delta(\boldsymbol{r} - \boldsymbol{r}_0)\, \boldsymbol{\sigma} | \Phi' \rangle \equiv \langle \Phi'' | \delta(\boldsymbol{r} - \boldsymbol{r}_0)\, \boldsymbol{\mu} | \Phi' \rangle$$

may be interpreted as a magnetic moment density. It will be seen that the divergence of this magnetic current vanishes, and therefore that it gives no contribution to the equation of continuity.

30. Application: Hyperfine Structure and Dipole-Dipole Coupling

We now consider an electron in the electric field of an atom, described by a certain electrostatic potential $\varphi(r)$, and examine the effect of the field created by the magnetic moment $\boldsymbol{M}$ of the nucleus. The field created by a magnetic dipole $\boldsymbol{M}$ situated at the origin of coordinates may be represented by the vector potential:

$$\boldsymbol{A} \equiv \frac{\boldsymbol{M} \times \boldsymbol{r}}{r^3} \tag{XX.191}$$

$$\equiv \operatorname{curl} (\boldsymbol{M}/r). \tag{XX.191'}$$

The presence of this field leads to an additional term $-e\boldsymbol{\alpha} \cdot \boldsymbol{A}$ in the Dirac Hamiltonian.

To determine the effect of this field in the non-relativistic approximation we can calculate the non-relativistic limit of the operator $-e\boldsymbol{\alpha} \cdot \boldsymbol{A}$ by the method of the preceding paragraph. We can equally well directly examine the modifications of the Pauli Hamiltonian (XX.188) due to the presence of $\boldsymbol{M}$. These two methods are equivalent. We shall adopt the second one here.

If we retain only terms linear in $\boldsymbol{M}$, the Pauli Hamiltonian contains the two supplementary terms:

$$I_a = -\frac{e}{2m} (\boldsymbol{p} \cdot \boldsymbol{A} + \boldsymbol{A} \cdot \boldsymbol{p})$$

$$I_b = -\frac{e}{2m} (\boldsymbol{\sigma} \cdot \mathscr{H}) = -\boldsymbol{\mu} \cdot \mathscr{H} ;$$

where $\mathscr{H}$ is the field created by the dipole $\boldsymbol{M}$.

I_a is a spin-orbit coupling term (spin of the nucleus, orbit of the electron). Clearly, [cf. eq. (XX.191')] div $\boldsymbol{A}=0$ whence:

$$I_a = -\frac{e}{m}\boldsymbol{A}\cdot\boldsymbol{p}$$

Substituting expression (XX.191) onto the right-hand side, and introducing the orbital angular momentum of the electron $\boldsymbol{L} \equiv \boldsymbol{r}\times\boldsymbol{p}$, one finds:

$$I_a = -\frac{e\boldsymbol{M}\cdot\boldsymbol{L}}{mr^3}. \tag{XX.192}$$

I_b is the spin-spin or dipole-dipole coupling term. It can be calculated with the aid of (XX.191'):

$$\begin{aligned} I_b &= -\boldsymbol{\mu}\cdot(\nabla\times\boldsymbol{A}) = -\boldsymbol{\mu}\cdot\left[\nabla\times\left(\nabla\times\frac{\boldsymbol{M}}{r}\right)\right] \\ &= (\boldsymbol{\mu}\cdot\boldsymbol{M})\,\triangle\left(\frac{1}{r}\right) - [(\boldsymbol{\mu}\cdot\nabla)\,(\boldsymbol{M}\cdot\nabla)]\left(\frac{1}{r}\right). \end{aligned} \tag{XX.193}$$

When $r\neq 0$, I_b can easily be calculated by performing the differentiation, which gives:

$$-\frac{3(\boldsymbol{M}\cdot\boldsymbol{r})\,(\boldsymbol{\mu}\cdot\boldsymbol{r}) - (\boldsymbol{M}\cdot\boldsymbol{\mu})\,r^2}{r^5}.$$

Considered as a simple function, expression (XX.193) has a singularity in $1/r^3$ at the origin. To determine the action of the operator I_b, it is convenient to examine the result of integration of the product of this quantity by a regular function $f(\boldsymbol{r})$ in a small domain about the point $r=0$. To this effect, we write I_b in the form

$$I_b = \tfrac{2}{3}\,(\boldsymbol{\mu}\cdot\boldsymbol{M})\,\triangle\left(\frac{1}{r}\right) - [(\boldsymbol{\mu}\cdot\nabla)\,(\boldsymbol{M}\cdot\nabla) - \tfrac{1}{3}\,(\boldsymbol{\mu}\cdot\boldsymbol{M})\,\triangle]\left(\frac{1}{r}\right). \tag{XX.193'}$$

The second term in this expression is a second order tensor operator in the space of functions of $\boldsymbol{r}$; if, to effect the integration mentioned above, $f(\boldsymbol{r})$ is expanded into spherical harmonics. Only coefficients of spherical harmonics of order 2 will contribute to the integration over the angles; these vanish at the origin at least as rapidly as r^2; the contribution of the second term of (XX.193') also vanishes at the origin in spite of the singularity in $1/r^3$. With the aid of identity

(A.12), the first term may be put in the form $-(\frac{8}{3}\pi)\,(\boldsymbol{\mu}\cdot\boldsymbol{M})\delta(\boldsymbol{r})$. Thus, for any $\boldsymbol{r}$, including the origin:

$$I_b = -\frac{8\pi}{3}\,(\boldsymbol{M}\cdot\boldsymbol{\mu})\,\delta(\boldsymbol{r}) - \frac{1}{r^3}\left[3\left(\boldsymbol{M}\cdot\frac{\boldsymbol{r}}{r}\right)\left(\boldsymbol{\mu}\cdot\frac{\boldsymbol{r}}{r}\right) - (\boldsymbol{M}\cdot\boldsymbol{\mu})\right]. \quad \text{(XX.194)}$$

Expressions (XX.192) and (XX.194) are valid in the non-relativistic imit and allow the determination of the hyperfine structure of atomic levels to within v^2/c^2. In particular, the contribution of the s-electrons to the hyperfine structure is given by the contact term $-(\frac{8}{3}\pi)\,(\boldsymbol{\mu}\cdot\boldsymbol{M})\delta(\boldsymbol{r})$.

31. Higher-order Corrections and the Foldy-Wouthuysen Transformation

To the lowest order in v/c, the Dirac theory is equivalent to the two-component Pauli theory. It is possible to obtain higher-order relativistic corrections by starting as before from equations (XX.182–183). To this effect, one replaces $1/M'$ by its expansion in powers of $[(W-e\varphi)/2m]$:

$$\frac{1}{M'} = \frac{1}{m}\left[1 - \frac{W-e\varphi}{2m} + \left(\frac{W-e\varphi}{2m}\right)^2 - \ldots\right].$$

Since:

$$\left\langle\frac{W-e\varphi}{2m}\right\rangle \simeq \left\langle\frac{\boldsymbol{\pi}^2}{4m^2}\right\rangle \simeq O\left(\frac{v^2}{c^2}\right),$$

this is essentially a power series expansion in v^2/c^2. If this expansion is stopped at the first term one obtains the Pauli theory, as in § 29. Higher-order relativistic corrections are obtained by taking the expansion beyond the first term. However, once corrections of the order of v^2/c^2 are taken into account, the Dirac theory in the form (XX.182–183) ceases to be formally equivalent to a two-component theory. This is because:

(*i*) the contribution from the small components can no longer be neglected, neither in the normalization, nor in the calculation of matrix elements of even operators;

(*ii*) equation (XX.183) is properly speaking no longer an eigenvalue equation (cf. note, p. 935).

Although the method is not thereby absolutely condemned, its application and the interpretation of the results becomes rather delicate. Foldy and Wouthuysen [1]) have proposed another method which allows one to approximate to the Dirac theory by a two-component theory to any given order in v/c. It essentially consists in effecting a suitably chosen unitary transformation on the wave functions and operators of the Dirac theory. In the new "representation" – which we shall call the FW "representation" – the Dirac Hamiltonian is an "even" operator to the given order in v/c, so that in this approximation the small and the large components are completely decoupled in the wave equation. One may therefore simply ignore the small components, thereby obtaining the desired two-component theory. The operators of the two-component theory are then obtained from the "even" operators of the FW "representation" and not from the operators of the old "representation". One is thus led to a new interpretation of the operators of non-relativistic mechanics, notably of the position operator, an interpretation in many ways more satisfying than the old one.

The rest of this section is devoted to the FW transformation and to its application to the non-relativistic approximation of the Dirac equation.

32. FW Transformation for a Free Particle

In the case of a free particle, the small and large components can be completely decoupled in all orders of v/c.

We consider the Dirac Hamiltonian

$$H_0 \equiv \boldsymbol{\alpha}\cdot\boldsymbol{p}+\beta m.$$

Let Γ_+ and Γ_- be the projectors onto the positive and negative energy solutions respectively:

$$\Gamma_\pm \equiv \frac{1}{2}\left[1 \pm \frac{H_0}{E_p}\right] = \frac{1}{2}\left[1 \pm \frac{\boldsymbol{\alpha}\cdot\boldsymbol{p}+\beta m}{E_p}\right]. \qquad \text{(XX.195)}$$

$$E_p \equiv \sqrt{m^2+\boldsymbol{p}^2}.$$

---

[1]) L. L. Foldy and S. A. Wouthuysen, Phys. Rev. **78** (1958) 29.

Denote by B_+ and B_- the projectors onto the spaces of the large and small components respectively:

$$B_\pm \equiv \tfrac{1}{2}(1 \pm \beta).$$

By definition, the operator U, which takes us over to the FW "representation", transforms Γ_+ into B_+ and Γ_- into B_-. With primes denoting the vectors and operators of the FW "representation", we therefore have:

$$U^\dagger U = UU^\dagger = 1$$
$$\Gamma'_\pm \equiv U\Gamma_\pm U^\dagger = B_\pm.$$

We also require that U be invariant under translation, rotation and reflection. It is left to the reader to show that U is then defined to within a phase factor. Fixing this phase one obtains

$$U = \sqrt{\frac{2E_p}{m+E_p}}\,\frac{1}{2}\left[1+\beta\frac{H_0}{E_p}\right] \qquad \text{(XX 196)}$$

$$= \sqrt{\frac{m+E_p}{2E_p}} + \beta\frac{\boldsymbol{\alpha}\cdot\boldsymbol{p}}{\sqrt{2E_p(m+E_p)}}. \qquad \text{(XX.196')}$$

It will easily be verified that this expression has all the desired properties.

Since U is time-independent, the Hamiltonian H_F which governs the evolution of the states in the FW "representation" is given by the equation

$$H_F = UH_DU^\dagger,$$

which gives, with the aid of (XX.196),

$$H_F = \beta E_p \equiv \beta(m^2+\boldsymbol{p}^2)^{\frac{1}{2}}. \qquad \text{(XX.197)}$$

Since H_F is an even operator, the large components Φ' and the small components χ' are completely decoupled in the equation of motion:

$$\mathrm{i}\frac{\partial\Phi'}{\partial t} = E_p\,\Phi' \qquad \text{(XX.198}a\text{)}$$

$$\mathrm{i}\frac{\partial\chi'}{\partial t} = -E_p\,\chi'. \qquad \text{(XX.198}b\text{)}$$

If we limit ourselves to positive energy solutions – *a fortiori* to non-

relativistic energies — the Dirac theory is exactly equivalent to the two-component theory represented by equation (XX.198a) to all orders in v/c.

The operator U commutes with $\boldsymbol{p}$, $\boldsymbol{J}$ and the parity operator P, but not with $\boldsymbol{r}$. In a representation with $\boldsymbol{r}$ diagonal, it is an integral operator with matrix element:

$$\langle \boldsymbol{r}|U|\boldsymbol{r}'\rangle = \int\int \langle \boldsymbol{r}|\boldsymbol{p}\rangle \, \mathrm{d}\boldsymbol{p} \, \langle \boldsymbol{p}|U|\boldsymbol{p}'\rangle \, \mathrm{d}\boldsymbol{p}' \, \langle \boldsymbol{p}'|\boldsymbol{r}'\rangle,$$

whence, with use of (XX.196′),

$$\langle \boldsymbol{r}|U|\boldsymbol{r}'\rangle = (2\pi)^{-3} \int \left[\sqrt{\frac{m+E_p}{2E_p}} + \beta \frac{\boldsymbol{\alpha}\cdot\boldsymbol{p}}{\sqrt{2E_p(m+E_p)}} \right] \mathrm{e}^{\mathrm{i}\boldsymbol{p}\cdot(\boldsymbol{r}-\boldsymbol{r}')} \, \mathrm{d}\boldsymbol{p}$$

$\langle \boldsymbol{r}|U|\boldsymbol{r}'\rangle$ is a function of $(\boldsymbol{r}-\boldsymbol{r}')$ that practically vanishes for $|\boldsymbol{r}-\boldsymbol{r}'| \gg 1/m$, but which takes non-negligible values when $|\boldsymbol{r}-\boldsymbol{r}'|$ is smaller or of the order of $1/m$. The FW transformation is therefore a non-local transformation in which the spinor $\Psi'(\boldsymbol{r})$, the transform of $\Psi(\boldsymbol{r})$, is obtained by taking a certain average over the values taken by Ψ in a volume about $\boldsymbol{r}$ whose linear dimensions are of the order of $1/m$, the Compton wavelength of the particle.

The position of the particle is represented in the FW "representation" by the operator

$$\boldsymbol{r}' \equiv U\boldsymbol{r}U^{\dagger}.$$

This operator is different to $\boldsymbol{r}$. Following Foldy and Wouthuysen, we shall call the quantity represented by $\boldsymbol{r}$ in the FW "representation" the *average position*. In the ordinary "representation", it is represented by a certain operator $\boldsymbol{R}$ and one has $\boldsymbol{R}' \equiv \boldsymbol{r}$, therefore:

$$\boldsymbol{R} \equiv U^{\dagger}\boldsymbol{r}U.$$

In the Dirac representation, $\boldsymbol{R}$ is a non-local operator whose action on the spinor $\Psi(\boldsymbol{r})$ consists, roughly, in multiplying by $\boldsymbol{r}$ and replacing the value at each point by a certain average over the values of the spinor in a domain of order $1/m$ about the point, whence the name of average position given above.

If Q' is an "even" operator of the FW "representation", the corresponding observable $Q_{n.r.}$ in the above-defined two-component theory is obtained by keeping only the matrix elements of Q' between vectors of the space of the large components: $Q_{n.r.} \equiv B_+Q'B_+$. In particular, the observable $\boldsymbol{r}$ representing the position in the two-

component theory corresponds to the "average position" $\boldsymbol{R}$ and not to the position operator of the Dirac theory proper [1]).

33. FW Transformation for a Particle in a Field

In the presence of a field the Dirac Hamiltonian takes the form:

$$H = \beta m + \mathscr{I} + \mathscr{P}$$
$$\mathscr{I} \equiv \boldsymbol{\alpha}\cdot\boldsymbol{\pi} = \boldsymbol{\alpha}\cdot(\boldsymbol{p} - e\boldsymbol{A}), \qquad \mathscr{P} \equiv e\varphi.$$

In general, there exists no "representation" in which the Hamiltonian is exactly "even" but by applying successive unitary transformations one may obtain "representations" in which the respective Hamiltonians have an "odd" part of higher and higher order in v/c.

To see this, let us make the unitary transformation

$$U = \exp(\beta\mathscr{I}/2m).$$

The Hamiltonian H_1 which governs the evolution of states in the new representation is given by the equation

$$H_1 = UHU^\dagger - \mathrm{i}U\,\partial U^\dagger/\partial t.$$

By making use of the fact that $\beta\mathscr{I}$ anticommutes with $(\beta m + \mathscr{I})$ and that $U^\dagger = \exp(-\beta\mathscr{I}/2m)$, one obtains:

$$\begin{aligned} U(\beta m+\mathscr{I})U^\dagger &= U^2(\beta m+\mathscr{I}) \\ &= \beta m[\cos(\mathscr{I}/m)+(\mathscr{I}/m)\sin(\mathscr{I}/m)] \\ &\quad + m[(\mathscr{I}/m)\cos(\mathscr{I}/m) - \sin(\mathscr{I}/m)]. \end{aligned}$$

The terms $U\mathscr{P}U^\dagger$ and $\mathrm{i}U\,\partial U^\dagger/\partial t$ may be expanded into a power series

---

[1]) In accordance with the interpretation given here, the orbital angular momentum $\boldsymbol{r}\times\boldsymbol{p}$ and the spin $\boldsymbol{\sigma}$ of the two-component theory do not correspond to the orbital angular momentum and spin of the Dirac theory, but to the "average angular momentum" $\boldsymbol{R}\times\boldsymbol{p}$ and the "average spin" $\boldsymbol{\Sigma}$. Here $\boldsymbol{\Sigma}$ is the operator whose correspondent in the FW "representation" is $\boldsymbol{\sigma}$; $\boldsymbol{\Sigma}' \equiv \boldsymbol{\sigma}$. The reader may verify that each of the components of the average spin and each of the components of the average angular momentum commutes with the free particle Hamiltonian; the spin and the orbital angular momentum proper do not have this property. Note also that:

$$\boldsymbol{J} \equiv (\boldsymbol{r}\times\boldsymbol{p}) + \tfrac{1}{2}\boldsymbol{\sigma} = (\boldsymbol{R}\times\boldsymbol{p}) + \tfrac{1}{2}\boldsymbol{\Sigma}.$$

in $\mathscr{J}/m$ by using the following operator identity, valid for any two operators A and B (cf. Problem VIII.4):

$$e^{A} B e^{-A} = B + [A, B] + \frac{1}{2}[A, [A, B]] + \ldots + \frac{1}{n!}\overbrace{[A, [A, \ldots [A, [A, B]] \ldots]]}^{n \text{ brackets}} + \ldots$$

We shall only give the result of the calculation of H' when $\mathscr{J}$ is time-independent. Since we are concerned with a first approximation, we put: $H' = H_1$. One finds:

$$H_1 = \beta m + \mathscr{P}_1 + \mathscr{J}_1$$

$$\mathscr{P}_1 = \mathscr{P} + \beta \frac{\mathscr{J}^2}{2m} - \frac{1}{8} m \left[\frac{\mathscr{J}}{m}, \left[\frac{\mathscr{J}}{m}, \frac{\mathscr{P}}{m}\right]\right] - \frac{1}{8} \beta m \left(\frac{\mathscr{J}}{m}\right)^4 + \ldots$$

$$\mathscr{J}_1 = m\left(\frac{1}{2} \beta \left[\frac{\mathscr{J}}{m}, \frac{\mathscr{P}}{m}\right] - \frac{1}{3}\left(\frac{\mathscr{J}}{m}\right)^3\right) + \ldots$$

The terms given in these expansions of the "even" and "odd" parts of H_1 allow the determination of $\mathscr{P}_1$ to within $(\mathscr{J}/m)^6$ or $(\mathscr{P}/m)\,(\mathscr{J}/m)^4$ whichever is the larger and the determination of $\mathscr{J}_1$ to within $(\mathscr{J}/m)^5$ or $(\mathscr{P}/m)\,(\mathscr{J}/m)^3$ whichever is the larger. The "odd" part of H_1 is therefore smaller than that of H by a factor of the order of the larger of $\mathscr{P}/m$ or $(\mathscr{J}/m)^2$; in the non-relativistic limit, $\mathscr{P}/m$ and $\mathscr{J}/m$ are of the order of $(v/c)^2$ and v/c respectively: $\mathscr{J}_1$ is therefore of the order of $(v/c)^3$.

We now effect upon H_1 the operation that we have effected upon H; i.e. we make a new unitary transformation with the operator

$$U_1 = \exp(\beta \mathscr{J}_1/2m),$$

and denote the new Hamiltonian by H_2. Its "odd" part $\mathscr{J}_2$ is smaller than that of H_1 by a factor of the order of $\mathscr{P}_1/m$ or $(\mathscr{J}_1/m)^2$, whichever is the larger: in the non-relativistic limit, $\mathscr{P}_1/m$ is of the order of $(v/c)^2$, $(\mathscr{J}_1/m)^2$ of the order of $(v/c)^6$ and $\mathscr{J}_2$ is therefore of the order of $(v/c)^5$. If one neglects terms of this order, H_2 is an "even" operator given by the formula:

$$H_2 \simeq \beta m + \mathscr{P}_2 + O(v^5)$$

$$\simeq \beta m + \mathscr{P} + \beta \frac{\mathscr{J}^2}{2m} - \frac{1}{8} m \left[\frac{\mathscr{J}}{m}, \left[\frac{\mathscr{J}}{m}, \frac{\mathscr{P}}{m}\right]\right] - \frac{1}{8} \beta m \left(\frac{\mathscr{J}}{m}\right)^4 + O(v^5)$$

$$\simeq \beta m + e\varphi + \frac{1}{2m} \beta(\boldsymbol{\sigma} \cdot \boldsymbol{\pi})^2 - \frac{e}{8m^2} [(\boldsymbol{\sigma} \cdot \boldsymbol{\pi}), [(\boldsymbol{\sigma} \cdot \boldsymbol{\pi}), \varphi]] - \frac{1}{8m^3} \beta(\boldsymbol{\sigma} \cdot \boldsymbol{\pi})^4 + O(v^5).$$

Similarly, if one neglects terms of order $(v/c)^3$, H_1 is the "even" operator given by

$$H_1 \simeq \beta m + e\varphi + \frac{1}{2m}\beta(\boldsymbol{\sigma}\cdot\boldsymbol{\pi})^2 + \mathrm{O}(v^3).$$

We may now pass over to the two-component theory as in the case of the free particle. To within $(v/c)^5$, the positive energy solutions are represented by the wave functions Φ' of the space of the large components which obey the equation

$$\mathrm{i}\frac{\partial \Phi'}{\partial t} = (m + H'_{n.r.})\,\Phi',$$

where $(m+H'_{n.r.})$ is the projection of the above approximate expression for H_2 onto the space of the large components; i.e.

$$H'_{n.r.} = e\varphi + \frac{1}{2m}(\boldsymbol{\sigma}\cdot\boldsymbol{\pi})^2 - \frac{e}{8m^2}[(\boldsymbol{\sigma}\cdot\boldsymbol{\pi}), [(\boldsymbol{\sigma}\cdot\boldsymbol{\pi}), \varphi]] - \frac{1}{8m^3}(\boldsymbol{\sigma}\cdot\boldsymbol{\pi})^4. \quad \text{(XX.199)}$$

The first two terms are the Hamiltonain $H_{n.r.}$ of Pauli theory. The last two terms are relativistic corrections of order $(v/c)^2$ to the non-relativistic energy $H_{n.r.}$.

A simple calculation gives

$$(\boldsymbol{\sigma}\cdot\boldsymbol{\pi})^4 = (\boldsymbol{\pi}^2 - e(\boldsymbol{\sigma}\cdot\mathscr{H}))^2 \quad \text{(XX.200)}$$

$$[(\boldsymbol{\sigma}\cdot\boldsymbol{\pi}), [(\boldsymbol{\sigma}\cdot\boldsymbol{\pi}), \varphi]] = \operatorname{div}\mathscr{E} + 2\boldsymbol{\sigma}\cdot(\mathscr{E}\times\boldsymbol{\pi}), \quad \text{(XX.201)}$$

which allows $H'_{n.r.}$ to be put in a more familiar form.

By successive application of a sufficient number of these unitary transformations, one may thus construct a two-component theory giving the positive energy states to any desired order in v/c. Each new transformation reduces the error by a factor of $(v/c)^2$. The study of the convergence of the series is rather delicate; it is very likely that it is an asymptotic expansion in most cases. Roughly speaking, it is a power series expansion in the operators $\mathbf{p}/m$, that is $(\hbar/mc)$ grad [and in $\partial/m\partial t$, that is, $(\hbar/mc^2)\partial/\partial t$]. The rate of convergence of the series therefore depends on the smallness of the variation of the potential $(\boldsymbol{A}, \varphi)$ over a distance of the order of $\hbar/mc$ [and over a time interval of the order of $\hbar/mc^2$, the interval necessary to travel one Compton wave length at the velocity of light].

34. Electron in a Central Electrostatic Potential

As an application of the technique described in the preceding paragraph, consider an electron in a central electrostatic potential $V(r) \equiv e\varphi(r)$. In this case, $A(r) = 0$ and the Hamiltonian of the Pauli theory is just the Hamiltonian of the ordinary Schrödinger theory:

$$H_{n.r.} = \frac{\mathbf{p}^2}{2m} + V(r).$$

If one wishes to continue the calculations to the order immediately above v^2/c^2, one must replace $H_{n.r.}$ by $H'_{n.r.}$. This amounts to adding the last two terms of (XX.199) to $H_{n.r.}$. In the present case

$$e\mathscr{E} = -\operatorname{grad} V = -\frac{\mathbf{r}}{r}\frac{\mathrm{d}V}{\mathrm{d}r}$$

$$e \operatorname{div} \mathscr{E} = -\triangle V$$

which gives, taking into account relation (XX.200) and (XX.201),

$$H'_{n.r.} = H_{n.r.} - \frac{p^4}{8m^3c^2} + \frac{\hbar^2}{4m^2c^2}\frac{1}{r}\frac{\mathrm{d}V}{\mathrm{d}r}(\boldsymbol{\sigma}\cdot\mathbf{L}) + \frac{\hbar^2}{8m^2c^2}\triangle V. \quad \text{(XX.202)}$$

The first correction term, $-p^4/8m^3c^2$, is the relativistic correction to the kinetic energy $p^2/2m$. The second is a spin-orbit coupling term [cf. formula (XIII.95)]. The third term, $\hbar^2\triangle V/8m^2c^2$, is a correction to the central potential known as the Darwin term; if $V(r) = -Ze^2/r$ (pure Coulomb potential), the Darwin term is equal to

$$(\pi Ze^2\hbar^2/2m^2c^2)\,\delta(\mathbf{r})$$

and affects only the s-states.

35. Discussions and Conclusions

In the presence, as in the absence, of a field, the operators of the two component non-relativistic theory are the projections of operators of the FW "representation" on the space of the large components. In particular, the operator $\mathbf{r}$ of the non-relativistic theory can be identified with what we have called the "average position" $\mathbf{R}$. In the Dirac theory, the interaction of the particle with the electromagnetic potential is a local interaction, in other words the particle interacts with the electromagnetic potential at its position $\mathbf{r}$. When we pass to the "FW representation", where $\mathbf{r}$ represents the "average

position", this interaction is transformed into a non-local interaction which has contributions from the values taken by the electromagnetic potential in a domain about the particle of approximate dimensions $\hbar/mc$; if the potential $(\boldsymbol{A}, \varphi)$ varies little in this domain, this interaction can be represented by a Taylor expansion involving the value of the potential and its successive derivatives at the point $\boldsymbol{r}$. A Hamiltonian such as $H'_{n.r.}$ [eqs. (XX.199) or (XX.202)] contains the first terms of this expansion.

Thus in the non-relativistic limit, the Dirac electron appears not as a point charge, but as a distribution of charge and current extending over a domain of linear dimensions $\hbar/mc$. This explains the appearance of interaction terms characteristic of the presence of a magnetic moment (interaction $-\boldsymbol{\mu}\cdot\boldsymbol{\mathscr{H}}$, spin-orbit interaction) and of an extended charge distribution (Darwin term).

Finally, we note that the application of the non-relativistic approximation to potentials that are singular at the origin such as $\boldsymbol{A} = \boldsymbol{M}\times\boldsymbol{r}/r^3$ or $\varphi = -Ze/r$ is not rigorously justified since in the neighborhood of the point $r=0$ the quantities $e\boldsymbol{A}/m$ and $e\varphi/m$ cease to be small. If the method of successive approximations described in this section was continued sufficiently far, terms sufficiently singular in the origin to give an infinite contribution to the energy would make their appearance in the non-relativistic Hamiltonian. The solution to this difficulty is suggested by the preceding discussion. In the non-relativistic Hamiltonian, $\boldsymbol{A}$ and φ are replaced by a certain average of these quantities over a domain of linear dimensions of the order of $\hbar/mc$. If the non-relativistic approximation is really justified, this amounts to effecting a cut-off of the singularity at a distance $\hbar/mc$ from the origin in all the singular expressions encountered in the calculation. In order for the non-relativistic approximation to be applicable to the two cases mentioned above, it suffices that [1]):

$$e|\boldsymbol{A}| \ll mc^2, \qquad e\varphi \ll mc^2$$

at the point $r=\hbar/mc$.

---

[1]) If m_N is the mass of the atomic nucleus $|\boldsymbol{M}| \simeq Ze\hbar/m_Nc$; the quantity eA/mc^3 is of the order of $(e^2/\hbar c)(Zm/m_N)$ at the point $r = \hbar/mc$, i.e. 10^{-5} to 10^{-6}; our calculation of the hyperfine coupling is therefore entirely justified. With regard to the example of § 34, the quantity $e\varphi/mc^2$ being of the order of $Ze^2/\hbar c$, the calculation is valid if $Z \ll 137$.

VI. NEGATIVE ENERGY SOLUTIONS AND POSITRON THEORY

Θαλασσα! θαλασσα! (*Anabasis*, IV.8).

36. Properties of Charge Conjugate Solutions

The concept of charge conjugation defined in § 19 will be useful to us in the following discussion. Charge conjugation is an antilinear and reciprocal correspondence between wave functions representing the evolution of two different particles in the same electromagnetic potential $(\boldsymbol{A}, \varphi)$: these particles have the same mass m but opposite charges $+e$ and $-e$.

If a physical quantity associated with the first particle is represented by $Q(e)$ the same quantity associated with the second particle is represented by $Q(-e)$. Thus the momentum is represented in both cases by $\boldsymbol{p} \equiv -i\nabla$, and the energy is represented by

$$H(e) \equiv \boldsymbol{\alpha}\cdot(\boldsymbol{p}-e\boldsymbol{A})+\beta m+e\varphi$$

in the one case and by

$$H(-e) \equiv \boldsymbol{\alpha}\cdot(\boldsymbol{p}+e\boldsymbol{A})+\beta m-e\varphi$$

in the other.

Consider a solution $\Psi(\boldsymbol{r}, t)$ and the charge-conjugate solution $\Psi^C(\boldsymbol{r}, t)$. We wish to compare the physical properties of the states represented by these solutions. One knows that

$$\Psi^C = K_C\Psi, \tag{XX.203}$$

where K_C is the antilinear operator defined by equation (XX.124). The notation $\langle Q\rangle$ will be used to denote the average value of Q in the state Ψ and the notation $\langle Q\rangle_C$ to denote the average value of the same operator in the state Ψ^C. Ψ and Ψ^C being supposed normalized to unity, we have

$$\langle Q\rangle = \langle\Psi|Q|\Psi\rangle$$
$$\langle Q\rangle_C = \langle\Psi^C|Q|\Psi^C\rangle.$$

From relation (XX.203),

$$\langle Q\rangle_C = (\langle\Psi|K_C{}^\dagger)(QK_C|\Psi\rangle)$$
$$= \langle\Psi|(K_C{}^\dagger QK_C)|\Psi\rangle^* = \langle\Psi|(K_C{}^\dagger Q^\dagger K_C)|\Psi\rangle$$

from which we obtain the relation between average values

$$\langle Q\rangle_C = \langle(K_C{}^\dagger Q^\dagger K_C)\rangle. \tag{XX.204}$$

By applying this relation and making use of the properties of the antiunitary transformation K_C, one finds the following relations between average values in the state Ψ and in the charge conjugate state:

$$\begin{aligned} &\langle\beta\rangle_C = -\langle\beta\rangle \qquad \langle\boldsymbol{\alpha}\rangle_C = \langle\boldsymbol{\alpha}\rangle \qquad \langle\boldsymbol{\sigma}\rangle_C = -\langle\boldsymbol{\sigma}\rangle \\ &\langle\boldsymbol{r}\rangle_C = \langle\boldsymbol{r}\rangle \qquad \langle\boldsymbol{p}\rangle_C = -\langle\boldsymbol{p}\rangle \qquad \langle\boldsymbol{L}\rangle_C = -\langle\boldsymbol{L}\rangle \\ &\langle P(\boldsymbol{r}_0)\rangle_C = \langle P(\boldsymbol{r}_0)\rangle \qquad \langle\boldsymbol{j}(\boldsymbol{r}_0)\rangle_C = \langle\boldsymbol{j}(\boldsymbol{r}_0)\rangle \qquad \langle\boldsymbol{J}\rangle_C = -\langle\boldsymbol{J}\rangle \\ &\langle H(-e)\rangle_C = -\langle H(e)\rangle. \end{aligned} \tag{XX.205}$$

It is seen that the two charge-conjugate solutions have the *same probability density and the same current density at all points* — thus opposite charge densities and electric-current densities — *but opposite energies*: charge conjugation changes the sign of the energy.

37. Abnormal Behavior of the Negative Energy Solutions

After these preliminaries, we are in a position to discuss the question of negative energies in detail.

We first consider the free particle case. The solutions of the Dirac equation were given in § 23. The energy spectrum is made up of two continuous bands $(-\infty, -mc^2)$ and $(mc^2, +\infty)$ separated by an interval of $2mc^2$ (Fig. XX.2*a*). The first of these bands corresponds to negative energy states: $E = -E_p = -\sqrt{m^2+p^2}$, and the second to positive energy states.

We propose to study the motion of a packet of free waves. It will be shown that in general it is only in the average that the center of the packet follows the classical trajectory. To this effect, we integrate the equations of motion in the Heisenberg "representation" which in this case are:

$$\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} = \mathrm{i}[H, \boldsymbol{r}] = \boldsymbol{\alpha} \tag{XX.206}$$

$$\begin{aligned} \frac{\mathrm{d}\boldsymbol{\alpha}}{\mathrm{d}t} &= \mathrm{i}[H, \boldsymbol{\alpha}] = \mathrm{i}(H\boldsymbol{\alpha} + \boldsymbol{\alpha}H) - 2\mathrm{i}\boldsymbol{\alpha}H \\ &= 2\mathrm{i}\boldsymbol{p} - 2\mathrm{i}\boldsymbol{\alpha}H, \end{aligned} \tag{XX.207}$$

$\boldsymbol{p}$ and H being constant in time, equation (XX.207) is easily integrated to give

$$\boldsymbol{\alpha}(t) = \left(\boldsymbol{\alpha}(0) - \frac{\boldsymbol{p}}{H}\right) \mathrm{e}^{-2\mathrm{i}Ht} + \frac{\boldsymbol{p}}{H}.$$

The dependence of $d\boldsymbol{r}/dt$ on t being thus explicitly given, equation (XX.206) can easily be integrated to give:

$$\boldsymbol{r}(t) = \boldsymbol{r}(0) + \frac{\boldsymbol{p}}{H}\, t + \mathrm{i}\left(\boldsymbol{\alpha}(0) - \frac{\boldsymbol{p}}{H}\right) \frac{\mathrm{e}^{-2\mathrm{i}Ht}}{2H}. \qquad \text{(XX.208)}$$

Equation (XX.208) gives the operator $\boldsymbol{r}$ of the Heisenberg "representation" at time t as a function of the values taken by the operators $\boldsymbol{r}$ and $\boldsymbol{\alpha}$ at the initial time $t=0$. From it we can obtain the law of motion of the center $\langle \boldsymbol{r} \rangle$ of any wave packet, which it will be instructive to compare with the classical law:

$$\boldsymbol{r}_{\mathrm{cl}}(t) = \boldsymbol{r}_{\mathrm{cl}}(0) + \left(\frac{\boldsymbol{p}}{H}\right)_{\mathrm{cl}} t.$$

Instead of the classical uniform rectilinear motion, the free wave packet follows a complicated motion resulting from the addition of a uniform rectilinear motion of velocity $\langle \boldsymbol{p}/H \rangle$ and a rapidly oscillatory motion,

$$\left\langle \mathrm{i}\left(\boldsymbol{\alpha}(0) - \frac{\boldsymbol{p}}{H}\right) \frac{\mathrm{e}^{-2\mathrm{i}Ht}}{2H} \right\rangle,$$

whose amplitude and period are of the order of $\hbar/2mc$ and $\hbar/2mc^2$ respectively. This oscillatory motion is called "Zitterbewegung".

The "Zitterbewegung" term vanishes if the packet is a superposition of only positive or only negative energy waves. To see this it suffices to show that

$$\Gamma_{\pm}\left(\boldsymbol{\alpha} - \frac{\boldsymbol{p}}{H}\right) \frac{\mathrm{e}^{-2\mathrm{i}Ht}}{2H}\, \Gamma_{\pm} = 0;$$

where Γ_+ and Γ_- are the projectors onto the states of positive and negative energies respectively [definition (XX.195)]. One finds successively

$$[H, \boldsymbol{\alpha}] = 2\boldsymbol{p} - 2\boldsymbol{\alpha}H$$

$$[\Gamma_{\pm}, \boldsymbol{\alpha}] = \pm \frac{\boldsymbol{p}}{E_p} \mp \boldsymbol{\alpha} \frac{H}{E_p},$$

and since $H\Gamma_{\pm} = \pm E_p \Gamma_{\pm}$, it can be deduced that

$$0 \equiv \Gamma_{\pm}[\Gamma_{\pm}, \boldsymbol{\alpha}]\, \Gamma_{\pm} = \Gamma_{\pm}\left(\frac{\boldsymbol{p}}{H} - \boldsymbol{\alpha}\right) \Gamma_{\pm},$$

from which we obtain the enunciated property by using the fact

that H commutes with Γ_+ and Γ_-. "The Zitterbewegung" is therefore caused by interference between the positive and negative energy components of the wave packet.

The "Zitterbewegung" is a curious effect related to negative energies but does not, in itself, constitute a difficulty. The difficulty appears when one studies the motion of a wave packet formed exclusively of negative energy states. In this case, the "Zitterbewegung" disappears; the center of the packet describes a uniform rectilinear motion of velocity:

$$\boldsymbol{v} = \left\langle \frac{\boldsymbol{p}}{H} \right\rangle = -\left\langle \frac{\boldsymbol{p}}{E_p} \right\rangle$$

in the opposite direction to its momentum $\langle \boldsymbol{p} \rangle$. In particular, in a non-relativistic limit ($H \simeq -mc^2$), one has the relation $\boldsymbol{v} = -\langle \boldsymbol{p} \rangle / m$, i.e. the particle behaves as if it had a negative mass $-m$.

This type of difficulty is even more apparent when one studies the motion of wave packets in a static field.

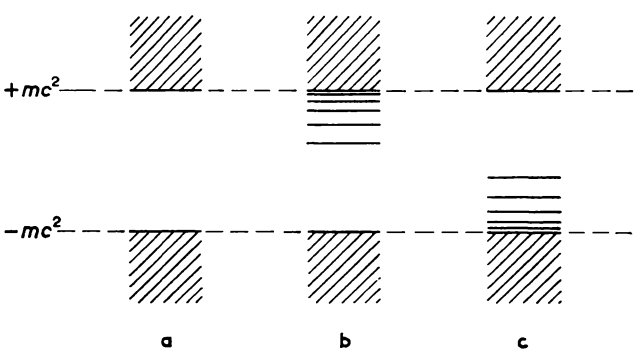


Fig. XX.2. Energy spectrum of a Dirac electron: (*a*) free; (*b*) in the attractive potential $-Ze^2/r$; (*c*) in the repulsive potential Ze^2/r.

Consider, for example, an electron in the attractive Coulomb potential $-Ze^2/r$. The spectrum (cf. Fig. XX.2*b*) is made up of a continuous positive energy band from mc^2 to ∞, a series of positive energy levels smaller than mc^2 and a continuous negative energy band from $-mc^2$ to $-\infty$. To picture the negative energy states, recall that they correspond by charge conjugation to the states of a particle of the same mass and of opposite charge (that is, a positron) in the

same potential or, what amounts to the same, to the states of an electron in the repulsive potential Ze^2/r. In this correspondence the energy changes sign and the small and large components are interchanged, but the densities and the current densities remain the same [cf. eq. (XX.205)]. The spectrum of the electron in the repulsive potential Ze^2/r is shown in Figure XX.2*c*. The positive energy continuum in the repulsive potential corresponds to the negative energy continuum in the attractive potential.

Let us consider the motion of a packet of negative energy waves in the potential $-Ze^2/r$, assuming that the non-relativistic approximation is valid ($Ze^2 \ll 1$, energies near $-mc^2$); the motion is the same as that of the packet of positive energy waves that corresponds by charge conjugation. In particular, in the limit of very small velocities, the classical approximation may be applied (cf. § VI.5) and the motion at the center of the packet is essentially that of a classical electron in the potential with the opposite sign, that is, of *a particle of negative mass* $-m$ in the potential $-Ze^2/r$: the velocity points in the opposite direction to the momentum, the acceleration in the opposite direction to the force. Such a situation has never been observed experimentally.

38. Reinterpretation of the Negative Energy States. Theory of "Holes" and Positrons

As they stand, the negative energy solutions have no physical significance. If it were possible to completely decouple the positive and negative energy states the latter could simply be ignored. Such, however, is not the case.

Consider, for example, a free electron in a positive energy state E_+, and subject it during a time interval $(0, t)$ to a radio-frequency field of frequency ω. If t is sufficiently long and the intensity of the field not too strong, the resulting effect can be calculated by the method of § XVII.6; one finds a non-vanishing probability for the electron to make a transition to a state of energy $E_+ + \hbar\omega$ or $E_+ - \hbar\omega$. In particular if

$$\hbar\omega > E_+ + mc^2,$$

the second transition is made to a state of negative energy.

As another example, consider the complete spectrum of the hydrogen atom (Fig. XX.2*b*). Owing to the coupling of the electron with the electromagnetic field, there is always a certain probability of a radiative

transition from a given state of the atom to a state of lower energy. Consequently an electron in one of the bound states of the hydrogen atom can, even if isolated, make quantum jumps to states of negative energy with emission of one or several photons; further, since the spectrum has no lower bound, the hydrogen atom has no stable state [1]).

In order to avoid these difficulties Dirac has made the following suggestion. In what one calls the "vacuum", all of the states of negative energy are occupied by an electron. If an electron is added to this "vacuum", it will necessarily be in a positive energy state since all of the negative energy states are occupied and electrons obey Fermi–Dirac statistics.

The "vacuum" therefore appears as a completely degenerate Fermi gas of infinite density. In addition, it is supposed that it is completely unobservable, giving rise to no gravitational or electromagnetic effects. The observable physical properties of a given state will be the deviations of that state from this "vacuum". Thus the observable charge of the system (electron+"vacuum") is the difference between the total charge of the system and the charge of the "vacuum", i.e. the charge of the electron. Similarly the observable energy of this system is the difference between its total energy and the energy of the "vacuum", and is therefore the energy of the electron. Up to the present, therefore, the only effect of redefining the vacuum in this way, and reinterpreting measurable quantities accordingly, is to forbid transitions to the negative energies, owing to the exclusion principle [2]).

Let us now consider what will be observed when an electron of the negative energy "sea" is missing. Applying the above convention concerning observable physical properties, we can conclude that this "hole" will have a charge opposite to that of the missing electron. It will also have an energy of opposite sign, that is a positive energy, and a momentum in the opposite direction. These considerations are valid whether or not the missing electron is in an eigenstate of the Hamiltonian. If, in particular, the missing electron forms a wave packet moving with velocity $\boldsymbol{v}$, the "hole" moves with the same

---

[1]) O. Klein has formulated a celebrated paradox which exhibits in another way the existence of a non-vanishing probability of transition to negative energy states. Klein's paradox is expounded in many treatises, for example, in M. Born, *loc. cit.*, note 1, p. 4, Vol. I.

[2]) In particular, the Zitterbewegung effect is automatically eliminated.

velocity but opposite momentum: the "hole" therefore acts like a particle of positive mass $+m$ and charge $-e$. Such particles have been observed in nature: they are called *positrons*.

Under the action of an electromagnetic field or any other suitable perturbation, an electron from the negative energy "sea" can make a transition to a state of positive energy. The "hole" of negative energy appears as a positron. In such a transformation, a pair of particles of opposite sign is thus formed. The creation of positron-electron pairs has been observed experimentally.

Similarly, if there is a "hole" in the negative energy "sea", an electron in a positive energy state can make a transition to this unoccupied negative energy state with emission of photons. This phenomenon of annihilation of an electron-positron pair with emission of photons has also been observed experimentally.

39. Difficulties with the "Hole" Theory

The "hole" theory which was briefly outlined above, permits the reconciliation of the Dirac theory with the experimental facts: non-existence of negative energy states, existence of positrons, creation and annihilation of pairs. It therefore constitutes a considerable step forward. However, it has a number of limitations and difficulties.

First of all, it is incomplete. By postulating the occupation of the quasi-totality of negative energy states, *the theory ceases to be a one-particle theory*, even when it sets out to describe a single electron. The formalism of the Dirac theory of a single particle, as set forth in this chapter, is therefore insufficient for describing such a situation, and it is only in the framework of Field Theory that one can hope to obtain a self-consistent description.

The hole theory is only a first step in the direction of a correct theory of the quantized electron field. It has the merit of providing simple pictures and can therefore serve as a guide in the elaboration of the correct theory. But pitfalls and contradictions appear when it is pushed too far.

For example, having defined the "vacuum" as composed of an infinite number of electrons, it is inconsistent to assume that these electrons do not interact.

Another weak point of the theory is the apparently very unsymmetrical role played by the electrons and the positrons. One can also

construct a corresponding charge-conjugate theory where the positrons play the role of the particles and the electrons that of the holes without any of the physical consequences being changed. All of these difficulties can be avoided in the Field Theory formalism by starting from equations invariant with respect to charge conjugation.

Finally, we note that even the definition of the negative energy states depends on the applied electromagnetic potential. In the two cases considered in § 37, namely the free particle and the particle in a Coulomb field, the space of the negative energy states is not the same. If, for example, the ground-state wave function for the hydrogen atom is expanded in a series of plane waves, one finds a weak but non-vanishing contribution from plane waves of negative energy. In the above definition of the "vacuum", the states of negative energy considered are those of the free particle; indeed it is natural to define the vacuum in the absence of a field. The introduction of an external electromagnetic field modifies this state of the "vacuum" (pair creation), the latter acting like a polarizable medium, in such a way that an electric charge in the "vacuum" seems smaller than it really is. Such effects are also found in field theory. Hole theory predicts these effects but gives no reliable and self-consistent method for their calculation.

Note Added in Proof (cf. p. 913).

To be entirely correct, one should state the parity question as follows. To the transformation s of the orthochronous Lorentz group, there correspond two transformations, s' and s'', of the group $G^{(a)}$ of Lorentz transformation operators. These are respectively represented by the *two* parity operators P' and P'' given by $P' = c' \gamma^0 P^{(0)}$ and $P'' = c'' \gamma^0 P^{(0)}$, with $c' = -c''$. The set (c', c'') defines the *intrinsic parity* of the particle. According to the discussion of § 13, there are two possible intrinsic parities, $(+1, -1)$ and $(-1, +1)$. They correspond to two inequivalent irreducible representations of the group $G^{(a)}$. Since we are dealing here with a single particle, this notion of intrinsic parity is academic. However, it is relevant in Quantum Field Theory when the interaction between several different particles is considered.

A similar treatment applies if $G^{(b)}$ instead of $G^{(a)}$ is chosen as the group of Lorentz transformation operators (cf. notes, p. 904 and p. 908). One again finds two possible values of the intrinsic parity, namely $(+i, -i)$ and $(-i, +i)$.

It should be noted that, contrary to what is often stated in the literature, the number of possible intrinsic parities for spinors *is not four but two.*

EXERCISES AND PROBLEMS

1. Show that if Ψ satisfies the Klein–Gordon equation with a field A^μ, the equation of continuity is satisfied by the four-current:

$$j^\mu = \frac{\mathrm{i}}{2m}[\Psi^*(D^\mu\Psi) - \Psi(D^\mu\Psi)^*] \equiv \frac{\mathrm{i}}{2m}[\Psi^*(\partial^\mu\Psi) - \Psi(\partial^\mu\Psi^*)] - \frac{e}{m}A^\mu\Psi^*\Psi$$

(cf. Problem IV.1).

2. Consider a hydrogen atom in which the electron is replaced by a particle of the same mass and the same charge obeying the Klein–Gordon equation. The levels E of the discrete spectrum are then given by the eigenvalue equation:

$$\left[\Delta + m^2 - \left(E - \frac{e^2}{r}\right)^2\right]\Psi(\mathbf{r}) = 0.$$

Show that this equation can be solved exactly by separating the angular and radial variables, and that the levels of the discrete spectrum depend on the quantum numbers n and l according to the formula:

$$E^{nl} = m\left(1 + \frac{e^4}{(n-\varepsilon_l)^2}\right)^{-\frac{1}{2}} \qquad \varepsilon_l = l + \tfrac{1}{2} - [(l+\tfrac{1}{2})^2 - e^4]^{\frac{1}{2}}$$
$$[n = 1, 2, \ldots, \infty; \qquad l = 0, 1, \ldots, n-1].$$

Compare this spectrum with that given by the non-relativistic Schrödinger theory.

3. Show that all of the (not necessarily unitary) γ^A matrices defined in Table 1 (§ 10), have a determinant equal to 1.

4. If B is the matrix defined at the end of § 10, show that $BB^* = B^*B = -I$. [Show first that

(*i*) BB^* is a multiple of the unit element, and therefore that $BB^* = B^*B = \pm I$;

(*ii*) the matrix BB^* is the same whatever the system of 4 unitary matrices γ^μ used to define B.]

5. Prove the following properties of the (antiunitary) charge conjugation operator K_C defined in § 19

$$K_C\, p_\mu\, K_C^\dagger = -p_\mu, \qquad K_C\, J_{\alpha\beta}\, K_C^\dagger = -J_{\alpha\beta}$$
$$K_C\, P\, K_C^\dagger = -P, \qquad K_C\, K_T\, K_C^\dagger = -K_T.$$

From these, deduce that with the choice made for the phase of the transformation operators in § 17, K_C commutes with the operators of translation

and of the proper Lorentz transformations and that it anticommutes with the spatial reflections and time reversal. How must the choice of phases be modified in order to have K_C commute with all of these transformations.

6. From the Dirac Hamiltonian deduce the equation of motion (XX.147) and (XX.148) for the operators $\boldsymbol{r}$ and $\boldsymbol{\pi}$ in the Heisenberg representation. Similarly deduce the equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[(\boldsymbol{r}\times\boldsymbol{\pi})+\frac{1}{2}\boldsymbol{\sigma}\right]=\boldsymbol{r}\times\boldsymbol{F}, \qquad \frac{\mathrm{d}M}{\mathrm{d}t}=\boldsymbol{\alpha}\cdot\boldsymbol{F}\equiv e\boldsymbol{\alpha}\cdot\boldsymbol{\mathscr{E}},$$

where $\boldsymbol{F}$ is the "Lorentz force": $\boldsymbol{F}\equiv e(\boldsymbol{\mathscr{E}}+\boldsymbol{\alpha}\times\boldsymbol{\mathscr{H}})$; compare these with equations (XX.22) and (XX.21′) from classical dynamics.

7. In the absence of a field, any solution of the Dirac equation is a solution of the Klein–Gordon equation. Show by giving a counter example that the converse is not true.

8. Prove identities (XX.169).

9. Expand a Dirac plane wave of momentum $\boldsymbol{p}$ directed along the z axis into free spherical waves.

10. Make a systematic search for the wave functions of the hydrogen atom such that the radial functions F and G are multiples of each other. Verify that the levels found correspond to $n'=0$ (therefore $J=n-\frac{1}{2}$) and $l=n-1$. One finds (notations of § 27):

$$E_{n,n-\frac{1}{2}}=m\left(1-\frac{e^4}{n^2}\right)^{\frac{1}{2}}$$

$$F=Cst.\times\varrho^s\,\mathrm{e}^{-\varrho}, \qquad G=-\nu F$$

with

$$s=\sqrt{n^2-e^4}, \qquad \varkappa=me^2/n, \qquad \varrho=\varkappa r, \qquad \nu=\varkappa/(E+m)\simeq e^2/2n.$$

11. Following the method described in § 27, calculate the levels of the hydrogen atom predicted by the Dirac theory.

12. Compare the fine structure of the levels of the hydrogen atom as given by the Dirac theory and as given by the Klein–Gordon theory of a particle of the same mass and same charge in the same Coulomb field (cf. Problem XX.2).

13. Calculate the relativistic corrections of order v^2/c^2 given by expression (XX.202) for the levels $2s_{1/2}$, $2p_{1/2}$ and $2p_{3/2}$ of the hydrogen atom. Verify that in this approximation the states $2s_{1/2}$ and $2p_{1/2}$ remain at the same level and compare the results with those given by the exact treatment of § 27.

CHAPTER XXI

FIELD QUANTIZATION. RADIATION THEORY

Se non è vero, è bene trovato

1. Introduction

The purpose of this last chapter is twofold: on the one hand to familiarize the reader with the concept of a quantized field and provide an introduction to the Quantum Theory of fields, its physical significance, its successes and its limitations; on the other, to give an elementary treatment of the Quantum Theory of electromagnetic radiation and the coupling of radiation with atomic and nuclear systems.

We shall establish the Quantum Theory of a field by the same method as was used for the non-relativistic Quantum Theory of particle systems. The method is based on the correspondence principle and starts from Hamilton's canonical formalism for the classical field. The only difference is that here it is applied to a system with an infinite number of degrees of freedom. The resulting formalism is called the Hamiltonian formalism.

The main shortcoming of this formalism is that it is not formally covariant [1]). It is possible to put the theory in a completely covariant form by starting from covariant equations for the classical field and adopting a quantization method that preserves their covariant character. Such a formalism has indisputable advantages, both practical and conceptual. Many calculations then have a simple character owing to the fact that the expressions automatically conserve a condensed and symmetrical form. On the other hand, the covariant formalism requires the construction of an elaborate mathematical framework. The Hamiltonian formalism uses more familiar concepts and is better suited for an introduction.

The chapter is divided into four sections, of which the first two are devoted to the simplest type of field, the real scalar field. Section I is devoted to the free field; in it we give the method of quantizing the classical field and its variants, the definition of state-vector space

[1]) This obviously does not imply that the relativity principle is violated. However, the relativistic invariance of the theory is not *a priori* evident and must be demonstrated. The demonstration will not be given here.

and the interpretation of the corresponding quantum states. In section II, we consider the coupling of such a field with a system of particles, and with the aid of this simple model examine the characteristic properties of a field in interaction with an atomic or nuclear system; the applications treated in this section have been chosen to illustrate the methods, the physical significance and the difficulties of Quantum Field Theory. The last two sections are devoted to electromagnetic radiation. In section III we recall the classical theory and in section IV we briefly treat the quantum theory and its applications. Most of the techniques and discussions relating to the scalar field are easily extended to radiation; they are repeated very briefly for this latter case, with special attention to the differences, which are due to the vector character of the electro-magnetic potential and to the fact that the mass of the photon is null [1]).

I. QUANTIZATION OF A REAL SCALAR FIELD

2. Classical Free Field. Normal Vibrations

A real classical scalar field is defined at each instant t by giving its amplitude $\Phi(\boldsymbol{r}, t)$ at each point $\boldsymbol{r}$ of space. It can be regarded as

[1]) There exist many treatises on Quantum Electrodynamics. We cite in particular: W. Heitler, *The Quantum Theory of Radiation* (Oxford University Press, second edition, 1944); A. I. Akhiezer and V. B. Berestetsky, *Quantum Electrodynamics* (translated from the Russian by Consultants Bureaux Inc. (1956)); J. M. Jauch and F. Röhrlich, *The Theory of Photons and Electrons* (Addison Wesley, Cambridge, 1956); N. N. Bogoliubov and D. V. Shirkov, *Introduction to the Theory of Quantized Fields* (translated from the Russian by G. M. Volkoff; Interscience Publishers Inc., New York, 1959); S. S. Schweber, *An Introduction to Relativistic Quantum Field Theory* (Row, Peterson and Company, 1961); W. E. Thirring, *Principles of Quantum Electrodynamics* (translated from the German by J. Bernstein; Academic Press, New York, 1958). Heitler uses the Hamiltonian formalism and treats the principal applications of the theory to atomic physics. Akhiezer and Berestetsky, and Jauch and Röhrlich give a very complete account of the covariant formalism and also a very complete review of its applications. Bogoliubov and Shirkov, and in a more complete way, Schweber, give a unified exposition of Quantum Field Theory in the covariant formalism including a thorough discussion of the renormalization problem and of the implications of causality in Field Theory. Thirring gives an overall view of the covariant formalism and a detailed discussion of its physical significance and its difficulties. For the measurability of field operators, one may also consult the classical articles of N. Bohr and L. Rosenfeld, Det. Klg. Danske Viddenskab. Selskab. 12, no. 8 (1933) and Phys. Rev. 78 (1950) 794.

a dynamical system with an infinite number of degrees of freedom. To each point of space there corresponds a particular coordinate of the system, namely the amplitude of the field at that point.

The dynamics of such a system is not essentially different to that of a system with a finite number of degrees of freedom. However, since the coordinates are labelled with the aid of *continuous* parameters – for example the three components x, y, z, of the point $\boldsymbol{r}$ – differentiation with respect to time is replaced by partial differentiation. The evolution of the system is determined by an equation of motion of the form:

$$\ddot{\Phi}(\boldsymbol{r}, t) \equiv \frac{\partial^2 \Phi}{\partial t^2} = F[\Phi(\boldsymbol{r}, t), \dot{\Phi}(\boldsymbol{r}, t)],$$

where F is a functional of the amplitude $\Phi(\boldsymbol{r}, t)$ and of its partial derivative: $\dot{\Phi} \equiv \partial\Phi/\partial t$, both taken at the same instant t. The dynamical state of the system is determined for any given time if we are given its position and its velocity at an initial time t_0, i.e. its coordinates $\Phi(\boldsymbol{r}, t_0)$ and their respective derivatives with respect to the time:

$$\dot{\Phi}(\boldsymbol{r}, t_0) \equiv \left.\frac{\partial \Phi(\boldsymbol{r}, t)}{\partial t}\right|_{t=t_0}.$$

Of all the equations of motion invariant with respect to the transformations of the inhomogeneous Lorentz group, the simplest is the Klein–Gordon equation:

$$[\Box + \mu^2]\Phi \equiv \ddot{\Phi} - \triangle\Phi + \mu^2\Phi = 0. \qquad \text{(XXI.1)}$$

This equation is linear and homogeneous in Φ; it is the law of motion for a free field. μ is a constant. We shall see further along that μ is the mass of the corpuscles associated with the quantized field (§ XXI.6).

If the term $\triangle\Phi$ did not appear in the left-hand side of eq. (XXI.1) the motion of the different coordinates of the system would be independent, and each would effect a harmonic oscillation of frequency μ with amplitude and phase depending on the initial conditions. The term $\triangle\Phi$ couples these harmonic oscillations.

Such a set of coupled oscillators can be transformed into a set of independent oscillators by the introduction of normal coordinates, as will now be shown.

We denote by $f_1, f_2, \ldots, f_i, \ldots$, a complete orthonormal set of real functions of r satisfying the eigenvalue equation of the Hermitean operator $(-\triangle)$. The corresponding eigenvalue spectrum extends from 0 to $+\infty$. We denote the eigenvalue corresponding to the function f_i by k_i^2. One has, therefore

$$(\Delta + k_i^2) f_i = 0 \tag{XXI.2}$$

$$\int f_i(\boldsymbol{r})\, f_j(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r} = \delta_{ij} \tag{XXI.3}$$

$$\sum_i f_i(\boldsymbol{r})\, f_i(\boldsymbol{r}') = \delta(\boldsymbol{r} - \boldsymbol{r}'). \tag{XXI.4}$$

From the orthogonality and closure relations (XXI.3) and (XXI.4), we obtain the expansions

$$\Phi(\boldsymbol{r}, t) = \sum_i q_i(t)\, f_i(\boldsymbol{r}) \tag{XXI.5}$$

$$q_i(t) = \int f_i(\boldsymbol{r})\, \Phi(\boldsymbol{r}, t)\, \mathrm{d}\boldsymbol{r}. \tag{XXI.6}$$

The quantities $q_1, q_2, \ldots, q_i, \ldots$ are the desired normal coordinates. From equations (XXI.6) and (XXI.1),

$$\ddot{q}_i + \omega_i^2 q_i = 0, \tag{XXI.7}$$

where

$$\omega_i = (k_i^2 + \mu^2)^{\frac{1}{2}}. \tag{XXI.8}$$

Thus the equation of motion of each q_i is indeed that of an independent harmonic oscillator and the corresponding frequency is ω_i. In accordance with equation (XXI.5), the field amplitude is a linear superposition of these independent oscillators.

In all of this, we have supposed that the basis functions f were labelled with a discrete index. In fact, since the operator $(-\triangle)$ has a continuous spectrum (extending from 0 to $+\infty$), these functions depend on a set of indices, at least one of which is continuous. It is easy to repeat the above analysis with continuous indices; apart from a few obvious modifications, the details of which will not be given here, the results remain essentially the same. However, the existence of continuous indices complicates somewhat the quantum treatment. In order to avoid these inessential complications, we shall suppose that the field is defined in a finite volume of space and satisfies convenient boundary conditions on the surface of this volume, and that the physical quantities we wish to calculate may be deduced from the quantities which we actually do calculate by increasing this

volume indefinitely. This type of artifice has already been described in § V.11. It obviously is not rigorous. It should also be noted that it partly destroys, in an artificial way, the invariance properties of the theory.

3. Quantization of the Free Field

The passage to normal coordinates suggest a very simple method for quantizing the real scalar field. It consists in associating with each mode of normal vibration of the classical field a quantum oscillator of the same frequency (see Chapter XII).

Let us write the classical Hamiltonian in normal coordinates. Let p_i be the momentum conjugate to q_i. Equation (XXI.7) is then equivalent to the Hamilton equations:

$$\dot{q}_i = p_i, \qquad \dot{p}_i = -\omega_i^2 q_i.$$

These can be obtained by taking the following Hamiltonian for the ith degree of freedom:

$$h_i = \tfrac{1}{2}(p_i^2 + \omega_i^2 q_i^2), \tag{XXI.9}$$

whence the following expression for the total Hamiltonian:

$$H = \sum_i h_i. \tag{XXI.10}$$

The corresponding quantized field is obtained by replacing the real dynamical variables q_i and p_i by observables obeying the commutation relations ($\hbar = 1$):

$$[q_i, p_j] = \mathrm{i}\delta_{ij}. \tag{XXI.11}$$

To each mode of vibration there corresponds a series of discrete equidistant levels according to the formula:

$$E_{n_i} = (n_i + \tfrac{1}{2})\omega_i \qquad (n_i = 0, 1, 2, \ldots, \infty). \tag{XXI.12}$$

Here it is convenient to introduce annihilation and creation operators, which we denote by b_i and $b_i^\dagger$ respectively:

$$\begin{aligned} b_i &= (2\omega_i)^{-\frac{1}{2}}(\omega_i q_i + \mathrm{i}p_i) \\ b_i^\dagger &= (2\omega_i)^{-\frac{1}{2}}(\omega_i q_i - \mathrm{i}p_i). \end{aligned} \tag{XXI.13}$$

The commutation rules (XXI.11) between observables are equivalent to the commutation rules between Hermitean conjugate operators:

$$[b_i, b_j^\dagger] = \delta_{ij} \tag{XXI.14}$$

and the eigenvectors of h_i are obtained by the repeated action of $b_i^\dagger$ on the eigenvector of the ground state [see eq. (XII.20)].

A complete set of eigenstates of H is obtained by taking the tensor product of the eigenstates of the partial Hamiltonians $h_1, h_2, \ldots, h_i, \ldots$. Each state is labelled with the quantum numbers $n_1, n_2, \ldots, n_i, \ldots$; its energy is the sum of the partial energies

$$E_{n_1 n_2 \ldots n_i \ldots} = \sum_i (n_i + \tfrac{1}{2})\,\omega_i. \tag{XXI.15}$$

Its eigenvector, normalized to unity, is obtained from the vector $|0\rangle$ representing the ground state by the formula

$$|n_1 n_2 \ldots n_i \ldots\rangle = \prod_i \frac{(b_i^\dagger)^{n_i}}{(n_i!)^{\frac{1}{2}}}|0\rangle. \tag{XXI.16}$$

The field at the point $\boldsymbol{r}$ is a certain Hermitean operator $\Phi(\boldsymbol{r})$, defined by the expression

$$\Phi(\boldsymbol{r}) = \sum_i q_i f_i(\boldsymbol{r}) \tag{XXI.17}$$

deduced by correspondence from the classical relation (XXI.5). It is useful to express it in terms of the b_i and $b_i^\dagger$; from equations (XXI.13) and (XXI.17),

$$\Phi(\boldsymbol{r}) = \sum_i (2\omega_i)^{-\frac{1}{2}}(b_i + b_i^\dagger) f_i. \tag{XXI.18}$$

Up to now the field has appeared as a set of quantized oscillators. In the stationary state considered above, each quantum number n represents the number of vibrational quanta relative to a particular normal mode of vibration. But one may also give the eigenstates of H, and more generally the operators of the system, a corpuscular interpretation of which the essential was set forth in § XII.6. n_i then represents the number of corpuscles in the state f_i of energy ω_i; it is the eigenvalue of a certain operator N_i, the "number of corpuscles in the state f_i" operator, defined by

$$N_i = b_i^\dagger b_i. \tag{XXI.19}$$

b_i is the annihilation operator for a corpuscle in the state f_i, $b_i^\dagger$ the creation operator for a corpuscle in the state f_i.

The interpretation is self-consistent only if the total energy of a given dynamical state is the sum of the energies of the corpuscles

present in that state. As has already been noted in § XII.6, it is sufficient for this to subtract the constant $\frac{1}{2}\omega_i$ from each term h_i of the sum on the right-hand side of (XXI.10). This modification of the Hamiltonian obviously has no effect on the equations of motion, and the new Hamiltonian has the desired property, i.e.,

$$H = \sum_i N_i \omega_i. \tag{XXI.20}$$

In particular, the vacuum is the state of zero energy namely the state represented by the ket vector $|0\rangle$. Similarly, the vector $b_i^\dagger |0\rangle$ represents a state having a single corpuscle and its energy is the energy ω_i of that corpuscle.

The corpuscles considered here are indistinguishable and the dynamical state of the total system is completely determined by giving the numbers $n_1, n_2, \ldots, n_i, \ldots$ of these corpuscles in each of the individual states $f_1, f_2, \ldots, f_i, \ldots$ that can be occupied. These "occupation numbers" can take any integral value from 0 to $+\infty$. The corpuscles therefore obey Bose–Einstein statistics (cf. § XIV.6) and the quantized field represents a *system of bosons* of undetermined number.

We conclude with a remark concerning the ground state of the system. This vacuum of the quantized field differs profoundly from the classical vacuum. In the case of a classical field, the vacuum, or state of minimum energy, is a state for which the field vanishes at every point. In the case of the quantized field, the field at a given point $\Phi(\mathbf{r})$ is an operator which does not commute with the Hamiltonian. Its vacuum expectation value vanishes as may easily be seen using the properties of the b_i and the $b_i^\dagger$, but its mean square deviation does not. Indeed, using expression (XXI.18) of the field, one finds,

$$\langle \Phi^2(\mathbf{r}) \rangle_0 \equiv \sum_i \sum_j \tfrac{1}{2} (\omega_i \omega_j)^{-\frac{1}{2}} \langle 0| (b_i + b_i^\dagger)(b_j + b_j^\dagger) |0\rangle f_i(\mathbf{r}) f_j(\mathbf{r}),$$

and since

$$\langle 0|(b_i + b_i^\dagger)(b_j + b_j^\dagger)|0\rangle = \langle 0|b_i b_j^\dagger|0\rangle = \delta_{ij},$$

we have

$$\langle \Phi^2 \rangle_0 = \sum_i \frac{f_i^2(\mathbf{r})}{2\omega_i}.$$

The right-hand side is a series of positive terms, and is therefore a non-zero positive quantity. In fact it can be shown that this series

is divergent [1]). The appearance of divergent quantities is characteristic of systems with an infinite number of degrees of freedom and is a source of serious difficulties in Quantum Field Theory. We shall have occasion to return to this point in section II.

4. Lagrangian of the Field. Momentum Conjugate to $\Phi(\mathbf{r})$

Before we continue, let us return to the classical field and briefly describe the generalization of the Lagrangian and Hamiltonian formalisms to this type of system having an infinite number of degrees of freedom. If the coordinates are labelled by a set of discrete indices, the generalization is obvious; it is the one that we used with the normal coordinates. The situation is more delicate if the coordinates depend on a continuous index, in particular if the "position" of the field is defined by specifying its amplitude at each point of space. It is precisely this choice of coordinates that interests us here.

The Lagrangian function L is a function of the coordinates of the system and of the components of its velocity; in the present case, it is a functional of Φ and $\dot{\Phi}$. We assume that it may be represented by an expression of the form

$$L = \int \mathscr{L}\, \mathrm{d}\mathbf{r}. \tag{XXI.21}$$

The Lagrangian density $\mathscr{L}$ depends on $\mathbf{r}$ and t through Φ, through the three components of its gradient and through $\dot{\Phi}$:

$$\mathscr{L} \equiv \mathscr{L}\left(\Phi, \frac{\partial \Phi}{\partial x^k}, \dot{\Phi}\right).$$

Knowing L, one defines the action integral

$$I = \int_{t_1}^{t_2} L\, \mathrm{d}t = \int_{t_1}^{t_2} \mathrm{d}t \int \mathrm{d}\mathbf{r}\, \mathscr{L}. \tag{XXI.22}$$

I is a certain functional of $\Phi(\mathbf{r}, t)$. The equations of motion are given by the principle of least action [eq. (I.12)]:

$$\delta I = 0 \qquad (\text{supposing } \delta\Phi(t_1) = \delta\Phi(t_2) = 0). \tag{XXI.23}$$

---

[1]) On the other hand, the integral of the field over a finite volume has a finite fluctuation; this fluctuation increases with the smallness of the volume of integration, becoming infinite in the limit where this volume vanishes (cf. Problem XXI.1).

From relation (XXI.22), it is clear that this law of motion obeys the relativity principle *if $\mathscr{L}$ transforms like a scalar in the transformations of the Lorentz group*. This condition on $\mathscr{L}$ makes it much easier to find the Lagrangian function.

Lagrange's equations are replaced in the present case by a partial differential equation for the amplitude $\Phi(\mathbf{r}, t)$ that is easily deduced from the principle of least action. One finds:

$$\frac{\partial}{\partial t}\frac{\partial \mathscr{L}}{\partial \dot{\Phi}} + \sum_k \frac{\partial}{\partial x^k}\frac{\partial \mathscr{L}}{\partial(\partial\Phi/\partial x^k)} - \frac{\partial \mathscr{L}}{\partial \Phi} = 0. \tag{XXI.24}$$

This equation is the equation of motion for the field.

To rigorously define the field coordinates, one must imagine the whole of space divided into infinitesimal cells. To each cell $(\mathbf{r}, \mathbf{r}+\mathrm{d}\mathbf{r})$ there corresponds a coordinate of the system, namely the quantity $\Phi(\mathbf{r}, t)\,\mathrm{d}\mathbf{r}$. The law of motion fixes the variation of these coordinates as a function of time. In accordance with the usual definition, the conjugate momentum $\Pi(\mathbf{r}, t)$ of the coordinate $\Phi(\mathbf{r}, t)\,\mathrm{d}\mathbf{r}$ is given by the functional derivative of L with respect to the velocity component $\dot{\Phi}(\mathbf{r}, t)\,\mathrm{d}\mathbf{r}$, i.e.:

$$\Pi(\mathbf{r}, t) \equiv \frac{\delta L}{\delta \dot{\Phi}(\mathbf{r}, t)\,\mathrm{d}\mathbf{r}} = \frac{\partial \mathscr{L}}{\partial \dot{\Phi}}. \tag{XXI.25}$$

The concept of a functional derivative used here is a natural generalization of the concept of a partial derivative; δL is the variation of L when $\dot{\Phi}$ varies by the infinitesimal quantity $\delta\dot{\Phi}$ in a small volume of extension $\mathrm{d}\mathbf{r}$ about the point $\mathbf{r}$, while at the same time remaining constant in the rest of space.

Having defined the conjugate momentum for each coordinate, one obtains the Hamiltonian function H by applying the usual definition [eq. (I.13)]; it is a functional of Φ and Π:

$$H \equiv H[\Phi, \Pi] \equiv \int \Pi\dot{\Phi}\,\mathrm{d}\mathbf{r} - L,$$

or again

$$H = \int \mathscr{H}\,\mathrm{d}\mathbf{r}$$
$$\mathscr{H} \equiv \mathscr{H}\left(\Phi, \frac{\partial\Phi}{\partial x^k}, \Pi\right) = \Pi\dot{\Phi} - \mathscr{L}. \tag{XXI.26}$$

$\mathscr{H}$ is called the Hamiltonian density. Hamilton's canonical equations are given by [cf. eq. (I.14)]:

$$\begin{aligned} \dot{\Phi} &= \frac{\delta \mathscr{H}}{\delta \Pi} \\ \dot{\Pi} &= -\frac{\delta \mathscr{H}}{\delta \Phi} + \sum_k \frac{\partial}{\partial x^k} \frac{\delta \mathscr{H}}{\delta(\partial \Phi / \partial x^k)}. \end{aligned} \tag{XXI.27}$$

In the case of the free scalar field, the function

$$\begin{aligned} \mathscr{L} &= \tfrac{1}{2} [(\partial_\mu \Phi)(\partial^\mu \Phi) - \mu^2 \Phi^2] \\ &= \tfrac{1}{2} [\dot{\Phi}^2 - (\nabla \Phi)^2 - \mu^2 \Phi^2] \end{aligned} \tag{XXI.28}$$

can be taken as the Lagrangian density since substitution of this expression in equation (XXI.24) gives equation (XXI.1).

In this case the momentum conjugate to Φ is:

$$\Pi = \dot{\Phi}. \tag{XXI.29}$$

Applying definition (XXI.26), one obtains for the Hamiltonian function:

$$H = \tfrac{1}{2} \int (\Pi^2 + (\text{grad } \Phi)^2 + \mu^2 \Phi^2) \, d\mathbf{r}. \tag{XXI.30}$$

The function H defined above is the same as the one given in terms of the normal coordinates by equation (XXI.10). To show this, we need only to replace Φ and Π by their expansions in terms of the q_i and the p_i, and to make use of the properties of the f. The expansion of Φ is given by equation (XXI.5). As for Π, from (XXI.25) one obtains:

$$\begin{aligned} \Pi(\mathbf{r}, t) \, d\mathbf{r} &= \frac{\delta L}{\delta \dot{\Phi}(\mathbf{r}, t)} \\ &= \sum_i \frac{\partial L}{\partial \dot{q}_i} \frac{\delta \dot{q}_i}{\delta \dot{\Phi}(\mathbf{r}, t)} \end{aligned}$$

and, since by definition $p_i = \partial L / \partial \dot{q}_i$ and since, according to equation (XXI.6)

$$\dot{q}_i = \int f_i(\mathbf{r}) \, \dot{\Phi}(\mathbf{r}, t) \, d\mathbf{r},$$

from which

$$\frac{\delta \dot{q}_i}{\delta \dot{\Phi}(\mathbf{r}, t)} = f_i(\mathbf{r}) \, d\mathbf{r},$$

we have:

$$\Pi(\mathbf{r}, t) = \sum_i p_i f_i(\mathbf{r}). \tag{XXI.31}$$

The substitutions mentioned above therefore give:

$$H = \tfrac{1}{2}\sum_i \sum_j [p_i p_j \int f_i f_j \, \mathrm{d}\boldsymbol{r} + q_i q_j \int (\nabla f_i \cdot \nabla f_j) \, \mathrm{d}\boldsymbol{r} + \mu^2 q_i q_j \int f_i f_j \, \mathrm{d}\boldsymbol{r}]. \quad \text{(XXI.32)}$$

Integrating by parts and applying equation (XXI.2), we obtain

$$\int (\nabla f_i \cdot \nabla f_j) \, \mathrm{d}\boldsymbol{r} = -\int f_i \triangle f_j \, \mathrm{d}\boldsymbol{r} = k_i^2 \int f_i f_j \, \mathrm{d}\boldsymbol{r}.$$

Substituting this last expression into the right-hand side of (XXI.32) and using the orthogonality relations for the f [eq. (XXI.3)] and the definition of ω_i [eq. (XXI.8)] we finally obtain the stated result:

$$H = \tfrac{1}{2}\sum_i (p_i^2 + \omega_i^2 q_i^2).$$

[Recall that this expression differs from expression (XXI.20) by an infinite constant.]

In order to quantize the field, it makes no difference whether we take the q_i and the p_i, the b_i and the $b_i^\dagger$ or the $\Phi(\boldsymbol{r})$ and the $\Pi(\boldsymbol{r})$ as our fundamental variables. The commutation relations for the Φ and the Π have a particularly simple form. They can be obtained from the commutation relations for the b and the $b^\dagger$. Equations (XXI.13) give:

$$p_i = \tfrac{1}{2}\mathrm{i}\sqrt{2\omega_i}(b_i^\dagger - b_i).$$

Substituting this expression in expansion (XXI.31), one finds

$$\Pi(\boldsymbol{r}) = \sum_i \tfrac{1}{2}\mathrm{i}\sqrt{2\omega_i}\,(b_i^\dagger - b_i) f_i. \quad \text{(XXI.33)}$$

The commutation relations are obtained by using expansions (XXI.18) and (XXI.33), the commutation relations of the b and the $b^\dagger$ and the closure relation (XXI.4). One obtains

$$\begin{aligned} &[\Phi(\boldsymbol{r}), \Phi(\boldsymbol{r}')] = [\Pi(\boldsymbol{r}), \Pi(\boldsymbol{r}')] = 0 \\ &[\Phi(\boldsymbol{r}), \Pi(\boldsymbol{r}')] = \mathrm{i}\delta(\boldsymbol{r} - \boldsymbol{r}'). \end{aligned} \quad \text{(XXI.34)}$$

5. Complex Basis Functions

We have seen in the preceding paragraph how the field can be quantized without having to make use of the normal coordinates. Nevertheless, the use of normal coordinates usually simplifies the calculations and also the interpretation of the theory of the quantized field. Since the eigenvalues of $(-\triangle)$ are degenerate, there is a large arbitrary in the choice of the basis functions and normal coordinates.

Furthermore, it is possible, by making a very simple generalization, to quantize the field with the aid of complex basis functions. In this paragraph, we shall first give an account of this more general method of quantization, and then show that the quantized field obtained is independent of the choice of the basis functions.

Denote by $u_1, u_2, \ldots, u_s, \ldots$ a complete orthonormal set of basis functions and by ω_s the corresponding energy. Thus [cf. eq. (XXI.2–4) and (XXI.8)],

$$\int u_s^* \, u_t \, \mathrm{d}\boldsymbol{r} = \delta_{st} \qquad \sum_s u_s(\boldsymbol{r}) \, u_s^*(\boldsymbol{r}') = \delta(\boldsymbol{r}-\boldsymbol{r}') \tag{XXI.35}$$

$$(\triangle + k_s^2) u_s = 0 \tag{XXI.36}$$

$$\omega_s = (k_s^2 + \mu^2)^{\frac{1}{2}}. \tag{XXI.37}$$

To each mode u_s there are associated two Hermitean conjugate operators a_s and $a_s^\dagger$. These operators obey the commutation rules [cf. eq. (XXI.14)]:

$$\begin{aligned} [a_s, a_t] &= [a_s^\dagger, a_t^\dagger] = 0, \\ [a_s, a_t^\dagger] &= \delta_{st}, \end{aligned} \tag{XXI.38}$$

which permit the construction of the space of the dynamical states of the system. A quantized harmonic oscillator is thus associated with each mode u_s. a_s and $a_s^\dagger$ are to be interpreted respectively as annihilation and creation operators of a corpuscle in the state u_s. The observables of the system are functions of the annihilation and creation operators. In particular, the observable

$$N_s \equiv a_s^\dagger a_s$$

represents the number of corpuscles in the state u_s. The field $\Phi(\boldsymbol{r})$ is given by the expansion [cf. eq. (XXI.18)]:

$$\Phi(\boldsymbol{r}) = \sum_s (2\omega_s)^{-\frac{1}{2}} \, (a_s \, u_s(\boldsymbol{r}) + a_s^\dagger \, u_s^*(\boldsymbol{r})), \tag{XXI.39}$$

the conjugate momentum, $\Pi(\boldsymbol{r})$, by the expansion [cf. eq. (XXI.33)]:

$$\Pi(\boldsymbol{r}) = \sum_s \tfrac{1}{2} \mathrm{i} (2\omega_s)^{+\frac{1}{2}} \, (a_s^\dagger \, u_s^*(\boldsymbol{r}) - a_s \, u_s(\boldsymbol{r})), \tag{XXI.40}$$

and the Hamiltonian H governing the evolution of the system by [cf. eq. (XXI.20)]

$$H = \sum_s a_s^\dagger \, a_s \, \omega_s. \tag{XXI.41}$$

We shall now establish the relation between the a and $a^\dagger$ and the operators b and $b^\dagger$ of § 3 and show that the system defined here is indeed the same as the quantum system defined in § 3.

We have used s to denote the set of quantum numbers that distinguish between the u functions. By convention, this set contains a quantum number that determines the energy and a set σ of additional quantum numbers that distinguish between functions corresponding to the same value of the energy; we shall use the wave number k to label the energy:

$$\omega = (k^2+\mu^2)^{\frac{1}{2}}. \tag{XXI.42}$$

Thus: $s \equiv (k, \sigma)$.

Similarly the set i of quantum numbers distinguishing between the f functions contains the wave number k fixing the energy and also a certain set ϱ of additional quantum numbers: $i \equiv (k, \varrho)$.

We pass from the $f_{k\varrho}$ functions corresponding to a given value of k to the $u_{k\sigma}$ functions corresponding to the same value of k by a unitary transformation:

$$u_{k\sigma} = \sum_{\varrho} T_{\sigma\varrho} f_{k\varrho}, \qquad f_{k\varrho} = \sum_{\sigma} (T^\dagger)_{\varrho\sigma} u_{k\sigma} \tag{XXI.43}$$

($T \equiv$ unitary matrix).

All of the quantities appearing in what follows will relate to a well-defined value of k and we shall henceforth drop the index k.

A necessary and sufficient condition for expansions (XXI.18) and (XXI.39) to represent the same operator $\Phi(\boldsymbol{r})$, and for expansions (XXI.33) and (XXI.40) to represent the same operator $\Pi(\boldsymbol{r})$, is that for any value of k:

$$\sum_{\sigma} a_\sigma u_\sigma = \sum_{\varrho} b_\varrho f_\varrho .$$

Taking into account the second of relations (XXI.43), this is equivalent to writing

$$a_\sigma = \sum_{\varrho} T^*_{\sigma\varrho}\, b_\varrho \tag{XXI.44}$$

or:

$$a_\sigma^\dagger = \sum_{\varrho} T_{\sigma\varrho}\, b_\varrho^\dagger. \tag{XXI.44†}$$

Thus we have established the correspondence between the a and the b on the one hand, and the $a^\dagger$ and the $b^\dagger$ on the other. This correspondence is linear. More precisely, the linear correspondence between

the creation operators $b^\dagger$ and $a^\dagger$ [relation (XXI.44†)] is identical with the linear correspondence between the f and the u [eq. (XXI.43)].

It results from relations (XXI.44) and (XXI.44†) that expansions (XXI.18) and (XXI.39) represent the same field $\Phi(\boldsymbol{r})$. In order that the two methods of quantizing this field be equivalent, it is necessary that the a and the $a^\dagger$ thus defined obey commutation relations (XXI.38), and that the Hamiltonian H defined by equation (XXI.20) be expressed in term of these operators by expansion (XXI.41). These two results can easily be proved using the unitarity of the T matrix. From this latter property we have that for each value of k:

$$\begin{aligned}[a_\sigma, a_{\sigma'}^\dagger] &= \sum_{\varrho\varrho'} T_{\sigma\varrho}^* T_{\sigma'\varrho'} [b_\varrho, b_{\varrho'}^\dagger] \\ &= \sum_{\varrho\varrho'} T_{\sigma\varrho}^* T_{\sigma'\varrho'} \delta_{\varrho\varrho'} = \delta_{\sigma\sigma'},\end{aligned}$$

from which:

$$[a_s, a_{s'}^\dagger] = \delta_{ss'}.$$

Similarly for each value of k:

$$\begin{aligned}\sum_\sigma a_\sigma^\dagger a_\sigma &= \sum_{\varrho\varrho'} \sum_\sigma T_{\sigma\varrho} T_{\sigma\varrho'}^* b_\varrho^\dagger b_{\varrho'} \\ &= \sum_{\varrho\varrho'} \delta_{\varrho\varrho'} b_\varrho^\dagger b_{\varrho'} = \sum_\varrho b_\varrho^\dagger b_\varrho,\end{aligned}$$

from which we obtain the desired expression for H.

6. Plane Waves. Definition of the Momentum

Among the different sets of complex basis functions one has notably the plane waves. They correspond, as will be shown below, to individual states of well-defined momentum.

In order to introduce discrete indices, we shall suppose the field enclosed in a cube of side L, and require the functions u_s to be periodic at the edges of the cube, i.e.:

$$u_s(\tfrac{1}{2}L, y, z) = u_s(-\tfrac{1}{2}L, y, z)$$

with two similar equations for the arguments y and z. These conditions are a simple generalization of the periodicity condition described in § V.11.

As basis functions we may then take the plane waves:

$$u_{\boldsymbol{k}} = L^{-3/2} e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \qquad \text{(XXI.45)}$$

whose wave vector has components that are multiples of $2\pi/L$:

$$k_x = \frac{2\pi n_x}{L}, \qquad k_y = \frac{2\pi n_y}{L}, \qquad k_z = \frac{2\pi n_z}{L}.$$

A complete orthonormal system of basis functions is obtained by giving each of the numbers n_x, n_y, n_z all possible integral values:

$$n_x, n_y, n_z = 0, \pm 1, \pm 2, \pm 3, \ldots.$$

With each of these plane waves, we associate an annihilation operator $a_{\boldsymbol{k}}$ and a creation operator $a_{\boldsymbol{k}}^\dagger$. The field $\Phi(\boldsymbol{r})$ is represented by the expansion [cf. eq. (XXI.39)]:

$$\Phi(\boldsymbol{r}) = \sum_{\boldsymbol{k}} (2\omega_{\boldsymbol{k}} L^3)^{-\frac{1}{2}} (a_{\boldsymbol{k}} e^{i\boldsymbol{k}\cdot\boldsymbol{r}} + a_{\boldsymbol{k}}^\dagger e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}) \tag{XXI.46}$$

its conjugate moment [cf. eq. (XXI.40)] by:

$$\Pi(\boldsymbol{r}) = \sum_{\boldsymbol{k}} i\left(\frac{\omega_{\boldsymbol{k}}}{2L^3}\right)^{\frac{1}{2}} (a_{\boldsymbol{k}}^\dagger e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} - a_{\boldsymbol{k}} e^{i\boldsymbol{k}\cdot\boldsymbol{r}}) \tag{XXI.47}$$

and the Hamiltonian [cf. eq. (XXI.41)] by:

$$H = \sum_{\boldsymbol{k}} (a_{\boldsymbol{k}}^\dagger a_{\boldsymbol{k}})\, \omega_{\boldsymbol{k}}, \tag{XXI.48}$$

the sums being extended over all possible values of k i.e., over all possible values of n_x, n_y, n_z.

As was explained in § V.11, the correct results are obtained by letting L increase indefinitely. In this limit the summations are replaced by integrations. In a small band $(\boldsymbol{k}, \boldsymbol{k}+\delta\boldsymbol{k})$, there are $(L/2\pi)^3\, \delta\boldsymbol{k}$ possible sets of values for the three numbers n_x, n_y, n_z. If we assume – in the spirit of the passage to the limit $L \to \infty$ – that the terms over which we sum may be treated as continuous functions of $\boldsymbol{k}$, each symbol $\sum$ may be replaced by the symbol

$$\int \left(\frac{L}{2\pi}\right)^3 d\boldsymbol{k}.$$

$(L/2\pi)^3$ is the density of individual states in wave vector space. One commonly reserves the term *level density* for the number of individual states per unit solid angle and per unit energy; the level density is a certain function $\varrho_L(\omega)$ of the energy. The number $\varrho_L\, d\Omega\, d\omega$ of states

with wave vector in the solid angle $(\Omega, \Omega+\mathrm{d}\Omega)$, and energy in the band $(\omega, \omega+\mathrm{d}\omega)$ is given by

$$\varrho_L \,\mathrm{d}\Omega \,\mathrm{d}\omega = \left(\frac{L}{2\pi}\right)^3 k^2 \frac{\mathrm{d}k}{\mathrm{d}\omega} \,\mathrm{d}\omega \,\mathrm{d}\Omega$$

$$= \left(\frac{L}{2\pi}\right)^3 \omega k \,\mathrm{d}\omega \,\mathrm{d}\Omega.$$

Therefore,

$$\varrho_L(\omega) = \left(\frac{L}{2\pi}\right)^3 \omega k \tag{XXI.49}$$

[cf. eq. (XVII.52)].

In actual fact, it is possible to completely avoid the artifice of the cube, and from the beginning use plane waves extending over all space. The wave vector $\boldsymbol{k}$ may then take all values without restriction and the waves in question depend on the 3 *continuous indices* k_x, k_y, k_z. Normalization with Kronecker symbols is then replaced by normalization with δ functions.

One then takes as basis functions the functions:

$$u(\boldsymbol{k}) \equiv u(\boldsymbol{k}; \boldsymbol{r}) = \frac{\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}}{(2\pi)^{3/2}}$$

and associates with each of them the Hermitean conjugate operators $a(\boldsymbol{k})$ and $a^\dagger(\boldsymbol{k})$. The different formulas are obtained from the preceding ones by everywhere replacing $u_{\boldsymbol{k}}$ by $(2\pi/L)^{3/2}u(\boldsymbol{k})$, $a_{\boldsymbol{k}}$ by $(2\pi/L)^{3/2}a(\boldsymbol{k})$ and the symbol $\sum$ by $\int (L/2\pi)^3 \,\mathrm{d}\boldsymbol{k}$. Thus the characteristic properties of the basis functions [cf. eq. (XXI.35–37)] are written:

$$\int u^*(\boldsymbol{k}; \boldsymbol{r})\, u(\boldsymbol{k}'; \boldsymbol{r}) \,\mathrm{d}\boldsymbol{r} = \delta(\boldsymbol{k}-\boldsymbol{k}') \qquad \int u(\boldsymbol{k}; \boldsymbol{r})\, u^*(\boldsymbol{k}; \boldsymbol{r}') \,\mathrm{d}\boldsymbol{k} = \delta(\boldsymbol{r}-\boldsymbol{r}')$$

$$(\triangle + k^2) u(\boldsymbol{k}; \boldsymbol{r}) = 0$$

$$\omega(k) = (k^2+\mu^2)^{\frac{1}{2}}.$$

The commutation relations for the operators a and $a^\dagger$ are [cf. eq. (XXI.38)]:

$$[a(\boldsymbol{k}), a(\boldsymbol{k}')] = [a^\dagger(\boldsymbol{k}), a^\dagger(\boldsymbol{k}')] = 0$$

$$[a(\boldsymbol{k}), a^\dagger(\boldsymbol{k}')] = \delta(\boldsymbol{k}-\boldsymbol{k}').$$

The fields $\Phi(\boldsymbol{r})$ and $\Pi(\boldsymbol{r})$ are represented by the integrals [cf. eq. (XXI.39) and (XXI.40)]:

$$\Phi(\boldsymbol{r}) = (2\pi)^{-3/2} \int \frac{\mathrm{d}\boldsymbol{k}}{\sqrt{2\omega}} \left(a(\boldsymbol{k})\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} + a^\dagger(\boldsymbol{k})\, \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\right)$$

$$\Pi(\boldsymbol{r}) = (2\pi)^{-3/2}\, \mathrm{i} \int \sqrt{\frac{\omega}{2}}\, \mathrm{d}\boldsymbol{k} \left(a^\dagger(\boldsymbol{k})\, \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} - a(\boldsymbol{k})\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\right)$$

and the Hamiltonian [cf. eq. (XXI.41)] by:

$$H = \int a^{\dagger}(\boldsymbol{k})\, a(\boldsymbol{k})\, \omega(k)\, \mathrm{d}\boldsymbol{k}.$$

It remains to establish the relation between the momentum $\boldsymbol{W}$ of the field and the expansion in plane waves.

By definition the vector operator $\boldsymbol{W}$ is related to infinitesimal transformations by formula (XV.41). In a finite translation $\mathscr{T}(\boldsymbol{a})$ (notations of § XV.9), the operator $\Phi(\boldsymbol{r}_0)$ giving the field amplitude at a given point $\boldsymbol{r}=\boldsymbol{r}_0$ is transformed into the operator $\Phi(\boldsymbol{r}_0+\boldsymbol{a})$ giving the field amplitude at the point obtained from $\boldsymbol{r}_0$ in the translation [1]). Similarly $\Pi(\boldsymbol{r}_0)$ is transformed into $\Pi(\boldsymbol{r}_0+\boldsymbol{a})$. Applying this law for the transformation of observables to the case when $\boldsymbol{a}$ is equal to an infinitesimal vector $\boldsymbol{\epsilon}$, one finds:

$$\begin{aligned} [\boldsymbol{\epsilon}\cdot\boldsymbol{W},\, \Phi(\boldsymbol{r}_0)] &= \mathrm{i}\boldsymbol{\epsilon}\cdot\nabla\Phi|_{\boldsymbol{r}=\boldsymbol{r}_0} \\ [\boldsymbol{\epsilon}\cdot\boldsymbol{W},\, \Pi(\boldsymbol{r}_0)] &= \mathrm{i}\boldsymbol{\epsilon}\cdot\nabla\Pi|_{\boldsymbol{r}=\boldsymbol{r}_0}. \end{aligned} \tag{XXI.50}$$

Since these commutation rules must be satisfied for any $\boldsymbol{r}_0$, they define $\boldsymbol{W}$ to within a constant additive vector. Thus the component W_x along the x axis is the operator defined to within a constant by the commutation relations

$$[W_x, \Phi(\boldsymbol{r}_0)] = \mathrm{i}\left.\frac{\partial\Phi}{\partial x}\right|_{\boldsymbol{r}=\boldsymbol{r}_0} \tag{XXI.51}$$

$$[W_x, \Pi(\boldsymbol{r}_0)] = \mathrm{i}\left.\frac{\partial\Pi}{\partial x}\right|_{\boldsymbol{r}=\boldsymbol{r}_0}. \tag{XXI.52}$$

The preceding relations are satisfied if

$$W_x = -\int \Pi\,\frac{\partial\Phi}{\partial x}\,\mathrm{d}\boldsymbol{r} + Cst. \tag{XXI.53}$$

An integration by parts gives the equivalent expression

$$W_x = \int \frac{\partial\Pi}{\partial x}\,\Phi\,\mathrm{d}\boldsymbol{r} + Cst. \tag{XXI.54}$$

Substituting expression (XXI.53) into the left-hand side of equation

[1]) Note the difference with the law for the transformation of the observable x defined in § XV.9 (cf. note, p. 645). Here $\boldsymbol{r}_0$ is a parameter fixing the point in space where the field amplitude is measured.

(XXI.51) one finds, taking into account commutation relations (XXI.34):

$$\left[-\int \Pi(\boldsymbol{r})\frac{\partial\Phi(\boldsymbol{r})}{\partial x}\,\mathrm{d}\boldsymbol{r},\ \Phi(\boldsymbol{r}_0)\right] = -\int [\Pi(\boldsymbol{r}),\Phi(\boldsymbol{r}_0)]\frac{\partial\Phi(\boldsymbol{r})}{\partial x}\,\mathrm{d}\boldsymbol{r}$$

$$= \mathrm{i}\int \delta(\boldsymbol{r}-\boldsymbol{r}_0)\frac{\partial\Phi(\boldsymbol{r})}{\partial x}\,\mathrm{d}\boldsymbol{r}$$

$$= \mathrm{i}\left.\frac{\partial\Phi}{\partial x}\right|_{\boldsymbol{r}=\boldsymbol{r}_0}.$$

One similarly shows that relation (XXI.52) is satisfied if we substitute for W_x expression (XXI.54).

Analogous expressions can be written down for W_y and W_z. The three components of $\boldsymbol{W}$ are thereby obtained to within a constant that is fixed by requiring that $\boldsymbol{W}$ be a vector operator. We thus find for the momentum of the field the following equivalent expressions:

$$\boldsymbol{W} = -\int \Pi(\nabla\Phi)\,\mathrm{d}\boldsymbol{r} \qquad \text{(XXI.55)}$$

$$= +\int (\nabla\Pi)\,\Phi\,\mathrm{d}\boldsymbol{r}. \qquad \text{(XXI.56)}$$

[*N.B.* The classical theory leads to a formally identical definition of the momentum.]

Let us now express $\boldsymbol{W}$ in terms of the $a_{\boldsymbol{k}}$ and $a_{\boldsymbol{k}}^{\dagger}$ operators. To this effect, we substitute for Φ and Π in the right-hand side of equation (XXI.55) their respective expansions [eqs. (XXI.46) and (XXI.47)] and make use of the orthogonality relations for the plane waves. We obtain:

$$\boldsymbol{W} = \tfrac{1}{2}\sum_{\boldsymbol{k}} \boldsymbol{k}(a_{\boldsymbol{k}}^{\dagger}a_{\boldsymbol{k}} + a_{\boldsymbol{k}}a_{\boldsymbol{k}}^{\dagger} - a_{\boldsymbol{k}}a_{-\boldsymbol{k}} - a_{-\boldsymbol{k}}^{\dagger}a_{\boldsymbol{k}}^{\dagger})$$

or, taking into account the commutation relations for the a and $a^{\dagger}$,

$$\boldsymbol{W} = \sum_{\boldsymbol{k}} (a_{\boldsymbol{k}}^{\dagger}a_{\boldsymbol{k}})\,\boldsymbol{k} + \tfrac{1}{2}\sum_{\boldsymbol{k}} \boldsymbol{k}(1 - a_{-\boldsymbol{k}}a_{\boldsymbol{k}} - a_{-\boldsymbol{k}}^{\dagger}a_{\boldsymbol{k}}^{\dagger}).$$

The symbols $\sum_{\boldsymbol{k}}$ indicate summations over all possible values of $\boldsymbol{k}$. In the second sum, the expression in parentheses does not change when one changes $\boldsymbol{k}$ to $-\boldsymbol{k}$; therefore, the two terms that correspond to a given value of $\boldsymbol{k}$ and its negative differ by a sign and the second sum vanishes. Thus:

$$\boldsymbol{W} = \sum_{\boldsymbol{k}} (a_{\boldsymbol{k}}^{\dagger}a_{\boldsymbol{k}})\,\boldsymbol{k}. \qquad \text{(XXI.57)}$$

From this expression one easily obtains the commutation relations of $\boldsymbol{W}$ with the a and the $a^\dagger$:

$$[\boldsymbol{W}, a_{\boldsymbol{k}}] = -\boldsymbol{k} a_{\boldsymbol{k}} \tag{XXI.58}$$

$$[\boldsymbol{W}, a_{\boldsymbol{k}}^\dagger] = +\boldsymbol{k} a_{\boldsymbol{k}}^\dagger. \tag{XXI.59}$$

Expressions (XXI.57–59) have a simple interpretation if we postulate that a corpuscle in state $u_{\boldsymbol{k}}$ has momentum $\boldsymbol{k}$. Since $(a_{\boldsymbol{k}}^\dagger a_{\boldsymbol{k}})$ is the number of corpuscles in the state $u_{\boldsymbol{k}}$, formula (XXI.57) means simply that the total momentum $\boldsymbol{W}$ of the field is equal to the sum of the momenta of each of its constituent corpuscles. Similarly, relations (XXI.58) and (XXI.59) are consistent with the interpretation of $a_{\boldsymbol{k}}$ and $a_{\boldsymbol{k}}^\dagger$ respectively as operators of annihilation and creation of a corpuscle of momentum $\boldsymbol{k}$. Indeed if $|\boldsymbol{w}\rangle$ is an eigenvector of the total momentum corresponding to the eigenvalue $\boldsymbol{w}$; then from equation (XXI.58) the vector $a_{\boldsymbol{k}}|\boldsymbol{w}\rangle$ (if it is not a null vector, that is, if the field contains at least one corpuscle of momentum $\boldsymbol{k}$) satisfies the eigenvalue equation:

$$\begin{aligned}\boldsymbol{W} a_{\boldsymbol{k}} |\boldsymbol{w}\rangle &= a_{\boldsymbol{k}}(\boldsymbol{W}-\boldsymbol{k}) |\boldsymbol{w}\rangle = a_{\boldsymbol{k}}(\boldsymbol{w}-\boldsymbol{k}) |\boldsymbol{w}\rangle \\ &= (\boldsymbol{w}-\boldsymbol{k}) a_{\boldsymbol{k}} |\boldsymbol{w}\rangle.\end{aligned}$$

In the same way we obtain from equation (XXI.59):

$$\boldsymbol{W} a_{\boldsymbol{k}}^\dagger |\boldsymbol{w}\rangle = (\boldsymbol{w}+\boldsymbol{k}) a_{\boldsymbol{k}}^\dagger |\boldsymbol{w}\rangle.$$

It will be noted that the corpuscles of momentum $\boldsymbol{k}$ contained in the field have an energy $\omega = (k^2+\mu^2)$. *They are therefore corpuscles of mass* μ.

7. Spherical Waves. Definition of the Angular Momentum

The plane waves are the basis functions most often used. However, one can just as well use spherical waves. The basis functions are then of the form

$$u_{klm} \equiv Cst.\ j_l(kr)\ Y_l^m(\theta, \varphi) \tag{XXI.60}$$

and depend on the continuous index k and the two discrete indices l and m. In order to have only discrete indices, we can enclose the field in a sphere of radius R and require that the basis functions vanish at the surface; the correct results are obtained by letting R become indefinitely large.

The entire analysis of the preceding paragraph can be repeated with these new basis functions. One thus introduces the annihilation and creation operators a_{klm} and $a^{\dagger}_{klm}$ of a corpuscle in the state represented by the spherical wave u_{klm}. It can be shown that the spherical waves are related to individual states of well-defined angular momentum just as the plane waves are related to individual states of well-defined momentum. For the total angular momentum $\boldsymbol{l}$ of the field, one obtains the expression:

$$\boldsymbol{l} = -\int \Pi(\boldsymbol{l}\Phi)\,\mathrm{d}\boldsymbol{r} \qquad \text{(XXI.61)}$$

$\boldsymbol{l}$ here represents the operator $-\mathrm{i}(\boldsymbol{r}\times\boldsymbol{\nabla})$ and acts on the function Φ to its right:

$$\boldsymbol{l}\Phi = -\mathrm{i}\boldsymbol{r}\times(\boldsymbol{\nabla}\Phi).$$

We can also use the equivalent expression:

$$\boldsymbol{l} = \int (\boldsymbol{l}\Pi)\,\Phi\,\mathrm{d}\boldsymbol{r}.$$

The a_{klm} and $a^{\dagger}_{klm}$ are respectively interpreted as annihilation and creation operators of a corpuscle of energy $\omega=(k^2+\mu^2)^{\frac{1}{2}}$ and of angular momentum (lm). In addition the $(2l+1)$ operators $a^{\dagger}_{klm}$ corresponding to a given value of l and of k are the standard components of an irreducible tensor operator of order l; the $(2l+1)$ operators a_{klm} are, except for order and sign, the standard components of the Hermitean conjugate tensor operator (definition of § XIII.30).

8. Space and Time Reflections

Among the transformations which leave the equations of motion invariant one also has the space and time reflections. Each of these is completely defined by giving the law for the transformation of the field $\Phi(\boldsymbol{r})$ and its conjugate momentum $\Pi(\boldsymbol{r})$; the corresponding (unitary or antiunitary) transformation operator is then determined up to a phase.

In the case of the *space reflection*, the law in question is

$$\Phi(\boldsymbol{r}) \to \Phi(-\boldsymbol{r}), \qquad \Pi(\boldsymbol{r}) \to \Pi(-\boldsymbol{r}). \qquad \text{(I)}$$

This law is characteristic of the scalar field. The operator S_0 corresponding to this transformation is *unitary* (the commutation relations for the fields are therefore conserved in the transformation). Since all of the operators of the theory are functions of Φ and Π, it is easy to deduce from (I) their law of transformation in the space reflection. In particular, for the a and $a^{\dagger}$ operators associated

with the plane waves, one finds, in accordance with the interpretation given to these operators:

$$S_0 a_k S_0^\dagger = a_{-k} \qquad (II)$$
$$S_0 a_k^\dagger S_0^\dagger = a_{-k}^\dagger.$$

In the case of *time reversal*, the field Φ is conserved and its velocity changes sign, the law is therefore:

$$\Phi(r) \to \Phi(r), \qquad \Pi(r) \to -\Pi(r). \qquad (I')$$

The operator K corresponding to this transformation is *antiunitary*, since the commutation relations for the fields change sign in the transformation. The law for the transformation of the a and $a^\dagger$ operators associated with the plane waves can be obtained without difficulty from (I'); one obtains, in accordance with the interpretation given to these operators:

$$K a_k K^\dagger = a_{-k} \qquad (II')$$
$$K a_k^\dagger K^\dagger = a_{-k}^\dagger.$$

These creation and absorption operators therefore transform in an identical way in space reflection and time reversal. One must take care, however, not to confuse these two transformations the one of which is unitary, the other antiunitary.

Relations (I) or (II) and relations (I') or (II') respectively define S_0 and K to within a phase. To remove the arbitrary in the phase, we shall require the vector $|0\rangle$ to be invariant in each of these transformations:

$$S_0 |0\rangle = |0\rangle, \qquad K |0\rangle = |0\rangle.$$

II. COUPLING WITH AN ATOMIC SYSTEM

9. Coupling to a System of Particles

We now consider the field in interaction with a particle. Most of what follows would not be modified if this single particle were replaced by a system of several particles.

The dynamical variables of the whole system, particle+field, are functions of the fundamental variables of each subsystem. For the particle, which we assume spinless for simplicity, we shall take as fundamental variables its position vector $\boldsymbol{R}$ and its momentum $\boldsymbol{P}$. For the fundamental variables of the field, we may take $\Phi(r)$ and $\Pi(r)$ or we may take a complete set of creation and annihilation operators like those defined in the preceding paragraph; in what

follows we shall often use the operators a_k and $a_k^\dagger$ associated with the decomposition into plane waves (cf. § 6).

The states of the whole system span a space that is the product of the space $\mathscr{E}_{\text{part}}$ of the states of the particle alone and the space $\mathscr{E}_{\text{field}}$ of the states of the free field:

$$\mathscr{E} = \mathscr{E}_{\text{part}} \otimes \mathscr{E}_{\text{field}}$$

The Hamiltonian H of the total system is the sum of the three terms

$$H = H_{\text{part}} + H_{\text{field}} + H'. \tag{XXI.62}$$

The first two terms are those which would govern the evolution of each of the two subsystems if they were completely decoupled, the last represents the interaction energy.

Thus H_{field} is the Hamiltonian for the free field; several equivalent expressions have been given above; we note in particular equation (XXI.48). H_{part} is the Hamiltonian which would govern the evolution of the particle in the absence of the field Φ. Let us suppose that the particle in question is placed in a potential and denote its mass by M. We assume that

$$M \gg \mu$$

and that the motion of the particles may be treated in the non-relativistic approximation, i.e.:

$$H_{\text{part}} = \frac{\boldsymbol{P}^2}{2M} + V(R). \tag{XXI.63}$$

It remains to define H'. The simplest possible expression for H' is obtained by assuming the interaction proportional to the value taken by the field amplitude at the position of the particle $\boldsymbol{R}$:

$$H' = g\Phi(\boldsymbol{R}) \tag{XXI.64}$$

$$= gL^{-3/2} \sum_{\boldsymbol{k}} \frac{1}{\sqrt{2\omega_k}} (a_{\boldsymbol{k}}\, e^{i\boldsymbol{k}\cdot\boldsymbol{R}} + a_{\boldsymbol{k}}^\dagger\, e^{-i\boldsymbol{k}\cdot\boldsymbol{R}}). \tag{XXI.64'}$$

g is a dimensionless constant called the *coupling constant*. The second expression for H' is obtained from the first by using expansion (XXI.46).

This particular form of the interaction is almost uniquely determined by considerations of relativistic invariance. However, the theory to be investigated here contains from the start a non-relativistic approxi-

mation. This is implied not only because the term H_{part} lacks the covariant properties required by the relativity principle, but also because the very concept of a material system of a single particle — and more generally that of a system made up of a definite number of particles — ceases to be justified in Relativistic Quantum Mechanics. These two limitations of the theory must be kept in mind in choosing H'. The study made in Chapter XX of the non-relativistic approximation of the Dirac equation can here be used as a guide. It suggests that the interaction between a non-relativistic particle of mass M and a field is not a local interaction like that given by equation (XXI.64) but an interaction involving the values taken by the field in a domain of extension $1/M$ about the position of the particle. This suggests that we replace expression (XXI.54) for H' by the expression:

$$H' = g \int \Phi(\boldsymbol{r})\, \varrho(|\boldsymbol{r}-\boldsymbol{R}|)\, \mathrm{d}\boldsymbol{r}, \tag{XXI.65}$$

where $\varrho(\boldsymbol{r}) \equiv \varrho(r)$ is a real, spherically symmetrical function obeying the normalization condition

$$\int \varrho(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r} \equiv 4\pi \int_0^\infty \varrho(r) r^2\, \mathrm{d}r = 1,$$

and concentrated in a domain of radius $1/M$ about the origin (cf. Fig. XXI.1*a*). We obtain (XXI.64) by substituting $\delta(\boldsymbol{r}-\boldsymbol{R})$ for this function in the right-hand side of equation (XXI.65).

We shall give the name of cut-off function to the function $C(k)$ defined by

$$C(k) = \int \mathrm{e}^{+\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\, \varrho(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r}.$$

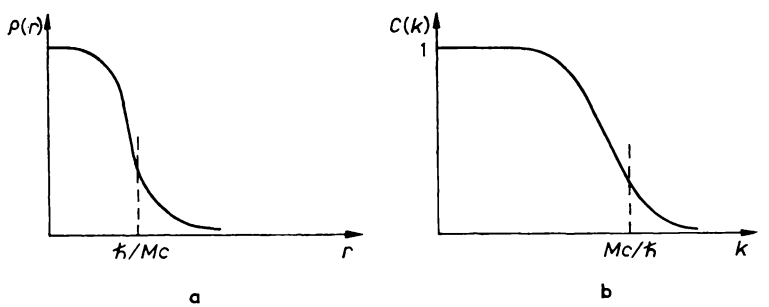


Fig. XXI.1. Form of $\varrho(r)$ and $C(k)$.

It is a real, symmetrical function, obeying the condition

$$C(0)=1,$$

and, according to well-known properties of Fourier transforms, concentrated in a domain of radius M about the point $k=0$ (cf. Fig. XXI.1*b*). Substituting expansion (XXI.46) for $\Phi(r)$ in the right-hand side of equation (XXI.65), one obtains, after a brief calculation:

$$H' = gL^{-3/2} \sum_k \frac{C(k)}{\sqrt{2\omega_k}} (a_k \, e^{ik\cdot R} + a_k^{\dagger} \, e^{-ik\cdot R}). \qquad \text{(XXI.66)}$$

This expansion differs from (XXI.64′) only by the presence of the factor $C(k)$ in each of its terms. This factor cuts off the contribution of high-frequency terms to the interaction, that is, the contribution of terms for which $k \gtrsim M$.

As will now be shown, this suppression of the high frequencies is consistent with the non-relativistic approximation. Each term of expansion (XXI.66) corresponds to a given momentum and energy transfer from the particle to the field and *vice versa*; the term $a_k e^{ik\cdot R}$ absorbs a field corpuscle of momentum k and energy ω_k and transfers this momentum and energy from the field to the particle; the term $a_k^{\dagger} e^{-ik\cdot R}$ creates a field corpuscle of momentum k and energy ω_k, diminishing by as much the momentum and energy of the particle. When $k \gtrsim M$, the transfers are sufficiently large that the law of conservation of the number of particles of mass M can be violated and the field can absorb the particle or create a second particle of mass M. The non-relativistic approximation is therefore only justified if the contribution from the high-frequency terms is negligible, and the results that it gives, if they are correct, must not be modified by the introduction of $C(k)$. *A fortiori*, they must be insensitive to the form of $C(k)$. In the following calculations we shall suppose that

$$C(k) = \begin{cases} 1 & \text{for} \quad 0 \leqslant k \leqslant K \\ 0 & \text{for} \quad k > K \end{cases} \qquad \text{(XXI.67)}$$

and shall give to the parameter K a value of the order of M.

Finally, let us examine the invariance properties of H. Clearly H', as defined by equation (XXI.65), is invariant with respect to any displacement of the whole system (field + particle) since it may easily be shown that it commutes with the total momentum $\boldsymbol{P} + \boldsymbol{W}$ and

with the total angular momentum $(\boldsymbol{R}\times\boldsymbol{P})+\boldsymbol{l}$ of the system. H' is also invariant under reflection if $\Phi(\boldsymbol{r})$ transforms into $\Phi(-\boldsymbol{r})$ in that operation (scalar field). H_{field} also has all of these invariance properties. If $V(R)=0$ the same is true of H_{part}, and therefore of H. If $V(R)\neq 0$, H_{part} is only invariant with respect to rotations and reflections; H is then also invariant under rotation and reflection, but not under translation.

The interest of the quantum system considered here is due to the fact that it exhibits in a simplified form the essential properties of an atom in interaction with the field of electromagnetic radiation; the particle here plays the part of the atom, the scalar field that of the electromagnetic radiation. The main difference is that the corpuscles of the electromagnetic field, the photons, have mass zero and spin 1, while the corpuscles of the field studied here have spin zero and a non-zero mass. In the rest of this section this simplified model will be used to study some of the typical properties of the atom in the presence of the electromagnetic field.

10. Weak Coupling and Perturbation Treatment

For g sufficiently small, H' may be treated as a small perturbation and the methods of Chapters XVI and XVII applied.

The unperturbed Hamiltonian is:

$$H_0 \equiv H_{\text{field}} + H_{\text{part}}.$$

In perturbation theory one employs a representation in which H_0 is diagonal. We shall adopt the representation whose basis vectors are obtained in the following way. Denote by $|\alpha\rangle, |\beta\rangle, \ldots |\lambda\rangle \ldots$ a complete orthonormal set of eigenvectors of H_{part} considered as an operator in $\mathscr{E}_{\text{part}}$, and by $E_\alpha, E_\beta, \ldots, E_\lambda, \ldots$ the corresponding eigenenergies (arranged in increasing order). By multiplying a given vector $|\lambda\rangle$ of this set by the vector $|0\rangle$ representing the corpuscular vacuum in $\mathscr{E}_{\text{field}}$, one obtains a vector of $\mathscr{E}$, which to simplify the writing will also be denoted by $|\lambda\rangle$. It is clear that:

$$H_0\,|\lambda\rangle = E_\lambda\,|\lambda\rangle.$$

A complete orthonormal set of eigenvectors of H_{field} in $\mathscr{E}_{\text{field}}$ can be obtained by repeated action of the creation operators $a_k^\dagger$ on the vector $|0\rangle$ [cf. eq. (XXI.16)]. In the same way a complete orthonormal

set of eigenvectors of H_0 in $\mathscr{E}$ can be obtained by repeated action of the creation operators $a_k^\dagger$ on each of the vectors $|\alpha\rangle$, $|\beta\rangle$, ..., $|\lambda\rangle$, ... defined above. This will be taken as the required basis. Denote by $|n\rangle$ a typical one of these basis vectors and by E_n the corresponding eigenenergy:

$$H_0 |n\rangle = E_n |n\rangle.$$

In the dynamical state $|n\rangle$, the particle is in a given eigenstate of H_{part} and the field contains a given number of corpuscles of well-defined momentum.

In practice, only the vectors corresponding to a very limited number of corpuscles will be involved in what follows; they can be labelled by the Greek letter indicating the dynamical state of the particle, and the wave vectors for each of the corpuscles in the field. With this notation $|\lambda \boldsymbol{k}\rangle$ represents the state obtained by associating a corpuscle of momentum $\boldsymbol{k}$ with the particle in the state λ, $|\lambda \boldsymbol{k}\boldsymbol{k}'\rangle$ the state obtained by associating the two corpuscles of momentum $\boldsymbol{k}$ and $\boldsymbol{k}'$ respectively with a particle in the state λ, etc. Note that:

$$\begin{aligned} H_0 |\lambda \boldsymbol{k}\rangle &= (E_\lambda + \omega_k) |\lambda \boldsymbol{k}\rangle \\ H_0 |\lambda \boldsymbol{k}\boldsymbol{k}'\rangle &= (E_\lambda + \omega_k + \omega_{k'}) |\lambda \boldsymbol{k}\boldsymbol{k}'\rangle. \end{aligned} \tag{XXI.68}$$

Also

$$|\lambda \boldsymbol{k}\rangle = a_k^\dagger |\lambda\rangle$$

and if $\boldsymbol{k} \neq \boldsymbol{k}'$,

$$|\lambda \boldsymbol{k}\boldsymbol{k}'\rangle = a_k^\dagger a_{k'}^\dagger |\lambda\rangle.$$

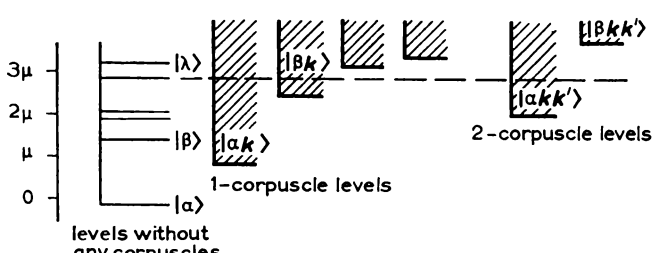


Fig. XXI.2. Typical diagram of the first levels of H_0.

A typical spectroscopic diagram giving the lowest levels of the Hamiltonian H_0 is shown in Figure XXI.2. Following the usual convention for this type of diagram (cf. Fig. XI.1, XII.1, etc.), the height of the levels above the ground state is equal to the difference

between their energy and that of the ground state; levels of the same energy are consequently at the same height. The levels in the figure are classified according to the nature of the corresponding eigenstates. In the left-hand column will be found the levels of states without any corpuscle. Next come the states with a single corpuscle, in several columns each of which corresponds to a given quantum state of the particle. The first corresponds to levels $|\alpha \boldsymbol{k}\rangle$, where the particle is in the ground state; it is a continuous band (in the limit when $L \to \infty$), starting from a minimum corresponding to the existence of a single corpuscle of zero momentum, and consequently situated at a distance μ above the level $|\alpha\rangle$. The second column corresponds to states $|\beta \boldsymbol{k}\rangle$ for which the particle is in its first excited state. And so forth. After the 1-corpuscle levels we have the 2-corpuscle levels, also arranged in a series of columns corresponding each to a well-defined quantum state of the particle. And so forth.

The perturbation H' couples these various levels. In the representation defined above this coupling term is represented by a particularly simple matrix. From (XXI.66) we see that *the only non-vanishing matrix elements of H' are those between basis vectors for which the number of corpuscles differs by one*. These matrix elements may easily be calculated using the properties of the creation and annihilation operators. One finds:

$$\langle \lambda | H' | \nu \boldsymbol{k} \rangle = \langle \nu \boldsymbol{k} | H' | \lambda \rangle^* = \left(\frac{2\pi}{L} \right)^{3/2} g \frac{C(k)}{\sqrt{2\omega_k}} \langle \lambda | U(\boldsymbol{k}) | \nu \rangle \quad \text{(XXI.69}a\text{)}$$

$$\begin{aligned} \langle \lambda \boldsymbol{k}' | H' | \nu \boldsymbol{k} \rangle &= \langle \nu \boldsymbol{k} | H' | \lambda \boldsymbol{k}' \rangle^* \\ &= \left(\frac{2\pi}{L} \right)^{3/2} g \frac{C(k)}{\sqrt{2\omega_k}} \langle \lambda | U^\dagger(\boldsymbol{k}') | \nu \rangle \quad (\boldsymbol{k} \neq \boldsymbol{k}'). \end{aligned} \quad \text{(XXI.69}b\text{)}$$

[If $\boldsymbol{k} = \boldsymbol{k}'$, a factor $\sqrt{2}$ must be added in the last of these expressions.] Here we have put:

$$U(\boldsymbol{k}) \equiv \frac{e^{i\boldsymbol{k}\cdot\boldsymbol{R}}}{(2\pi)^{3/2}}. \quad \text{(XXI.70)}$$

No matter how weak the coupling, it leads to a qualitative modification of the spectrum, since most of the bound states become unstable and can make transitions to states of lower energy with emission of one or several field-corpuscles. For example, consider the state $|\lambda\rangle$ indicated in Fig. XXI.3; it is at the same level as certain states of

the continuum, namely the 1-corpuscle states of type $|\alpha \boldsymbol{k}\rangle$ or $|\beta \boldsymbol{k}\rangle$ and the 2-corpuscle states of type $|\alpha \boldsymbol{k}\boldsymbol{k}'\rangle$. The introduction of the small perturbation H' couples the discrete eigenstates to the 1-corpuscle states, the 1-corpuscle states to the 2-corpuscle states, etc., and therefore makes possible "radiative" transitions from state $|\lambda\rangle$ to states of the continuum situated at the same level; these possible decay modes of the state $|\lambda\rangle$ are schematically shown in Fig. XXI.3. The only bound states that remain stable are those for which the energy above the ground state is too small for the emission of a corpuscle

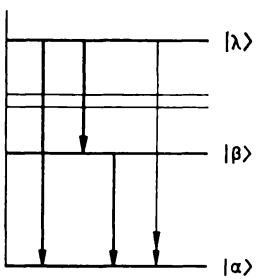


Fig. XXI.3. Radiative transitions from the level $|\lambda\rangle$. The double arrow on the right represents a transition to the ground state with emission of two corpuscles. It is a higher-order transition.

of mass μ to be energetically possible, that is, states whose energy is less than $E_\alpha + \mu$. In the example of Fig. XXI.2 the only stable state is the ground state.

The existence of the coupling term H' also leads to a certain shift of the levels of the bound states. It is, as we shall see, a second-order effect. It may nevertheless affect the stability and the decay modes of the different states of the particle since a small level shift may be enough to make certain "radiative" transitions energetically possible that previously were not, or to have the inverse effect.

11. Level Shifts

As a first application, let us calculate the level shifts mentioned at the end of the preceding paragraph. This simple perturbation calculation exhibits the main difficulties of field theory and provides some insight into the physical significance and the limitations of the theory.

We therefore consider a stable level, the ground state level for example, and we calculate the shift due to the existence of the coupling

term H' by applying the method of stationary perturbations. For simplicity, we assume the level in question non-degenerate. Since H' has all the invariance properties of H_{part}, the treatment of degenerate levels is not fundamentally different. To the first order, the shift is given by formula (XVI.12). It follows from the properties of H' given above that:

$$\langle\alpha|H'|\alpha\rangle = 0.$$

The calculations must therefore be taken to the second order. Denoting the second-order correction by δE_α, one has (cf. § XVI.6):

$$\delta E_\alpha = \langle\alpha|H' \frac{1}{E_\alpha - H_0} H'|\alpha\rangle. \tag{XXI.71}$$

We calculate this expression in the representation defined above. The only intermediate states involved are those having a single corpuscle. Taking relation (XXI.68), (XXI.69*a*) and (XXI.70) into account, we successively obtain:

$$\begin{aligned}\delta E_\alpha &= \sum_n \langle\alpha|H'|n\rangle \frac{1}{E_\alpha - E_n} \langle n|H'|\alpha\rangle \\ &= \sum_\nu \sum_{\boldsymbol{k}} \frac{|\langle\alpha|H'|\nu\boldsymbol{k}\rangle|^2}{E_\alpha - E_\nu - \omega_k} \\ &= g^2 \left(\frac{2\pi}{L}\right)^3 \sum_\nu \sum_{\boldsymbol{k}} \frac{C^2(k)}{2\omega_k} \frac{|\langle\alpha|U_{\boldsymbol{k}}|\nu\rangle|^2}{E_\alpha - E_\nu - \omega_k}\end{aligned}$$

whence, replacing the summation over $\boldsymbol{k}$ by an integration in accordance with the technique described in § 6, and adopting the cut-off function given by equation (XXI.67):

$$\delta E_\alpha = -\, g^2 \int_{k \leqslant K} \mathrm{d}\boldsymbol{k} \sum_\nu \frac{|\langle\alpha|U(\boldsymbol{k})|\nu\rangle|^2}{2\omega(\omega + E_\nu - E_\alpha)}. \tag{XXI.72}$$

Since we are dealing with the ground state, the energy denominator is positive for all of the terms in the sum on the right hand side of (XXI.72). We therefore have:

$$\delta E_\alpha < 0,$$

An upper limit for $|\delta E_\alpha|$ can be obtained by replacing the quantity $(E_\nu - E_\alpha)$ in the denominator by its minimum value, that is, by 0

(E_α is the ground state); the summation over ν is then easily effected using the closure relation: $\sum_\nu |\nu\rangle\langle\nu| = 1$; since $UU^\dagger = (2\pi)^{-3}$, this sum gives $(2\pi)^{-3}/2\omega^2$. We thus obtain, after integration over the angles

$$|\delta E_\alpha| < \frac{g^2}{4\pi^2} \int_0^K \frac{k^2}{\omega^2}\, \mathrm{d}k. \tag{XXI.73}$$

The integral in (XXI.73) can be calculated exactly; since $K \gg \mu$, it is slightly smaller than K. Consequently:

$$|\delta E_\alpha| < \frac{g^2 K}{4\pi^2}. \tag{XXI.74}$$

A better estimate of the right-hand side of (XXI.72) can be obtained in the following manner. If the closure relation is directly applied, one obtains, replacing $U(\boldsymbol{k})$ by its definition [eq. (XXI.70)]:

$$\delta E_\alpha = -\frac{g^2}{8\pi^3} \int_{k \leqslant K} \langle\alpha|\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}} (\omega + H_{\mathrm{part}} - E_\alpha)^{-1}\, \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}}\, |\alpha\rangle \frac{\mathrm{d}\boldsymbol{k}}{2\omega}. \tag{XXI.75}$$

The unitary operator $\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}}$ commutes with $\boldsymbol{R}$ and transforms $\boldsymbol{P}$ into $(\boldsymbol{P}-\boldsymbol{k})$. It follows that:

$$\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}}\, H_{\mathrm{part}}\, \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}} = \frac{(\boldsymbol{P}-\boldsymbol{k})^2}{2M} + V(R) = H_{\mathrm{part}} + \frac{k^2}{2M} - \frac{\boldsymbol{P}\cdot\boldsymbol{k}}{M}$$

whence:

$$\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}}\, (\omega + H_{\mathrm{part}} - E_\alpha)^{-1}\, \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}} = \left(\omega + \frac{k^2}{2M} + H_{\mathrm{part}} - E_\alpha - \frac{\boldsymbol{P}\cdot\boldsymbol{k}}{M}\right)^{-1}.$$

Substituting this expression into the right-hand side of (XXI.75), we find:

$$\delta E_\alpha = -\frac{g^2}{8\pi^3} \int_{k \leqslant K} \langle\alpha| \left(\omega + \frac{k^2}{2M} + H_{\mathrm{part}} - E_\alpha - \frac{\boldsymbol{P}\cdot\boldsymbol{k}}{M}\right)^{-1} |\alpha\rangle \frac{\mathrm{d}\boldsymbol{k}}{2\omega}.$$

The term $\boldsymbol{P}\cdot\boldsymbol{k}/M$ in the bracket on the right-hand side can be neglected since the average value of the velocity $\boldsymbol{P}/M$ of the particle in the state $|\alpha\rangle$ is much smaller than 1 (non-relativistic approximation). With this approximation we obtain, after integration over the angles:

$$\delta E_\alpha \simeq -\frac{g^2}{4\pi^2} \int_0^K \frac{k^2}{\omega^2} \left(1 + \frac{k^2}{2M\omega}\right)^{-1} \mathrm{d}k, \tag{XXI.76}$$

an expression which differs from the right-hand side of (XXI.73) only by the factor $[1+(k^2/2M\omega)]^{-1}$ in the integrand. Since

$\mu \ll K \approx M$, this factor falls roughly from 1 to $\frac{2}{3}$ over the interval of integration. We may therefore write:

$$\delta E_\alpha = -\frac{sg^2K}{4\pi^2} \qquad (s \lesssim 1). \tag{XXI.77}$$

The numerical constant s appearing in this formula has a value between $\frac{2}{3}$ and 1.

To estimate the size of the effect, we calculate its order of magnitude in the case when the various quantities in the model have numerical values of the order of those encountered in atoms [1]). M is then the mass of the electron and g^2 the fine-structure constant:

$$g^2 \simeq 10^{-1}. \tag{XXI.78}$$

Denote the distance from the level to its nearest neighbor by D/M:

$$\frac{D}{M} \simeq 10^{-4} - 10^{-5}. \tag{XXI.79}$$

With these figures,

$$|\delta E_\alpha/D| \approx g^2M/4\pi^2D \approx 100.$$

It is thus a very large effect, much larger than the shifts actually observed. This might lead one to question the validity of the perturbation treatment and the physical significance of its results.

However, a meaningful comparison between theory and experiment should take into account the following point. The mass M involved in a calculation of the unperturbed levels *is not* the mass that is experimentally observed. The latter is obtained by measuring the energy and the momentum of a "free" particle, that is, of a particle outside the potential $V(R)$. Such a "free" particle is nevertheless coupled to the field and therefore the measurement gives:

$$M_{\text{exp}} = M + \delta M,$$

where δM represents the contribution to the rest energy of the "free"

---

[1]) Since the forms of the couplings are not the same, the numerical results given by the scalar theory cannot be in exact agreement with those given by the Quantum Theory of Radiation. In order to compare comparable quantities, the constant g^2 is fixed so as to give a value for the dipole transition probability of the same order of magnitude as that for the dipole transitions in atoms [cf. eqs. (XXI.81) and (XXI.245)]; namely $g^2/4\pi \approx 1/137$.

particle due to the presence of the field. In the present case, a brief calculation gives

$$\delta M \approx -\frac{g^2 K}{4\pi^2}.$$

More precisely, the calculation of δM to the second order in the perturbation gives the right-hand side of equation (XXI.76) (Problem XXI.2); the shift δE_α is therefore almost entirely due to this "renormalization" of the mass [1]).

When the levels of the hydrogen atom are calculated with the Schrödinger or Dirac theories, one makes use of the experimental mass of the electron; owing to this, one accounts for the main part of the coupling of the electron with the radiation field and this explains why the calculated spectrum is so remarkably in agreement with the observed spectrum.

The "experimentally observed shift" $\delta E_\alpha'$ is the difference between the calculated shift δE_α and the shift $\delta E_\alpha^{(0)}$ obtained by merely replacing the theoretical mass by the experimental mass (also called the "renormalized" mass) in the particle Hamiltonian. In our model, this substitution of the "renormalized" mass consists in replacing H_{part} by

$$\delta M + \frac{\boldsymbol{P}^2}{2(M+\delta M)} + V(R)$$

and leads essentially to an overall displacement of all the levels by the quantity δM. From this we obtain:

$$\delta E_\alpha^{(0)} = \delta M \left(1 - \langle\alpha| \frac{\boldsymbol{P}^2}{2M^2} |\alpha\rangle\right).$$

If we stop at the second order in the perturbation method, the calculation of $\delta E_\alpha' = \delta E_\alpha - \delta E_\alpha^{(0)}$ presents no serious difficulties and gives for $|\delta E_\alpha'/D|$ a value of the order of 10^{-3}, which constitutes a reasonable order of magnitude.

The value of $\delta E_\alpha'$ thereby obtained must nevertheless be regarded with caution, since it proves to be very sensitive to the choice of the

---

[1]) Putting the cut-off at $K \approx M$, we find $\delta M/M \approx 10^{-3}$. The relative modification of the mass is therefore very small, which justifies *a posteriori* the use of the perturbation method. However, if the cut-off function is not introduced, δM is given by a divergent integral. We shall return to this difficulty later.

cut-off function. A realistic calculation must therefore start from a completely relativistic theory. In fact, we have here a difficult problem in Quantum Field Theory for which at the present time there is no completely satisfactory solution. In the Relativistic Theory, the interaction is entirely local and one obtains for δM an expression of the form: $\delta M = ZM$, where Z is a divergent integral (cf. note, p. 990). Similarly δE_α and $\delta E_\alpha{}^{(0)}$ are given by divergent integrals so that the expression for $\delta E_\alpha{}'$ is an indeterminate form of the type $\infty - \infty$.

In practice, the indetermination is removed by an *ad hoc* prescription based on considerations of relativistic invariance. One can thereby obtain results in remarkable agreement with experiment [1]). However, a rigorous justification for such calculational methods remains to be found.

Keeping in mind these limitations, we shall nevertheless continue our study of this simple model which correctly accounts for a large body of experimental facts. Effects such as the mass "renormalization" which can satisfactorily be treated only within the framework of the covariant formalism will be accounted for phenomenologically when they appear.

12. Emission of a Corpuscle

Let us turn now to the states which are unstable due to the coupling H'.

In this paragraph, we calculate the transition probabilities per unit time of the various "radiative transitions" by applying the time-dependent perturbation method (§ XVII.4).

To be definite, suppose that the spectrum of H_0 is that shown in Figure XXI.2 and consider the state $|\lambda\rangle$. The only transitions energetically possible (Fig. XXI.3) are the transitions to the ground state with emission of one or two corpuscles ($3\mu > E_\lambda - E_\alpha > 2\mu$) and transitions to the first excited state with emission of a single corpuscle ($2\mu > E_\lambda - E_\beta > \mu$). In what follows, we shall consider only transitions with emission of a single corpuscle. The emission of two corpuscles is a higher-order process and, other things being equal, g^2 times less frequent than the emission of a single corpuscle; it must therefore be left aside if we go only to the first order in the perturbation calculation.

The particle being initially in the state $|\lambda\rangle$, the probability per

[1]) Cf. note, p. 933.

unit time for it to make a transition to the state $|\alpha\rangle$ with emission of a corpuscle in the solid angle $(\Omega, \Omega+\mathrm{d}\Omega)$, i.e., the probability per unit time for the transition $\lambda \to \alpha\boldsymbol{k}$, where $\boldsymbol{k}$ is a vector in this solid angle, is given to the first order by the expression [cf. eq. (XVII.50)]:

$$w_{\lambda\to\alpha\boldsymbol{k}}\,\mathrm{d}\Omega = 2\pi\,|\langle\lambda|H'|\alpha\boldsymbol{k}\rangle|^2\varrho_L(\omega)\,\mathrm{d}\Omega, \tag{XXI.80}$$

in which the energy of the corpuscle, ω, obeys the conservation law:

$$\omega = \omega_{\lambda\alpha} \equiv E_\lambda - E_\alpha.$$

Taking into account relations (XXI.49), (XXI.69a) and (XXI.70), formula (XXI.80) gives

$$w_{\lambda\to\alpha\boldsymbol{k}} = 2\pi g^2\,\frac{|\langle\lambda|\,\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}}\,|\alpha\rangle|^2}{2\omega}\,\frac{\omega k}{(2\pi)^3} = \frac{g^2k}{8\pi^2}\,|\langle\lambda|\,\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}}\,|\alpha\rangle|^2. \tag{XXI.81}$$

By integrating over the angles, one obtains the probability of transition per unit time from λ to α, which we denote by $\Gamma_{\lambda\to\alpha}$:

$$\Gamma_{\lambda\to\alpha} \equiv \int w_{\lambda\to\alpha\boldsymbol{k}}\,\mathrm{d}\Omega = \frac{g^2k}{8\pi^2}\int|\langle\lambda|\,\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}}\,|\alpha\rangle|^2\,\mathrm{d}\Omega. \tag{XXI.82}$$

Analogous expressions are found for the transitions to the state β. The total transition probability per unit time Γ_λ is:

$$\Gamma_\lambda = \Gamma_{\lambda\to\alpha} + \Gamma_{\lambda\to\beta}. \tag{XXI.83}$$

Γ_λ is the inverse of the *lifetime* of the state $|\lambda\rangle$, as can be seen by the following semi-classical argument. Consider a statistical ensemble of particles and denote by $N_\lambda(t)$ the number of particles in the state $|\lambda\rangle$ at time t. The number of particles that effect a radiative transition in the time interval $(t, t+\mathrm{d}t)$ is $\Gamma_\lambda N_\lambda(t)\,\mathrm{d}t$. Consequently,

$$N_\lambda(t+\mathrm{d}t) = (1-\Gamma_\lambda\,\mathrm{d}t)N_\lambda(t)$$

or:

$$\frac{\mathrm{d}N_\lambda}{\mathrm{d}t} = -\Gamma_\lambda N_\lambda.$$

Consequently

$$N_\lambda(t) = N_\lambda(0)\,\mathrm{e}^{-\Gamma_\lambda t}. \tag{XXI.84}$$

This is the well-known *law of exponential decay* with a lifetime equal to $1/\Gamma_\lambda$. In the example treated here there are essentially two possible decay modes, $\lambda \to \alpha$ and $\lambda \to \beta$ and the *branching ratio* of the first to the second is $\Gamma_{\lambda\to\alpha}/\Gamma_{\lambda\to\beta}$.

Clearly, the above semi-classical argument cannot be regarded as

a proof of the exponential law, since the perturbation treatment for transition probabilities is valid only for $\Gamma_\lambda t \ll 1$ (cf. discussion at the end of § XVII.4), that is, for t small enough so that the state vector differs little from its initial value. For longer times, a more refined treatment is needed which will be given in § 13.

Let us return to the numerical values and notations of § 11 in order to determine the order of magnitude of the Γ. For not too large excitation energies we have $k \approx D$, and since the extension $\langle R \rangle$ of the wave functions of the particle is about $(2MD)^{-\frac{1}{2}}$ we have:

$$k\langle R \rangle \approx (D/2M)^{\frac{1}{2}} \simeq 5 \times 10^{-3} \ll 1.$$

$\langle \lambda | e^{i\mathbf{k}\cdot\mathbf{R}} | \alpha \rangle$ can therefore be evaluated by replacing the exponential by the first term of its Taylor expansion giving a non-null contribution (long-wavelength approximation): the order of this term is essentially fixed by the angular momentum *selection rules*. In the most favourable case (dipole transitions), one finds

$$|\langle \lambda | e^{i\mathbf{k}\cdot\mathbf{R}} | \alpha \rangle| \simeq |\langle \lambda | \mathbf{k}\cdot\mathbf{R} | \alpha \rangle| \approx k\langle R \rangle$$

whence:

$$\Gamma_{\lambda \to \alpha} \approx \frac{g^2 k}{2\pi} (k\langle R \rangle)^2 \approx 10^{-6}\, D.$$

If we take the same value for Γ_λ, we obtain a lifetime $1/\Gamma_\lambda$ very much larger (by a factor 10^6) than the characteristic period $1/D$ of the motion of the particle in its potential. It is therefore reasonable to treat, as we have done, the interaction H' responsible for the emission of the corpuscle as a small perturbation.

A formally very simple and very general expression for Γ_λ can be obtained by starting directly from formula (XXI.80) and using the definition of level densities, the conservation of energy and the closure relation. We obtain successively [1])

$$\begin{aligned} \Gamma_\lambda &= 2\pi \left(\frac{L}{2\pi}\right)^3 \sum_\nu \int \mathrm{d}\mathbf{k}\, |\langle \lambda | H' | \nu \mathbf{k} \rangle|^2\, \delta(E_\lambda - E_\nu - \omega) \\ &= 2\pi \left(\frac{L}{2\pi}\right)^3 \sum_\nu \int \mathrm{d}\mathbf{k}\, \langle \lambda | H' \delta(E_\lambda - H_0) | \nu \mathbf{k} \rangle \langle \nu \mathbf{k} | H' | \lambda \rangle \\ &= 2\pi \langle \lambda | H' \delta(E_\lambda - H_0)\, H' | \lambda \rangle. \end{aligned} \tag{XXI.85}$$

This expression will prove useful in what follows.

[1]) Expression (XXI.85) appears in many problems; it is the general expression for the transition probability per unit time to the first order in the perturbation calculation. It can be deduced directly from the theory of § XVII.4 starting from formula (§ XVII.40) and using the asymptotic form (XVII.43) of the function f which appears in this formula.

13. Quantum Theory of Decaying States. Line Width

In this paragraph we shall give the exact quantum treatment of the evolution of a decaying state. It will be shown that the results given by the simplified arguments of § 12 are essentially correct, and notably that the falling-off is given to an excellent approximation by an exponential law. We shall also be led to the important concepts of level shift and line width.

Thus, we suppose that at the initial time $t=0$ the system is in a bound state of H_0 corresponding to the eigenvalue E_λ, and we consider its evolution in the course of time.

Up to the present, we have supposed that E_λ was a non-degenerate eigenvalue of H_0. This simplifying hypothesis will be invoked when we are discussing the results, but the treatment that follows is also applicable to the degenerate case. We denote by $\mathscr{E}_\lambda$ the subspace spanned by the bound states of unperturbed energy E_λ, and by P_λ and Q_λ the projectors onto $\mathscr{E}_\lambda$ and onto its complementary space respectively:

$$H_0P_\lambda = P_\lambda H_0 = E_\lambda P_\lambda, \qquad P_\lambda+Q_\lambda = 1.$$

Let $U(t)$ be the evolution operator for the system:

$$U(t) \equiv \mathrm{e}^{-\mathrm{i}Ht}. \tag{XXI.86}$$

Our problem amounts to calculating $P_\lambda U(t) P_\lambda$ for $t>0$.

Let us introduce the resolvent:

$$G(z) \equiv \frac{1}{z-H}. \tag{XXI.87}$$

$G(z)$ is a function of the complex variable z whose singularities are the eigenvalues of H (the discrete values are simple poles, the continuous spectrum is a cut). One has [1])

$$U(t) = -\frac{1}{2\pi\mathrm{i}} \int_{-\infty}^{+\infty} \mathrm{e}^{-\mathrm{i}xt}\,[G(x^+) - G(x^-)]\,\mathrm{d}x \qquad (x^\pm \equiv x \pm \mathrm{i}\varepsilon). \tag{XXI.88}$$

[1]) Relation (XXI.88) is directly obtained by using the fact that for ε infinitesimal and positive [cf. eq. (A.15e)]

$$G(x^\pm) \equiv \frac{1}{x-H\pm\mathrm{i}\varepsilon} = \mathrm{PP}\,\frac{1}{x-H} \mp \mathrm{i}\pi\delta(x-H). \tag{XXI.89}$$

Note that for $t>0$ the contribution from $G(x^-)$ to the integral on the right-hand side of (XXI.88) vanishes.

It follows that the operators $\mathscr{U}$ and $\mathscr{G}$, defined in the space $\mathscr{E}_\lambda$ by the relations:

$$\mathscr{U}(t) = P_\lambda U(t) P_\lambda, \qquad \mathscr{G}(z) = P_\lambda G(z) P_\lambda, \tag{XXI.90}$$

are related by the formula:

$$\mathscr{U}(t) = -\frac{1}{2\pi i}\int_{-\infty}^{+\infty} e^{-izt}\,[\mathscr{G}(x^+) - \mathscr{G}(x^-)]\,dx \tag{XXI.91}$$
$$(x^\pm \equiv x \pm i\varepsilon).$$

In order to obtain a workable expression for $\mathscr{G}(z)$, we put:

$$H = H_1 + H''$$
$$H_1 \equiv P_\lambda H P_\lambda + Q_\lambda H Q_\lambda = H_0 + Q_\lambda H' Q_\lambda$$
$$H'' \equiv P_\lambda H Q_\lambda + Q_\lambda H P_\lambda = P_\lambda H' Q_\lambda + Q_\lambda H' P_\lambda.$$

Note that:

$$H_1 P_\lambda = P_\lambda H_1 = E_\lambda P_\lambda \qquad [Q_\lambda, H_1] = 0$$
$$P_\lambda H'' = H'' Q_\lambda \qquad Q_\lambda H'' = H'' P_\lambda.$$

From these relations, and from the operator identity:

$$(A-B)^{-1} \equiv A^{-1} + A^{-1} B (A-B)^{-1},$$

we obtain:

$$\frac{1}{z-H} = \frac{1}{z-H_1} + \frac{1}{z-H_1} H'' \frac{1}{z-H} \tag{XXI.92}$$

$$= \frac{1}{z-H_1} + \frac{1}{z-H_1} H'' \frac{1}{z-H_1} + \frac{1}{z-H_1} H'' \frac{1}{z-H_1} H'' \frac{1}{z-H} \tag{XXI.93}$$

$$P_\lambda \frac{1}{z-H} P_\lambda = \frac{P_\lambda}{z-E_\lambda} + \frac{1}{z-E_\lambda}\left(P_\lambda H'' \frac{1}{z-H_1} H'' P_\lambda\right) P_\lambda \frac{1}{z-H} P_\lambda. \tag{XXI.94}$$

Relation (XXI.94) is a relation between operators of the space $\mathscr{E}_\lambda$, namely:

$$\mathscr{G}(z) = \frac{1}{z-E_\lambda}\,[1 + \mathscr{W}(z)\,\mathscr{G}(z)], \tag{XXI.95}$$

where:

$$\mathscr{W}(z) \equiv P_\lambda H'' \frac{1}{z-H_1} H'' P_\lambda = P_\lambda H' Q_\lambda \frac{1}{z-Q_\lambda H Q_\lambda} Q_\lambda H' P_\lambda. \tag{XXI.96}$$

More generally, $U(t)$ and $G(z)$ are related by the transformation formulas

$$G(z) = -i\int_0^{\eta\infty} e^{izt}\,U(t)\,dt \qquad U(t) = \frac{1}{2\pi i}\int_C e^{-izt}\,G(z)\,dz$$

[η sign of $\operatorname{Im} z$; C = contour in the complex plane going from $\infty + i\varepsilon$ to $\infty - i\varepsilon$ with all the singularities of $G(z)$ i.e., all the eigenvalues of H, on its left], which are easily demonstrated in the representation where H is diagonal.

From equation (XXI.95) we obtain:

$$\mathscr{G}(z) = \frac{1}{z - E_\lambda - \mathscr{W}(z)}. \tag{XXI.97}$$

Equations (XXI.96), (XXI.97) and (XXI.91) are exact relations which can serve as a starting point for a calculation of $\mathscr{U}(t)$.

Equations (XXI.96) and (XXI.97) define $\mathscr{W}(z)$ and $\mathscr{G}(z)$ respectively as functions of the complex variable z in the plane cut by the eigenvalues of the continuous spectrum of H. We are especially interested in the behavior of these functions in the neighborhood of the cut. Using relation (XXI.89), the Hermitean and antihermitean parts of $\mathscr{W}(x^\pm)$ can be separated. One obtains:

$$\mathscr{W}(x^\pm) = \Delta(x) \mp \tfrac{1}{2}\mathrm{i}\Gamma(x) \tag{XXI.98}$$

$$\Delta(x) \equiv P_\lambda H' \,\mathrm{PP}\, \frac{1}{x - Q_\lambda H Q_\lambda} H' P_\lambda \tag{XXI.99}$$

$$\Gamma(x) \equiv 2\pi P_\lambda H' \delta(x - Q_\lambda H Q_\lambda) H' P_\lambda. \tag{XXI.100}$$

Note that $\Gamma(x)$ is a positive definite Hermitean operator.

We shall henceforth suppose that E_λ is a non-degenerate eigenvalue [1]) of H_0.

---

[1]) In most practical cases, the degeneracy of E_λ is due to an invariance of H_0 which remains when the coupling H' is added. For example in the radioactive decay of a nucleus of spin J, the degeneracy of order $2J + 1$ is due to invariance under rotation, which remains when the coupling responsible for the decay is added to the unperturbed Hamiltonian. Since $\mathscr{E}_\lambda$ is then an irreducible subspace, and the operators $\mathscr{W}(z)$, $\mathscr{G}(z)$ and $\mathscr{U}(t)$ are invariant, the latter are multiples of the unit operator, and they can be calculated in the same way as in the non-degenerate case. Each state of the space $\mathscr{E}_\lambda$ follows the same law of radioactive decay with the same lifetime and the same level shift.

If, on the other hand, these symmetry conditions do not hold, there can be several lifetimes and several shifts. This is what occurs for the first excited level of the hydrogen atom (cf. discussion of the Lamb shift at the end of § XX.27). Another example is that of the K^0 boson. The treatment of the general case is essentially the same as that of the non-degenerate case. With the same approximations, we obtain for $\mathscr{U}(t)$ the expression:

$$\mathscr{U}(t) \simeq \exp\left[-\mathrm{i}(E_\lambda + \mathscr{W}(E_\lambda^+))\, t\right]$$

which is the generalization of (XXI.109). The real and imaginary parts of the eigenvalues of the (non-Hermitean) operator $\mathscr{W}(E_\lambda^+)$ respectively give the different level shifts and corresponding lifetimes.

By hypothesis, the system is in the corresponding eigenstate $|\lambda\rangle$ at time $t=0$. We wish to determine its state vector at any later time t

$$|\Psi(t)\rangle = U(t)\,|\lambda\rangle \tag{XXI.101}$$

and in particular the component of this vector along $|\lambda\rangle$.

The operators $\mathscr{U}(t)$, $\mathscr{G}(z)$, $\mathscr{W}(z)$, $\Delta(x)$ and $\Gamma(x)$ can be replaced in the preceding work by their average value over $|\lambda\rangle$, and treated as simple functions. Equations (XXI.90), (XXI.99) and (XXI.100) are then respectively replaced by:

$$\mathscr{U}(t) \equiv \langle\lambda|U(t)|\lambda\rangle \qquad \mathscr{G}(z) \equiv \langle\lambda|G(z)|\lambda\rangle \tag{XXI.102}$$

$$\Delta(x) \equiv \langle\lambda|\,H'\,\mathrm{PP}\left(\frac{1}{x-Q_\lambda H\,Q_\lambda}\right)H'\,|\lambda\rangle \tag{XXI.103}$$

$$\Gamma(x) \equiv 2\pi\langle\lambda|H'\delta(x-Q_\lambda H Q_\lambda)H'|\lambda\rangle, \tag{XXI.104}$$

and equations (XXI.97) and (XXI.91) give respectively:

$$\mathscr{G}(x^\pm) = [x-E_\lambda-\Delta(x) \pm \tfrac{1}{2}\mathrm{i}\Gamma(x)]^{-1} \tag{XXI.105}$$

$$\mathscr{U}(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}xt}\,F(x)\,\mathrm{d}x, \tag{XXI.106}$$

where:

$$F(x) \equiv \mathrm{i}(\mathscr{G}(x^+)-\mathscr{G}(x^-)) = \frac{\Gamma(x)}{(x-E_\lambda-\Delta(x))^2+\frac{1}{4}\Gamma(x)^2}. \tag{XXI.107}$$

Up to now, no approximation has been made. To evaluate $F(x)$ and $\mathscr{U}(t)$, we shall suppose the coupling weak and retain only the lowest order terms in H' in expressions (XXI.103) and (XXI.104) giving respectively $\Delta(x)$ and $\Gamma(x)$. For this we need only to replace $Q_\lambda H Q_\lambda$ by H_0 in these expressions. It is then easy to calculate $\Delta(x)$ and $\Gamma(x)$. The form of $\Gamma(x)$ is given in Fig. XXI.4. For Δ and Γ sufficiently small, $F(x)$ has a sharp maximum at about the point $x=E_\lambda$, and the main contribution to integral (XXI.106) comes from this region. This integral can therefore be evaluated by substituting for the slowly-varying functions $\Delta(x)$ and $\Gamma(x)$ their values at the point $x=E_\lambda$, namely:

$$\Delta(E_\lambda) \equiv \delta E_\lambda, \qquad \Gamma(E_\lambda) \equiv \Gamma_\lambda, \tag{XXI.108}$$

which amounts to replacing (cf. Fig. XXI.5) the function $F(x)$ by:

$$F_1(x) = \frac{\Gamma_\lambda}{(x-E_\lambda-\delta E_\lambda)^2+\frac{1}{4}\Gamma_\lambda^2}.$$

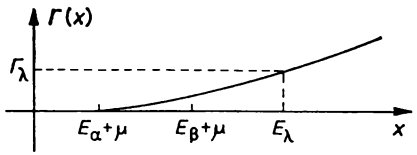


Fig. XXI.4. Form of $\Gamma(x)$. Note that $\Gamma(x) = 0$ for $x < E_\alpha + \mu$.

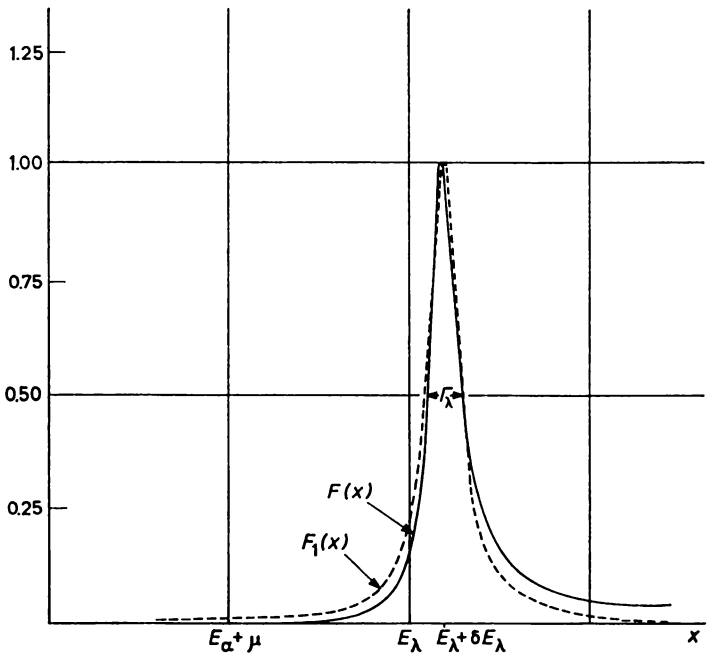


Fig. XXI.5. Typical form of $F(x)$ and $F_1(x)$.

Here the function Δ and Γ have been given an x dependence typical of the dipole transitions:

$$\Delta(x) \simeq Cst = \delta E_\lambda \qquad \frac{\Gamma(x)}{\Gamma_\lambda} \simeq \frac{(x - E_\lambda + \xi)^3}{\xi^3} \qquad (\xi = E_\lambda - E_\alpha - \mu).$$

We have taken $\delta E_\lambda/\xi = 0.2 \qquad \Gamma_\lambda/\xi = 0.1$.

The error in $U(t)$ is smaller than the absolute value of the quantity

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} |F(x) - F_1(x)| \, dx.$$

If Γ_λ and δE_λ are sufficiently small compared with $E_\lambda - E_\alpha - \mu$ and $E_\lambda - E_\beta - \mu$, it remains negligible in relative value except for very long times ($t \gg 1/\Gamma_\lambda$).

The integration is then easily carried out and gives:

$$\mathscr{U}(t) \simeq \mathrm{e}^{-\mathrm{i}(E_\lambda+\delta E_\lambda)t-\frac{1}{2}\Gamma_\lambda t}. \tag{XXI.109}$$

Therefore:

$$|\mathscr{U}(t)|^2 = \mathrm{e}^{-\Gamma_\lambda t}. \tag{XXI.110}$$

This is the expected exponential decay law.

The value of the *lifetime*, $1/\Gamma_\lambda$, is that which was calculated in § 12 [cf. eqs. (XXI.84–85), (XXI.104) and (XXI.108)].

According to formula (XXI.109), the modifications of the law of motion of the level λ due to the coupling H' consists essentially in adding to it the complex energy

$$\delta E_\lambda - \tfrac{1}{2}\mathrm{i}\Gamma_\lambda.$$

The real part, δE_λ, is the *level shift* proper; it has a form analogous to that of the shift of a stable level [cf. eqs. (XXI.71), (XXI.103) and (XXI.108)]. The imaginary part, characteristic of a decaying state, is to within a sign the half-width of the level; it is the part responsible for the existence of the exponential falling-off.

Let us now calculate the other components of $|\Psi(t)\rangle$ in the weak-coupling approximation. In this approximation the only non-vanishing components are along the vectors containing one and only one corpuscle:

$$|\Psi(t)\rangle = \mathscr{U}(t)|\lambda\rangle + \sum_{\nu\mathbf{k}} h_{\nu\mathbf{k}}(t)|\nu\mathbf{k}\rangle \tag{XXI.111}$$

$$h_{\nu\mathbf{k}}(t) = \langle\nu\mathbf{k}|U(t)|\lambda\rangle. \tag{XXI.112}$$

From equation (XXI.92),

$$\langle\nu\mathbf{k}|G(z)|\lambda\rangle = \langle\nu\mathbf{k}|(z-H_1)^{-1}H''|\lambda\rangle\mathscr{G}(z).$$

This equation is exact. The weak-coupling approximation consists in replacing H_1 by H_0 in the denominator on the right-hand side, which gives:

$$\langle\nu\mathbf{k}|G(z)|\lambda\rangle = \langle\nu\mathbf{k}|H'|\lambda\rangle(z-\omega-E_\nu)\mathscr{G}(z).$$

From this approximate value for the matrix element of $G(z)$ one obtains an approximate value for the corresponding element of $U(t)$ by applying formula (XXI.88), which gives an integral which may be evaluated by the residue method. For sufficiently large values of t,

the only significant contribution is the one from the pole on the real axis, which gives

$$h_{\nu\mathbf{k}}(t) \simeq \langle \nu\mathbf{k}|H'|\lambda\rangle \left.\frac{e^{-ixt}}{x - E_\lambda - \Delta(x) + \frac{1}{2}i\Gamma(x)}\right|_{x=\omega+E_\nu}.$$

$$(\Gamma t \gg 1)$$

The square of the modulus of this expression gives a probability of finding the system in the state $(\nu\mathbf{k})$ for the large values of t (large with respect to the decay lifetime). It can be put in the form:

$$|h_{\nu\mathbf{k}}(\infty)|^2 = |\langle \nu\mathbf{k}|H'|\lambda\rangle|^2 \frac{\Gamma(\omega + E_\nu)}{F(\omega + E_\nu)}, \qquad \text{(XXI.113)}$$

where $F(x)$ is the function defined by equation (XXI.107). According to the above-mentioned properties of $F(x)$ (cf. Fig. (XXI.5)), $|h_{\nu\mathbf{k}}(\infty)|^2$ takes appreciable values only for $\omega + E_\nu \simeq E_\lambda$, that is, for transitions which roughly conserve the unperturbed energy. For these – in the example treated here these are the ones for which $\nu = \alpha$ and $\nu = \beta$ – we can replace $F(x)$ by $F_1(x)$. Thus, for the transitions $\lambda \to \alpha\mathbf{k}$ we obtain the formula:

$$|h_{\nu\mathbf{k}}(\infty)|^2 = \frac{|\langle \alpha\mathbf{k}|H'|\lambda\rangle|^2}{(\omega + E_\alpha - E_\lambda - \delta E_\lambda)^2 + \frac{1}{4}\Gamma_\lambda^2}. \qquad \text{(XXI.114)}$$

According to this formula, the energy distribution of the corpuscles emitted in the transition is given by a *"Lorentz law" of width* Γ_λ centered about the point $E_\lambda + \delta E_\lambda - E_\alpha$, that is, centered about the Bohr frequency of the transition, corrected by the *shift* δE_λ.

The *line width* of the corpuscles emitted in the radiative transition is therefore equal, in our units where $\hbar = 1$, to the inverse of the lifetime in accordance with the time-energy uncertainty relation.

Several remarks should be made regarding the significance of these results.

In the first place, to compare a distribution law such as (XXI.114) with experimental results, one should indicate how the initial state $|\lambda\rangle$ can be experimentally realized. We shall return to this point in § 15.

Apart from this, the preceding treatment has rather severe limitations. If we take the numerical results of § 11, we see that δE_λ is of the same order as δE_α, and therefore much larger than the level distance. However, if one denotes by $\delta E_\lambda^{(0)}$ the effect due to mass renormalization, it can be shown that the difference

$$\delta E_\lambda' \equiv \delta E_\lambda - \delta E_\lambda^{(0)}$$

is a small correction compared to the level distance. The difficulty encountered here is exactly the same as the one discussed at some length in § 11; that discussion can be repeated here point by point.

Finally – and this is obviously related to the weak-coupling approximation in the calculation of the amplitudes $h_{\nu k}$ – we have not taken into account possible shifts in the levels towards which the transition is effected, nor *a fortiori* of the fact that some of these levels are themselves rendered unstable by the presence of the coupling H'.

For transitions towards stable levels or levels with a long lifetime compared with the lifetime of the initial state, formula (XXI.114) can be corrected so as to take the two last-mentioned effects into account. Denote by $E_\alpha^{(0)}, \ldots, E_\lambda^{(0)}, \ldots$ the levels obtained by substituting the experimental mass for M in the Hamiltonian H_{part}. The corrected formula is written:

$$|h_{\nu k}|^2_{\text{cor}} = \frac{|\langle \alpha \mathbf{k}|H'|\lambda\rangle|^2}{(\omega + E_\alpha^{(0)} + \delta E_\alpha' - E_\lambda^{(0)} - \delta E_\lambda')^2 + \frac{1}{4}\Gamma_\lambda^2}. \quad \text{(XXI.114')}$$

It is a formula of this type that is usually compared with experimental results. The agreement is excellent.

14. Elastic Scattering. Dispersion Formula

To conclude this section let us consider some simple collision problems.

As a first problem, we treat the elastic scattering of a field corpuscle by a particle in its ground state. Denote by $\mathbf{k}_i$ the wave vector of the incident corpuscle and by $\mathbf{k}_f$ a vector of the same length pointing in the direction in which the corpuscle is observed after the collision. The transition in question is therefore

$$i \to f \equiv (\alpha \mathbf{k}_i) \to (\alpha \mathbf{k}_f).$$

We put

$$k_i = k_f = k, \qquad \omega_i = \omega_f = \omega,$$

and denote the initial energy of the system by E:

$$E = \omega + E_\alpha. \quad \text{(XXI.115)}$$

The cross section is given by the general formula:

$$\frac{\mathrm{d}\sigma_{i\to f}}{\mathrm{d}\Omega} = \frac{2\pi}{\text{incident flux}} |T_{i\to f}|^2 \varrho_f(E). \quad \text{(XXI.116)}$$

$T_{i\to f}$ is the transition amplitude:

$$T_{i\to f} \equiv \langle f|T|i\rangle \equiv \langle \alpha \boldsymbol{k}_f|T|\alpha \boldsymbol{k}_i\rangle \tag{XXI.117}$$

$$T = H' + H' \frac{1}{E - H + \mathrm{i}\varepsilon} H'. \tag{XXI.118}$$

The incident flux is equal to the product of the incident velocity, k/ω, by the density of corpuscles in the state $|\alpha \boldsymbol{k}_i\rangle$, which gives, with the normalization adopted here: $L^{-3}k/\omega$. $\varrho_f(E)$ is the density of final states at energy E [cf. eq. (XXI.49)]. One therefore has:

$$\frac{2\pi \varrho_f(E)}{\text{incident flux}} = \frac{L^6 \omega^2}{4\pi^2}.$$

Whence:

$$\frac{\mathrm{d}\sigma_{i\to f}}{\mathrm{d}\Omega} = \frac{L^6 \omega^2}{4\pi^2} |T_{i\to f}|^2. \tag{XXI.119}$$

In all of this, we have assumed that the basic formulae of collision theory are entirely valid here: equations (XXI.116) and (XXI.118) which constitute our point of departure are essentially equivalent to equations (XIX.115) and (XIX.144) respectively. The few differences are due only to notation, to the units chosen and to the normalization of the wave functions involved in the definition of the transition amplitude.

In the present case, however, these formulae are not absolutely correct. They were established under the hypothesis that the evolution of the system *before* and *after* the collision was almost entirely governed by the Hamiltonian H_0, the interaction H' being completely negligible. Such a hypothesis, entirely justified in the collision problems of Chapters X and XIX, ceases to hold here.

Let us examine, for example, the state of the system before the collision. The particle is then in its ground state and the corpuscle is represented by a wave packet moving towards it. The wave packet has not yet reached the small region about the origin in which the particle moves; the interaction corpuscle-particle is therefore negligible and the corpuscle evolves as if it were free. In this respect, the present situation does not differ from that of ordinary collision theory. But the interaction H' nevertheless cannot be neglected, since the particle, even when separated from the incident corpuscle, is always coupled with the field; to be exact, its initial state is not the eigenstate $|\alpha\rangle$ of the Hamiltonian H_0 corresponding to the energy E_α, but the eigenstate $|\bar{\alpha}\rangle$ of the Hamiltonian H which coincides with $|\alpha\rangle$ in the limit $g \to 0$, and whose eigenenergy, $E_\alpha + \delta E_\alpha$, was calculated in § 11.

To be correct, the theory must take into account the difference between the "physical state" $|\tilde{\alpha}\rangle$ and the unperturbed state $|\alpha\rangle$. If the coupling is sufficiently weak, it leads essentially to an effect of renormalization of the mass, and the mechanism of the collision proper is not affected [1]).

Under these conditions, the usual formulae of collision theory are valid if we replace the mass M by the experimental mass $M+\delta M$ in the definition of the unperturbed Hamiltonian.

We shall now calculate the cross section by treating H' as a perturbation. We therefore replace T by its Born expansion (cf. eq. (XIX.143)). In the first order, the amplitude obviously vanishes:

$$\langle \alpha \boldsymbol{k}_f | H' | \alpha \boldsymbol{k}_i \rangle = 0.$$

We must therefore proceed to the second order. We denote the amplitude to the second order in H' by $T^{(B)}_{i\to f}$:

$$T^{(B)}_{i\to f} \equiv \langle \alpha \boldsymbol{k}_f | H' \frac{1}{E - H_0 + i\varepsilon} H' | \alpha \boldsymbol{k}_i \rangle \qquad \text{(XXI.120)}$$

$$= \sum_n \langle \alpha \boldsymbol{k}_f | H' | n \rangle \frac{1}{E - E_n + i\varepsilon} \langle n | H' | \alpha \boldsymbol{k}_i \rangle. \qquad \text{(XXI.120')}$$

In the second line, the summation is extended over all of the basis vectors of the Hamiltonian H_0. Owing to the very special form of H' most of the terms in the sum vanish. The virtual states whose contribution is different to zero can be put into two categories:

(*i*) the zero-corpuscle states $|\nu\rangle$;

(*ii*) the states $|\nu \boldsymbol{k}_i \boldsymbol{k}_f\rangle$ containing two corpuscles of momentum $\boldsymbol{k}_i$ and $\boldsymbol{k}_f$ respectively.

The first correspond to the transitions:

$$(\alpha \boldsymbol{k}_i) \to (\nu) \to (\alpha \boldsymbol{k}_f)$$

in which the particle absorbs the incident corpuscle before emitting the corpuscle in the final state; the second correspond to the transitions:

$$(\alpha \boldsymbol{k}_i) \to (\nu \boldsymbol{k}_i \boldsymbol{k}_f) \to (\alpha \boldsymbol{k}_f)$$

---

[1]) In a higher approximation, one must take into account the modifications of the interaction of the incident corpuscle and the particle by the presence of virtual corpuscles surrounding the latter. This modification is essentially represented by a modification of the coupling constant ("charge renormalization").

in which the final corpuscle is emitted before the initial corpuscle is absorbed.

Let us introduce the Bohr frequencies:

$$\omega_{\nu\alpha} = E_\nu - E_\alpha \qquad (>0). \tag{XXI.121}$$

One finds

$$T^{(B)}_{i\to f} = \sum_\nu \left[\frac{\langle \alpha \mathbf{k}_f | H' | \nu \rangle \langle \nu | H' | \alpha \mathbf{k}_i \rangle}{\omega - \omega_{\nu\alpha}} - \frac{\langle \alpha \mathbf{k}_f | H' | \nu \mathbf{k}_i \mathbf{k}_f \rangle \langle \nu \mathbf{k}_i \mathbf{k}_f | H' | \alpha \mathbf{k}_i \rangle}{\omega + \omega_{\nu\alpha}} \right]. \tag{XXI.122}$$

This gives, applying equation (XXI.69) and (XXI.70) with $C(k)=1$ (which is certainly legitimate if $k \ll M$):

$$T^{(B)}_{i\to f} = \frac{L^{-3} g^2}{2\omega} \sum_\nu \left[\frac{X_\nu}{\omega - \omega_{\nu\alpha}} - \frac{X_\nu'}{\omega + \omega_{\nu\alpha}} \right] \tag{XXI.123}$$

where:

$$X_\nu \equiv \langle \alpha | \, e^{-i\mathbf{k}_f \cdot \mathbf{R}} | \nu \rangle \langle \nu | \, e^{i\mathbf{k}_i \cdot \mathbf{R}} | \alpha \rangle \tag{XXI.124}$$

$$X_\nu' \equiv \langle \alpha | \, e^{i\mathbf{k}_i \cdot \mathbf{R}} | \nu \rangle \langle \nu | \, e^{-i\mathbf{k}_f \cdot \mathbf{R}} | \alpha \rangle. \tag{XXI.124'}$$

Substituting this expression into the right-hand side of (XXI.119), one obtains the cross section in the second-order Born approximation:

$$\frac{d\sigma^{(B)}_{i\to f}}{d\Omega} = \frac{g^4}{16\pi^2} \left| \sum_\nu \left(\frac{X_\nu}{\omega - \omega_{\nu\alpha}} - \frac{X_\nu'}{\omega + \omega_{\nu\alpha}} \right) \right|^2. \tag{XXI.125}$$

Each term in this series represents the contribution of one of the transitions described above. All things being equal, this contribution is the larger the nearer the incidence energy to the energy of the virtual state involved in that transition; when these two energies are equal, the corresponding term is infinite. There is one such critical value for ω for each transition of the first category in which the intermediate state can decay into the ground state with emission of a field-corpuscle.

When passing by one of these critical values, E_λ for example ($\omega_{\lambda\alpha} > \mu$), the denominator $\omega - \omega_{\lambda\alpha}$ vanishes and changes sign, the term $X_\lambda/(\omega - \omega_{\lambda\alpha})$ increases indefinitely and the expression for the cross section diverges. However small may be the coupling constant,

the Born approximation obviously loses all validity in this domain [1]); in actual fact the cross section remains finite, but passes by a very sharp maximum. It is, as we shall see, a resonance phenomenon entirely analogous to the ones studied in Chapter X (§ 14–16).

15. Resonance Scattering. Formation of a Metastable State

To calculate the elastic scattering cross section in the neighborhood of one of these critical energies, E_λ say, we return to the exact expression for the transition amplitude. According to definitions (XXI.117) and (XXI.118),

$$\langle f|T|i\rangle = \sum_{nn'} \langle f|H'|n\rangle \langle n| \frac{1}{E - H + i\varepsilon} |n'\rangle \langle n'|H'|i\rangle. \quad \text{(XXI.126)}$$

The Born approximation consists in replacing in each term of this expansion the exact Green's function $(E-H+i\varepsilon)^{-1}$ by the Green's function $(E-H_0+i\varepsilon)^{-1}$ of the unperturbed Hamiltonian [cf. eq. (XXI.120) and (XXI.120′)]. If the coupling constant is sufficiently small, this approximation is justified for all of the terms except the term $n=\lambda$, $n'=\lambda$, for which the factor $\langle\lambda|(E-H_0+i\varepsilon)^{-1}|\lambda\rangle$ increases indefinitely when the energy passes by the value E_λ. Denote this term by $A^{(\text{res})}$:

$$A^{(\text{res})} \equiv \langle f|H'|\lambda\rangle \langle\lambda| \frac{1}{E - H + i\varepsilon} |\lambda\rangle \langle\lambda|H'|i\rangle \quad \text{(XXI.127)}$$

and the difference by $A^{(\text{pot})}$:

$$\langle f|T|i\rangle \equiv A^{(\text{res})} + A^{(\text{pot})}. \quad \text{(XXI.128)}$$

We shall use the Born approximation to calculate $A^{(\text{pot})}$; the ex-

[1]) In the long-wavelength approximation ($k\langle R\rangle \ll 1$), one has $|X_\lambda| \simeq (k\langle R\rangle)^2$; if we put $\Delta\omega = |E - E_\lambda|$ and retain only the contribution from the dominant term, expression (XXI.125) gives:

$$\frac{d\sigma}{d\Omega} \simeq \frac{g^4}{16\pi^2} \left(\frac{k^2\langle R\rangle^2}{\Delta\omega}\right)^2.$$

According to a known criterion (§ XIX.7), the Born approximation ceases to be applicable when the total cross section is of the order of the geometrical cross section $4\pi\langle R\rangle^2$. This occurs here for:

$$\Delta\omega \lesssim \frac{g^2 k}{4\pi} (k\langle R\rangle).$$

With the numerical values given in § 11 this is an extremely narrow domain, of the order of $10^{-4} D$ (nevertheless much larger than the width Γ_λ).

pression obtained differs from (XXI.123) only by the absence of the term $X_\lambda/(\omega-\omega_{\lambda\alpha})$ in the summation over the states ν. It remains to calculate $A^{(\mathrm{res})}$.

By replacing the matrix elements of H' in (XXI.127) by their explicit form [cf. eq. (XXI.69) and (XXI.124)], we obtain:

$$A^{(\mathrm{res})} = L^{-3}g^2 \frac{X_\lambda}{2\omega} \mathcal{G}(E^+), \tag{XXI.129}$$

where $\mathcal{G}(E^+)$ is the average value of the Green's function introduced in § 13:

$$\mathcal{G}(E^+) \equiv \langle\lambda| \frac{1}{E-H+\mathrm{i}\varepsilon} |\lambda\rangle. \tag{XXI.130}$$

This function was calculated in § 13. The result is given by equation (XXI.105). Substituting the expression thus obtained in equation (XXI.129), one obtains:

$$A^{(\mathrm{res})} = \frac{L^{-3}g^2}{2\omega} \frac{X_\lambda}{E-E_\lambda-\Delta(E)+\frac{1}{2}\mathrm{i}\Gamma(E)}. \tag{XXI.131}$$

This expression for $A^{(\mathrm{res})}$ is exact. It differs from the one given by the Born approximation only by the presence of the complex term $\Delta(E)-\frac{1}{2}\mathrm{i}\Gamma(E)$ in the denominator. This additional term removes the singularity. Its properties were examined in § 13. By replacing it by its value for $E=E_\lambda$ [cf. eqs. (XXI.103), (XXI.104) and (XXI.108)], which is certainly legitimate in the weak-coupling approximation since its energy variation is negligible in the energy band of interest, one obtains:

$$A^{(\mathrm{res})} = \frac{L^{-3}g^2}{2\omega} \frac{X_\lambda}{E-E_\lambda-\delta E_\lambda+\frac{1}{2}\mathrm{i}\Gamma_\lambda} \tag{XXI.132}$$

This expression is characteristic of a scattering amplitude of width Γ_λ centered about the point $E_\lambda+\delta E_\lambda$.

The discussion of Chapter X on scattering resonances and their *relation to the existence of decaying states* can be repeated here. By further specifying the model, it is easy to obtain expressions practically identical to those given in § X.15 (providing $A^{(\mathrm{pot})}$ is neglected) for the scattering amplitude, the differential and total cross sections, and the "delay" in the transmission of the scattered wave (cf. Problem XXI.4). The above resonance is related to the decaying state $|\lambda\rangle$

whose properties have already been discussed in § 12 and 13.

To carry out a precise measurement of the cross section, and in particular to exhibit its characteristic variation with the energy, it is necessary that the energy uncertainty $\Delta\omega$ of the incident wave packet be sufficiently small that the scattering amplitude remains practically constant over the interval $\Delta\omega$; in the neighborhood of the resonance this requires:

$$\Delta\omega \ll \Gamma_\lambda. \tag{XXI.133}$$

The collision time, $1/\Delta\omega$, is then much larger than the lifetime $1/\Gamma_\lambda$, and the latter is unobservable.

If, instead of this, we have the complementary conditions:

$$\Delta\omega \gg \Gamma_\lambda, \tag{XXI.134}$$

it is possible to observe the time development of the phenomenon and to exhibit, after a transient period of the order of $1/\Delta\omega$, the exponentially decreasing law characteristic of the decaying state $|\lambda\rangle$. This is due to the rather special form of the scattering amplitude in the resonance region [eq. (XXI.132)], and can easily be established by repeating point by point the calculation of § X.16.

This result is a very general one, valid not only for scattering but for all collisions. Any collision resonance defined by its energy E and its width Γ is associated with a metastable state of the same energy and of lifetime $\hbar/\Gamma$, which can be observed under experimental conditions complementary to those in which one usually observes the resonance [1]).

16. Absorption of a Corpuscle (Photo-electric Effect). Radiative Capture

The problems treated thus far have involved only bound states of the particle. In this paragraph we shall treat two problems, absorp-

[1]) Here we have two complementary manifestations of the same property of the system, since they both follow from the same property of its resolvant $G(z)$. The latter is defined by the equation $G(z) = (z - H)^{-1}$ in the plane cut by the continuous spectrum of H: its only singularities in the cut plane are on the real axis, and are the eigenvalues of H. However its analytic continuation onto other Riemann sheets may have other singularities. Resonance and decaying state are due to the fact that $G(z)$ has a pole at point $z = E - \frac{1}{2}i\Gamma$ near the real axis on the second Riemann sheet of the lower half plane.

tion of a corpuscle and radiative capture, which deal with possible unbound states of the particle.

We shall suppose that $\lim_{R\to\infty} V(R) = 0$ and consequently that the spectrum of the particle subsystem contains a discrete set of bound states of negative energy $E_\alpha, E_\beta, \ldots, E_\lambda, \ldots$, and a continuum of unbound states of energy extending from 0 to $+\infty$. We denote by $\psi_{\mathbf{q}}^{(+)}$ and $\psi_{\mathbf{q}}^{(-)}$ the stationary waves associated with the wave vector $\mathbf{q}$ according to the definitions of § XIX.2. By definition, they are eigensolutions of H_{part} corresponding to the energy:

$$E = \frac{q^2}{2M}. \tag{XXI.135}$$

They may also be regarded as eigensolutions of H_0 representing stationary collision states of the total system when the coupling term H' is neglected. We shall denote the corresponding stationary collision states when the coupling is taken into account by $\Psi_{\mathbf{q}}^{(+)}$ and $\Psi_{\mathbf{q}}^{(-)}$.

Suppose now that the particle in its ground state is bombarded by a monoergic beam of corpuscles of velocity $\mathbf{k}/\omega$. If the energy ω of the incident corpuscles is large enough to "ionize the atom", that is, if

$$\omega > (-E_\alpha),$$

the corpuscle may simply be absorbed, transferring the whole of its energy to the particle; the latter will then be ejected with a kinetic energy E equal to the difference between the energy received ω and the binding energy $(-E_\alpha)$:

$$E = \omega + E_\alpha.$$

This process will be recognized as the photoelectric effect (§ I.4).

Let us calculate the cross section for the absorption of a corpuscle of energy $\mathbf{k}$ with ejection of the particle in a given direction. We denote by $\mathbf{q}$ a vector pointing in this direction whose length is related to the energy E by relation (XXI.135). Thus we are concerned with the transition:

$$i \to f \equiv (\alpha\mathbf{k}) \to (\mathbf{q}).$$

As in the elastic scattering case, we suppose that the formulas of ordinary collision theory can be applied [the same qualifications apply here as in § 14]. The cross section is therefore given by formula

(XXI.116) with the same expression for the incident flux, but with $\varrho_f(E) = Mq/(2\pi)^3$ and with the transition amplitude given by [cf. eq. (XIX.122)]:

$$T_{i\to f} \equiv \langle \boldsymbol{q}|T|\alpha\boldsymbol{k}\rangle = \langle \Psi_{\boldsymbol{q}}^{(-)}|H'|\alpha\boldsymbol{k}\rangle.$$

If one stops at the lowest order in H', $\Psi_{\boldsymbol{q}}^{(-)}$ may be replaced by $\psi_{\boldsymbol{q}}^{(-)}$ in this matrix element [cf. eq. (XIX.129–130)]; which gives

$$T_{i\to f} \simeq \langle \psi_{\boldsymbol{q}}^{(-)}|H'|\alpha\boldsymbol{k}\rangle = (L^{-3}g^2/2\omega)^{\frac{1}{2}}\langle \psi_{\boldsymbol{q}}^{(-)}|e^{i\boldsymbol{k}\cdot\boldsymbol{R}}|\alpha\rangle,$$

whence

$$\frac{d\sigma(\alpha\boldsymbol{k}\to\boldsymbol{q})}{d\Omega} \simeq \frac{g^2}{8\pi^2}\frac{Mq}{k}\,|\langle \psi_{\boldsymbol{q}}^{(-)}|\,e^{i\boldsymbol{k}\cdot\boldsymbol{R}}\,|\alpha\rangle|^2. \qquad \text{(XXI.136)}$$

In the same way we can calculate the cross section for the inverse reaction, in which the particle initially propagating with momentum $\boldsymbol{q}$ is captured in the state α by the potential $V(R)$, emitting a corpuscle of momentum $\boldsymbol{k}$:

$$\boldsymbol{q} \to \alpha\boldsymbol{k}.$$

To the first order in H', one finds:

$$\frac{d\sigma(\boldsymbol{q}\to\alpha\boldsymbol{k})}{d\Omega} \simeq \frac{g^2}{8\pi^2}\frac{Mk}{q}\,|\langle\alpha|\,e^{-i\boldsymbol{k}\cdot\boldsymbol{R}}|\psi_{\boldsymbol{q}}^{(+)}\rangle|^2. \qquad \text{(XXI.137)}$$

This expression can be obtained directly or can be deduced from the preceding one by applying the microreversibility relation (XIX.190) to the two reactions $\boldsymbol{q}\to\alpha\boldsymbol{k}$ and $(K\alpha)(-\boldsymbol{k})\to(-\boldsymbol{q})$, which are time reversal transforms one of the other. Indeed:

$$\langle\alpha|\,e^{-i\boldsymbol{k}\cdot\boldsymbol{R}}|\psi_{\boldsymbol{q}}^{(+)}\rangle = \langle\psi_{-\boldsymbol{q}}^{(-)}|\,e^{-i\boldsymbol{k}\cdot\boldsymbol{R}}(K|\alpha\rangle).$$

III. CLASSICAL THEORY OF ELECTROMAGNETIC RADIATION

17. The Equations of the Classical Maxwell–Lorentz Theory

The classical theory of radiation is based on Maxwell's equations [1]:

$$\operatorname{curl}\mathscr{E} + \frac{\partial\mathscr{H}}{\partial t} = 0 \qquad \text{(XXI.138}a\text{)}$$

$$\operatorname{div}\mathscr{H} = 0 \qquad \text{(XXI.138}b\text{)}$$

[1] Contrary to what was done in treating the scalar field, we shall not use rationalized units here, whence the presence of factors 4π in the right-hand sides of equations (XXI.139).

$$\operatorname{curl} \mathscr{H} - \frac{\partial \mathscr{E}}{\partial t} = 4\pi \boldsymbol{s} \qquad \text{(XXI.139}a\text{)}$$

$$\operatorname{div} \mathscr{E} = 4\pi s^0. \qquad \text{(XXI.139}b\text{)}$$

These equations determine the evolution of the electric and magnetic fields $\mathscr{E}$ and $\mathscr{H}$ in the presence of a distribution of electric charge of density s^0 and a distribution of electric current of density $\boldsymbol{s}$. $\boldsymbol{s}$ and s^0 satisfy the equation of continuity

$$\operatorname{div} \boldsymbol{s} + \frac{\partial s^0}{\partial t} = 0,$$

an evident consequence of equations (XXI.139), from which we obtain the law of conservation of electric charge.

Maxwell's equations are complemented by the Lorentz equation that determines the motion of electric charges in the presence of an electromagnetic field. According to this equation, the density of inertial force of a charge-carrying fluid is balanced at each instant by the electromagnetic force density:

$$\boldsymbol{f} \equiv s^0 \mathscr{E} + (\boldsymbol{s} \times \mathscr{H}). \qquad \text{(XXI.140)}$$

According to the classical Lorentz theory of the electron, matter is exclusively made up of particles, each of which has a determined mass and charge. To obtain the equation of motion for each particle, we need only to take the limit of the Lorentz equation when the charge is concentrated in an infinitely small volume. Consider for example a particle of mass m and of charge e and denote by $\boldsymbol{R}$, $\boldsymbol{V}$ and $\boldsymbol{\Pi}$ its position, velocity and mechanical momentum respectively [$\boldsymbol{\Pi} = M\boldsymbol{V}$; M is the relativistic mass according to definition (XX.20)]. The densities s^0 and $\boldsymbol{s}$ — more exactly, the contributions of the particle to these quantities — are given by the relations:

$$s^0 = e\,\delta(\boldsymbol{r} - \boldsymbol{R}), \qquad \boldsymbol{s} = e\boldsymbol{V}\delta(\boldsymbol{r} - \boldsymbol{R}). \qquad \text{(XXI.141)}$$

The electromagnetic force $\boldsymbol{F}$ felt by the particle is obtained by taking the integral of $\boldsymbol{f}$ over a small domain about $\boldsymbol{R}$; the Lorentz equation gives [cf. eq. (XX.22)]:

$$\frac{d\boldsymbol{\Pi}}{dt} = \boldsymbol{F} \equiv e(\mathscr{E} + \boldsymbol{V} \times \mathscr{H}). \qquad \text{(XXI.142)}$$

In this equation $\mathscr{E}$ and $\mathscr{H}$ represent the values taken by the electric and magnetic fields at the point $\boldsymbol{R}$. Recall that

$$\boldsymbol{V} \equiv \frac{\mathrm{d}\boldsymbol{R}}{\mathrm{d}t} \equiv \frac{\boldsymbol{\Pi}}{M} \equiv \frac{\boldsymbol{\Pi}}{\sqrt{m^2 + \boldsymbol{\Pi}^2}} \tag{XXI.143}$$

Equations (XXI.142) and (XXI.143) are the equations of motion of the particle in the presence of the field.

All of the preceding equations and definitions may be rewritten in covariant form. $\mathscr{E}$ and $\mathscr{H}$ form an antisymmetrical tensor $F_{\mu\nu}$ in accordance with definition (XX.9); s^0 and $\boldsymbol{s}$ form the current-density four-vector $s_\mu \equiv (s^0, \boldsymbol{s})$. The Maxwell–Lorentz equations become [1]):

$$\varepsilon^{\varkappa\lambda\mu\nu}\partial_\lambda F_{\mu\nu} = 0 \tag{XXI.138'}$$

$$\partial_\mu F^{\mu\nu} = 4\pi s^\nu \tag{XXI.139'}$$

and the equation of continuity becomes $\partial_\nu s^\nu = 0$. The force-density $\boldsymbol{f}$ is the space part of the four-vector:

$$f^\mu \equiv F^{\mu\nu}s_\nu. \tag{XXI.140'}$$

Similarly, $\boldsymbol{F}\,\mathrm{d}t$ is the space part of the four-vector $eF^{\mu\nu}\,\mathrm{d}x_\nu$; $\boldsymbol{\Pi}$ that of the four-vector $\Pi^\mu \equiv (M, \boldsymbol{\Pi})$ defined in § XX.4. Equation (XXI.142), complemented by an analogous equation for M, may be written in covariant form by introducing the proper time of the particle and its four-velocity.

18. Symmetries and Conservation Laws of the Classical Theory

The relativistic invariance of the Classical Theory leads to a certain number of symmetry properties and conservation laws. Here we shall consider in some detail the conservation of energy and of momentum, and briefly mention the conservation of angular momentum and the symmetry properties in reflection and time reversal. To simplify the discussion, we shall suppose that the system has only a single particle (an electron); the generalization to systems of several particles is obvious.

---

[1]) By introducing the dual of the electromagnetic tensor: $\hat{F}^{\varkappa\lambda} \equiv \frac{1}{2}\varepsilon^{\varkappa\lambda\mu\nu}F_{\mu\nu}$ equation (XXI.138') can be put in the simpler form:

$$\partial_\lambda \hat{F}^{\varkappa\lambda} = 0.$$

In order to obtain the energy and momentum conservation laws, we first show that f^μ, defined by equation (XXI.140′), is equal to the divergence of a certain tensor when $F^{\mu\nu}$ obeys the Maxwell–Lorentz equations.

From equation (XXI.139′),

$$f^\mu \equiv F^{\mu\nu} s_\nu = \frac{1}{4\pi} F^{\mu\nu}(\partial^\varrho F_{\varrho\nu}) = \frac{1}{4\pi}[\partial^\varrho(F^{\mu\nu}F_{\varrho\nu}) - F_{\varrho\nu}(\partial^\varrho F^{\mu\nu})]. \tag{XXI.144}$$

However, (XXI.138′) gives, taking into account that F is an antisymmetrical tensor:

$$F_{\varrho\nu}(\partial^\varrho F^{\mu\nu}) = \tfrac{1}{2}F_{\varrho\nu}(\partial^\varrho F^{\mu\nu} + \partial^\nu F^{\varrho\mu}) = \tfrac{1}{2}F_{\varrho\nu}(\partial^\mu F^{\varrho\nu}) = \tfrac{1}{4}\partial^\mu(F_{\varrho\nu}F^{\varrho\nu}) = \tfrac{1}{4}g^{\mu\varrho}\partial_\varrho(F_{\varkappa\lambda}F^{\varkappa\lambda}).$$

Substituting this expression into the right-hand side of (XXI.144) and putting:

$$T^{\mu\varrho} \equiv \frac{1}{4\pi}(F^{\mu\nu}F^\varrho{}_\nu - \tfrac{1}{4}g^{\mu\varrho}F_{\varkappa\lambda}F^{\varkappa\lambda}) \tag{XXI.145}$$

we obtain:

$$f^\mu = \partial_\varrho T^{\mu\varrho}. \tag{XXI.146}$$

This is the desired result.

$T^{\mu\varrho}$ is the "energy-momentum" tensor. Explicitly writing out the field-components in its definition, we have:

$$T^{00} = -\frac{1}{8\pi}(\mathscr{E}^2 + \mathscr{H}^2)$$

$$T^{k0} = T^{0k} = -\frac{1}{4\pi}[\mathscr{E} \times \mathscr{H}]$$

$$T^{kl} = T^{lk} = \frac{1}{4\pi}[\mathscr{E}_k\mathscr{E}_l + \mathscr{H}_k\mathscr{H}_l + \tfrac{1}{2}\delta_{kl}(\mathscr{H}^2 - \mathscr{E}^2)].$$

One defines the energy W^0 and the momentum $\boldsymbol{W}$ of the field by the following integrals extended over all space:

$$W^0 \equiv \frac{1}{8\pi}\int(\mathscr{E}^2 + \mathscr{H}^2)\,\mathrm{d}\boldsymbol{r} \tag{XXI.147a}$$

$$\boldsymbol{W} \equiv \frac{1}{4\pi}\int(\mathscr{E} \times \mathscr{H})\,\mathrm{d}\boldsymbol{r} \tag{XXI.147b}$$

i.e.:

$$W^{\mu} = -\int T^{\mu 0}\, \mathrm{d}\boldsymbol{r}. \qquad \text{(XXI.148)}$$

It is not *a priori* evident that W_{μ} is a four-vector. However, by using the energy and momentum conservation laws one can prove that it is.

To demonstrate the energy and momentum conservation laws, we put the equation of motion of the particle in the form

$$\frac{\mathrm{d}\Pi^{\mu}}{\mathrm{d}t} = \int f^{\mu}\, \mathrm{d}\boldsymbol{r}.$$

This equation is satisfied irrespective of the volume of integration V provided that it contains the point representing the position of the particle. Substituting for f^{μ} from (XXI.146), we obtain

$$\frac{\mathrm{d}\Pi^{\mu}}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\int_V T^{\mu 0}\, \mathrm{d}\boldsymbol{r}\right) + \int_V \left(\sum_k \frac{\partial T^{\mu k}}{\partial x^k}\right) \mathrm{d}\boldsymbol{r}.$$

The second integral can easily be transformed into an integral over the boundary surface of the volume V. Let us now extend the volume V to cover the whole of space. Under the hypothesis, entirely legitimate here, that the electromagnetic field vanishes sufficiently rapidly at infinity, the surface integral then tends to zero and one finds, taking definition (XXI.148) into account:

$$\frac{\mathrm{d}}{\mathrm{d}t}\,(\Pi^{\mu} + W^{\mu}) = 0.$$

This justifies the definition of the energy and momentum of the field adopted above, and constitutes the law of conservation of the energy and momentum of the total system (particle + field):

$$M + W^0 = Cst. \text{ in time.} \qquad \text{(XXI.149}a\text{)}$$

$$\boldsymbol{\Pi} + \boldsymbol{W} = Cst. \text{ in time.} \qquad \text{(XXI.149}b\text{)}$$

A similar argument leads to the following law for the conservation of the total angular momentum:

$$(\boldsymbol{R} \times \boldsymbol{\Pi}) + \boldsymbol{I} = Cst. \text{ in time} \qquad \text{(XXI.150)}$$

if we take the following definition for the angular momentum of the field:

$$\boldsymbol{I} \equiv \frac{1}{4\pi}\int [\boldsymbol{r} \times (\mathscr{E} \times \mathscr{H})]\, \mathrm{d}\boldsymbol{r}. \qquad \text{(XXI.151)}$$

If the system contains several particles, $\mathbf{\Pi}$ and M must be replaced by the sum of the momenta and the sum of the relativistic masses, and $(\mathbf{R} \times \mathbf{\Pi})$ by the sum of the angular momenta. With these obvious modifications, the three conservation laws (XXI.149*a*), (XXI.149*b*) and (XXI.150) remain valid.

Let us now consider the space and time reflections.

The equations of motion are invariant in a spatial reflection $(t \to t, \mathbf{r} \to -\mathbf{r})$ if we treat the charge e as a true scalar and the field $F^{\mu\nu}$ as a true tensor. With this convention s^0 transforms like a scalar function of three-dimensional space, $\mathbf{s}$ and $\mathscr{E}$ like polar-vector fields and $\mathscr{H}$ like an axial-vector field; in particular we have the transformation laws:

$$\mathscr{E}(t, \mathbf{r}) \to -\mathscr{E}(t, -\mathbf{r}), \quad \mathscr{H}(t, \mathbf{r}) \to +\mathscr{H}(t, -\mathbf{r}). \qquad \text{(XXI.152)}$$

The equations of motion are reversible in time $(t \to -t, \mathbf{r} \to \mathbf{r})$, if we treat the charge e as a scalar and $F^{\mu\nu}$ as a "pseudotensor". With this convention, s^0 and $\mathscr{E}$, considered as functions of $\mathbf{r}$, are invariant under time reversal, $\mathbf{s}$ and $\mathscr{H}$ change sign:

$$\mathscr{E}(t, \mathbf{r}) \to +\mathscr{E}(-t, \mathbf{r}), \quad \mathscr{H}(t, \mathbf{r}) \to -\mathscr{H}(-t, \mathbf{r}). \qquad \text{(XXI.153)}$$

19. Self-energy and Classical Radius of the Electron

At this point, we should mention a serious difficulty of the classical theory of radiation. In all of the preceding work, the charge of each particle is supposed to be concentrated in a very small volume, for example a small sphere whose radius a is made to tend to zero. In fact, the hypothesis of exact point-charges is not self consistent. Let us consider an isolated electron at rest at the origin of coordinates, and calculate the energy δm of the field created by this electron; this energy is commonly called the *electrostatic self-energy*. It is really a *mass renormalization* term, for in accordance with the argument of § 11, which applies here without change, the experimentally observed mass is:

$$m_{\text{exp}} = m + \delta m.$$

Since the electric field created in the external region $(r > a)$ is $e\mathbf{r}/r^3$, formula (XXI.147*a*) gives:

$$\delta m > \frac{e^2}{8\pi} \int \frac{\mathrm{d}\mathbf{r}}{r^4} = \frac{e^2}{2a}$$

and, since $\delta m < m_{\text{exp}}$, one has: $a > e^2/2m_{\text{exp}}$. The length:

$$r_0 = \frac{e^2}{m_{\text{exp}}} \qquad (\simeq 2.82 \times 10^{-13} \text{ cm})$$

is called the *classical radius of the electron*. The classical electron is therefore an object of finite dimensions, of the order of r_0. However, the theory is incapable of accounting for the stability of such objects, and *a fortiori*, for the details of their internal structure, i.e. the distribution of charges and currents in the region $r < r_0$. This is because the electromagnetic forces between the constitutive elements of the electron are in the main repulsive forces which tend to dissociate it; for the electron to be stable, they must be counterbalanced by cohesive forces of non-electromagnetic origin. The introduction of such forces in the case of a relativistic theory encounters serious difficulties. In practice, one contents oneself with *ad hoc* hypotheses concerning the internal structure of the electron; we must therefore expect to be able to correctly account only for phenomena which do not critically depend on the details of that structure, i.e. those involving sufficiently large wavelengths ($\lambda \gg r_0$) [1].

20. Electromagnetic Potential. Choice of the Gauge

Equations that are equivalent to equations (XXI.138) but simpler can be obtained by introducing the vector potential $\boldsymbol{A}$ and the scalar potential φ according to the definitions [cf. eq. (XX.7)]:

$$\mathscr{H} = \text{curl } \boldsymbol{A}, \qquad \mathscr{E} = -\frac{\partial \boldsymbol{A}}{\partial t} - \text{grad } \varphi. \qquad \text{(XXI.154)}$$

Equations (XXI.138) are then automatically satisfied. Equations (XXI.139) are equivalent to the equations:

$$\ddot{\boldsymbol{A}} - \triangle \boldsymbol{A} + \text{grad } (\dot{\varphi} + \text{div } \boldsymbol{A}) = 4\pi \boldsymbol{s} \qquad \text{(XXI.155}a\text{)}$$

$$\ddot{\varphi} - \triangle \varphi - \frac{\partial}{\partial t} (\dot{\varphi} + \text{div } \boldsymbol{A}) = 4\pi s^0. \qquad \text{(XXI.155}b\text{)}$$

Equations (XXI.154) define $\boldsymbol{A}$ and φ only to within an arbitrary

[1] The Quantum Theory also has an "ultraviolet catastrophe" of similar origin. Nevertheless, the analogy between the Classical and Quantum Theories must not be pushed too far. Effects like pair creation lead one to attribute the "quantum electron" with a radius of the order of $\hbar/mc$, that is 137 times greater than r_0.

function; if $G(t, \boldsymbol{r})$ is an arbitrary function of space and time, these equations remain satisfied if we make the substitutions:

$$\boldsymbol{A} \to \boldsymbol{A} - \operatorname{grad} G, \qquad \varphi \to \varphi + \frac{\partial G}{\partial t}. \tag{XXI.156}$$

Such a set of substitutions is called a *gauge transformation* (cf. § XX.20).

All these relations can be written in a completely covariant way. The potentials $\boldsymbol{A}$ and φ form a four-vector A^μ [cf. eq. (XX.6)]; according to equations (XXI.154), A^μ is the vector whose curl is $F^{\mu\nu}$ [cf. eq. (XX.8)], which defines it to within the gradient of an arbitrary function; the gauge transformation (XXI.156) is nothing else but the addition of a gradient:

$$A^\mu \to A^\mu + \partial^\mu G \qquad (G \text{ arbitrary}). \tag{XXI.157}$$

Equations (XXI.155) read

$$\Box A^\mu - \partial^\mu(\partial_\nu A^\nu) = 4\pi s^\mu. \tag{XXI.158}$$

The arbitrary in the gauge can be partly removed without violating the formal covariance of the theory by imposing the *Lorentz supplementary condition*

$$\partial_\nu A^\nu = 0. \tag{XXI.159}$$

The equation of motion of the potential then takes the simpler form

$$\Box A^\mu = 4\pi s^\mu. \tag{XXI.160}$$

Let χ be an arbitrary solution of the equation: $\Box\chi = 0$. The Lorentz condition fixes the gauge of the potential to within a function of this type. In other words, the definition of the field $F^{\mu\nu}$ and equations (XXI.159) and (XXI.160) are invariant in the *Lorentz gauge transformations* (sometimes called special gauge transformations):

$$A^\mu \to A^\mu + \partial^\mu \chi \qquad (\Box\chi = 0). \tag{XXI.161}$$

One can also obtain very simple equations of motion by imposing the condition

$$\operatorname{div} \boldsymbol{A} = 0. \tag{XXI.162}$$

Contrary to the Lorentz condition, (XXI.162) does not preserve the formal covariance of the theory. Nevertheless it has the advantage

of completely removing the arbitrary in the gauge. The gauge thus defined is currently called the *radiation gauge*. It is the one that will be adopted in what is to follow. Before studying radiation theory in this gauge, we must recall an important property of the decomposition of vector fields.

21. Longitudinal and Transverse Parts of a Vector Field

Any real or complex vector field $\boldsymbol{B}(\boldsymbol{r})$ can be considered as a superposition of two fields:

$$\boldsymbol{B}(\boldsymbol{r}) = \boldsymbol{B}_{\parallel}(\boldsymbol{r}) + \boldsymbol{B}_{\perp}(\boldsymbol{r}) \tag{XXI.163}$$

the one of which, $\boldsymbol{B}_{\parallel}$, is irrotational and the other, $\boldsymbol{B}_{\perp}$, divergenceless. One has:

$$\operatorname{curl} \boldsymbol{B}_{\parallel} = 0, \qquad \operatorname{div} \boldsymbol{B}_{\parallel} = \operatorname{div} \boldsymbol{B} \tag{XXI.164}$$

$$\operatorname{div} \boldsymbol{B}_{\perp} = 0, \qquad \operatorname{curl} \boldsymbol{B}_{\perp} = \operatorname{curl} \boldsymbol{B}. \tag{XXI.165}$$

If one limits oneself to square-integrable vector fields, this decomposition is unique [1]). By definition the fields $\boldsymbol{B}_{\parallel}$ and $\boldsymbol{B}_{\perp}$ thus defined constitute respectively the *longitudinal part and the transversal part of the field* $\boldsymbol{B}$. In addition we have

$$\boldsymbol{B}_{\parallel} = \operatorname{grad} U \qquad U(\boldsymbol{r}) = -\frac{1}{4\pi} \int \frac{\operatorname{div} \boldsymbol{B}(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} \, \mathrm{d}\boldsymbol{r}' \tag{XXI.166}$$

$$\boldsymbol{B}_{\perp} = \operatorname{curl} \boldsymbol{V} \qquad \boldsymbol{V}(\boldsymbol{r}) = \frac{1}{4\pi} \int \frac{\operatorname{curl} \boldsymbol{B}(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} \, \mathrm{d}\boldsymbol{r}'. \tag{XXI.167}$$

This decomposition can easily be effected using the Fourier transformation. Denote the Fourier transforms of $\boldsymbol{B}$, $\boldsymbol{B}_{\perp}$ and $\boldsymbol{B}_{\parallel}$ by $\boldsymbol{b}$, $\boldsymbol{b}_{\perp}$ and $\boldsymbol{b}_{\parallel}$ respectively:

$$\begin{aligned} \boldsymbol{b}(\boldsymbol{k}) &= (2\pi)^{-3/2} \int \boldsymbol{B}(\boldsymbol{r}) \, \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \, \mathrm{d}\boldsymbol{r} \\ \boldsymbol{b}_{\parallel}(\boldsymbol{k}) &= \ldots \qquad\qquad \text{etc.} \end{aligned} \tag{XXI.168}$$

Equations (XXI.164) and (XXI.165) are respectively equivalent to the equations

$$\boldsymbol{k} \times \boldsymbol{b}_{\parallel} = 0 \qquad \boldsymbol{k} \cdot \boldsymbol{b}_{\parallel} = \boldsymbol{k} \cdot \boldsymbol{b}$$

$$\boldsymbol{k} \cdot \boldsymbol{b}_{\perp} = 0 \qquad \boldsymbol{k} \times \boldsymbol{b}_{\perp} = \boldsymbol{k} \times \boldsymbol{b}.$$

---

[1]) We must also suppose $\boldsymbol{B}_{\parallel}$ and $\boldsymbol{B}_{\perp}$ square-integrable. This decomposition will sometimes be made for fields which are not square-integrable; the longitudinal and transverse parts will then be unambiguously defined by equations (XXI.166) and (XXI.167) respectively.

$\boldsymbol{b}_{\|}$ is the projection of $\boldsymbol{b}$ along the vector $\boldsymbol{k}$, whence the name of longitudinal part [1]) for $\boldsymbol{B}_{\|}$; $\boldsymbol{b}_{\perp}$ is the projection of $\boldsymbol{b}$ perpendicular to the vector $\boldsymbol{k}$, whence the name of transverse part for $\boldsymbol{B}_{\perp}$. Equations (XXI.166) and (XXI.167) are respectively equivalent to the equations:

$$\boldsymbol{b}_{\|} = \frac{\boldsymbol{k}(\boldsymbol{b}\cdot\boldsymbol{k})}{k^2} \qquad \boldsymbol{b}_{\perp} = \frac{\boldsymbol{k}\times(\boldsymbol{b}\times\boldsymbol{k})}{k^2}. \tag{XXI.169}$$

The Fourier transform of the function U [definition (XXI.166)] is:

$$u(\boldsymbol{k}) = -\,\mathrm{i}\,\frac{(\boldsymbol{b}\cdot\boldsymbol{k})}{k^2}.$$

[*N.B.* The Fourier transform of $1/r$ is $\sqrt{2\pi}/\pi k^2$.]

In a more general way, the square-integrable vector-field space — that is, the space of the wave function of a particle of spin 1 — is the direct sum of two orthogonal subspaces, the longitudinal-field space and the transverse-field space. The decomposition envisaged here consists in writing that $\boldsymbol{B}$ is the sum of its projections on these two complementary subspaces. These projections can automatically be obtained by choosing an appropriate basis. Denote by

$$\boldsymbol{L}^{(1)}, \boldsymbol{L}^{(2)}, \ldots, \boldsymbol{L}^{(\lambda)}, \ldots$$

a complete set of orthonormal longitudinal fields and by

$$\boldsymbol{T}^{(1)}, \boldsymbol{T}^{(2)}, \ldots, \boldsymbol{T}^{(\tau)}, \ldots$$

a complete set of orthonormal transverse fields:

$$\operatorname{curl} \boldsymbol{L}^{(\lambda)} = 0, \qquad \int (\boldsymbol{L}^{(\lambda)*}\cdot\boldsymbol{L}^{(\lambda')})\,\mathrm{d}\boldsymbol{r} = \delta_{\lambda\lambda'} \tag{XXI.170a}$$

$$\operatorname{div} \boldsymbol{T}^{(\tau)} = 0, \qquad \int (\boldsymbol{T}^{(\tau)*}\cdot\boldsymbol{T}^{(\tau')})\,\mathrm{d}\boldsymbol{r} = \delta_{\tau\tau'} \tag{XXI.170b}$$

$$\sum_{\lambda} L_i^{(\lambda)}(\boldsymbol{r})\, L_j^{(\lambda)*}(\boldsymbol{r}') + \sum_{\tau} T_i^{(\tau)}(\boldsymbol{r})\, T_j^{(\tau)*}(\boldsymbol{r}') = \delta_{ij}\,\delta(\boldsymbol{r}-\boldsymbol{r}') \tag{XXI.170c}$$

$$(i, j = x, y \text{ or } z).$$

Let us put

$$\beta_{\|}^{(\lambda)} = \int (\boldsymbol{L}^{(\lambda)*}\cdot\boldsymbol{B})\,\mathrm{d}\boldsymbol{r}, \qquad \beta_{\perp}^{(\tau)} = \int (\boldsymbol{T}^{(\tau)*}\cdot\boldsymbol{B})\,\mathrm{d}\boldsymbol{r}. \tag{XXI.171}$$

---

[1]) $\boldsymbol{B}_{\|}$ is not to be confused with the *radial* part of $\boldsymbol{B}$ which is the projection of $\boldsymbol{B}$ along the vector $\boldsymbol{r}$; $\boldsymbol{B}_{\mathrm{rad}} = \boldsymbol{r}(\boldsymbol{B}\cdot\boldsymbol{r})/r^2$. Note that the radial part of a purely transverse field does not, in general, vanish.

One then has:

$$\boldsymbol{B}_{\parallel} = \sum_{\lambda} \beta_{\parallel}^{(\lambda)} \boldsymbol{L}^{(\lambda)}, \qquad \boldsymbol{B}_{\perp} = \sum_{\tau} \beta_{\perp}^{(\tau)} \boldsymbol{T}^{(\tau)}. \tag{XXI.172}$$

The use of the Fourier transform amounts to choosing a particular basis, that of the fields of longitudinal and transverse plane waves. With each wave vector there is associated a longitudinal field:

$$\boldsymbol{L}^{(\boldsymbol{k})} = (2\pi)^{-3/2} \hat{\boldsymbol{k}}\, e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \qquad (\hat{\boldsymbol{k}} \equiv \boldsymbol{k}/k) \tag{XXI.173a}$$

and two orthogonal transverse fields:

$$\boldsymbol{T}^{(\boldsymbol{k}\varpi)} = (2\pi)^{-3/2} \boldsymbol{\epsilon}^{(\varpi)} e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \qquad (\varpi = 1, 2) \tag{XXI.173b}$$

$\boldsymbol{\epsilon}^{(1)}$ and $\boldsymbol{\epsilon}^{(2)}$ are two arbitrary unit vectors [1]) orthogonal to $\boldsymbol{k}$ and orthogonal to each other:

$$\boldsymbol{k}\cdot\boldsymbol{\epsilon}^{(1)} = \boldsymbol{k}\cdot\boldsymbol{\epsilon}^{(2)} = 0 \tag{XXI.174a}$$

$$\boldsymbol{\epsilon}^{(1)*}\cdot\boldsymbol{\epsilon}^{(1)} = \boldsymbol{\epsilon}^{(2)*}\cdot\boldsymbol{\epsilon}^{(2)} = 1, \qquad \boldsymbol{\epsilon}^{(1)*}\cdot\boldsymbol{\epsilon}^{(2)} = 0. \tag{XXI.174b}$$

The components of $\boldsymbol{B}$ in this basis [definitions (XXI.171)] are related to the vectors $\boldsymbol{b}_{\parallel}$ and $\boldsymbol{b}_{\perp}$ by the formulas [cf. eq. (XXI.169)]:

$$\boldsymbol{b}_{\parallel} = \beta_{\parallel}^{(\boldsymbol{k})} \hat{\boldsymbol{k}}, \qquad \boldsymbol{b}_{\perp} = \sum_{\varpi} \beta_{\perp}^{(\boldsymbol{k}\varpi)} \boldsymbol{\epsilon}^{(\varpi)}. \tag{XXI.175}$$

The plane waves here depend on the continuous index $\boldsymbol{k}$ and are normalized to $\delta(\boldsymbol{k}-\boldsymbol{k}')$. In order to have only discrete indices, one can, according to the usual artifice, enclose the system in a cube of side L; the constant $(2\pi)^{-3/2}$ in definitions (XXI.173*a*, *b*) must then be replaced by $L^{-3/2}$.

Among the other bases currently used are the spherical waves. We shall return to them in § 29.

22. Elimination of the Longitudinal Field

Among Maxwell's equations, equations (XXI.138*b*) and (XXI.139*b*) are not, properly speaking, equations of motion, but rather constraints imposed on the fields $\mathscr{E}$ and $\mathscr{H}$. They fix the longitudinal parts of these fields.

---

[1]) These vectors are not necessarily real. By definition a vector is real if its Cartesian components are real; the vector $\boldsymbol{k}$ is a real vector. The scalar product of two complex vectors $\boldsymbol{A}$ and $\boldsymbol{B}$ is by definition:

$$\boldsymbol{A}^*\cdot\boldsymbol{B} \equiv A_x^* B_x + A_y^* B_y + A_z^* B_z.$$

We see from equation (XXI.138*b*) that $\mathscr{H}_{\parallel} = 0$ and from equation (XXI.139*b*) that $\mathscr{E}_{\parallel}(t, \mathbf{r})$ *is the electrostatic field created by the charge distribution* $s_0(t, \mathbf{r})$. In order to define the dynamical state of the system, it is therefore sufficient to specify the charge distributions and currents – that is, the positions and the velocities of the particles – on the one hand, and the transverse fields $\mathscr{H}$ and $\mathscr{E}_{\perp}$ on the other. It is therefore possible to reformulate the theory eliminating completely the longitudinal part of the electromagnetic field.

Rather than proceed directly to this elimination, it is preferable to first introduce the potential $(\varphi, \boldsymbol{A})$ in the radiation gauge [cf. eq. (XXI.162)]. This is merely a change of variable. The field $\boldsymbol{A}$ is by definition purely transverse:

$$\boldsymbol{A}_{\parallel} = 0, \qquad \boldsymbol{A} = \boldsymbol{A}_{\perp}. \tag{XXI.176}$$

The old variables are given in terms of the new by equations (XXI.154), or:

$$\mathscr{E}_{\parallel} = -\operatorname{grad} \varphi \tag{XXI.177}$$

$$\mathscr{H} = \operatorname{curl} \boldsymbol{A}, \qquad \mathscr{E}_{\perp} = -\frac{\partial \boldsymbol{A}}{\partial t}. \tag{XXI.178}$$

φ is determined by equation (XXI.155*b*), which in this case is written:

$$\triangle \varphi = -4\pi s^0. \tag{XXI.179}$$

One finds:

$$\varphi(t, \mathbf{r}) = \int \frac{s^0(t, \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \, \mathrm{d}\mathbf{r}' \tag{XXI.180}$$

from which one obtains the expression mentioned above for $\mathscr{E}_{\parallel}$ in terms of s^0. To eliminate φ from equation (XXI.155*a*) it is sufficient to separately write down the equality of the longitudinal and transverse components. The first of these equalities:

$$\operatorname{grad} \dot{\varphi} = 4\pi \boldsymbol{s}_{\parallel} \tag{XXI.181}$$

is identically satisfied if φ satisfies equation (XXI.180) [1]).

The second is independent of φ, and it is the equation of motion:

$$\ddot{\boldsymbol{A}} - \triangle \boldsymbol{A} = 4\pi \boldsymbol{s}_{\perp}. \tag{XXI.182}$$

---

[1]) For this equality to be satisfied it is sufficient that the divergence of each side be equal: $\Delta \dot{\varphi} = 4\pi \operatorname{div} \boldsymbol{s}$; this equation can be deduced from equation (XXI.179) by differentiating both sides of the latter with respect to the time and using the equation of continuity.

It remains to eliminate the longitudinal field from the equations of motion of the particles [eqs. (XXI.142) and (XXI.143)], that is from the expression for the Lorentz force to which each of them is subjected. The different quantities attached to each particle are denoted by the same symbols as in § 17, with an index 1, 2, ..., n, ... to distinguish between them. Thus the contribution of the nth particle to the charge density is:

$$s_n{}^0 = e_n \delta(\boldsymbol{r} - \boldsymbol{R}_n).$$

Equation (XXI.179) is therefore written:

$$\triangle \varphi = -4\pi \sum_n e_n \, \delta(\boldsymbol{r} - \boldsymbol{R}_n) \tag{XXI.183}$$

and its solution is the Coulomb potential:

$$\varphi = \sum_n \frac{e_n}{|\boldsymbol{r} - \boldsymbol{R}_n|}. \tag{XXI.184}$$

Let $\boldsymbol{F}_i$ be the Lorentz force acting on the ith particle. We can write:

$$\begin{gathered} \boldsymbol{F}_i = \boldsymbol{F}_i{}^{\mathrm{tr}} + \boldsymbol{F}_i{}^{\mathrm{long}} \\ \boldsymbol{F}_i{}^{\mathrm{tr}} = \mathrm{e}_i(\boldsymbol{\mathscr{E}}_\perp(i) + \boldsymbol{V}_i \times \boldsymbol{\mathscr{H}}(i)) = e_i\left(-\frac{\partial \boldsymbol{A}(i)}{\partial t} + \boldsymbol{V}_i \times \mathrm{curl}\, \boldsymbol{A}(i)\right) \\ \boldsymbol{F}_i{}^{\mathrm{long}} = e_i \, \boldsymbol{\mathscr{E}}_{\|}(i) = -\, e_i \, \mathrm{grad}\, \varphi(i) \end{gathered} \tag{XXI.185}$$

$\boldsymbol{\mathscr{E}}_\perp(i)$, $\boldsymbol{\mathscr{H}}(i)$, ... here denote the values taken by $\boldsymbol{\mathscr{E}}_\perp$, $\boldsymbol{\mathscr{H}}$, ... at the point $\boldsymbol{R}_i$. The desired elimination is effected by substituting expression (XXI.184) into $\boldsymbol{F}^{\mathrm{long}}$.

However, this substitution cannot be effected immediately, since the potential φ diverges at the point $\boldsymbol{R}_i$. The difficulty is due to the fact that we have assumed point charges. $\boldsymbol{F}_i{}^{\mathrm{long}}$ is the electrostatic force on the ith particle due to all of the charges present; we denote by $\boldsymbol{F}_i{}'$ the contribution to $\boldsymbol{F}_i{}^{\mathrm{long}}$ from the charge on the ith particle itself, and by $\boldsymbol{F}_i{}^{\mathrm{coul}}$ the contribution from all the other charges. So long as the particles are sufficiently well separated, the calculation of $\boldsymbol{F}_i{}^{\mathrm{coul}}$ may be carried through supposing point charges, giving:

$$\boldsymbol{F}_i{}^{\mathrm{coul}} = \sum_{n \neq i} e_i \, e_n \frac{(\boldsymbol{R}_i - \boldsymbol{R}_n)}{|\boldsymbol{R}_i - \boldsymbol{R}_n|^3}. \tag{XXI.186}$$

But the hypothesis of point charges is obviously not justified for

the calculation of $\boldsymbol{F}_i'$. In any case, whatever may be the distribution of charge inside the ith particle, one finds:

$$\boldsymbol{F}_i' = 0.$$

Thus, the force on each particle is the Lorentz force of the transverse field [eq. (XXI.185)] and the electrostatic force due to the presence of all of the *other* charges [eq. (XXI.186)]:

$$\boldsymbol{F}_i = \boldsymbol{F}_i^{\mathrm{tr}} + \boldsymbol{F}_i^{\mathrm{coul}}. \tag{XXI.187}$$

23. Energy, Momentum, Angular Momentum

Having eliminated the longitudinal field from the equations of motion, we now wish to eliminate it from the conservation laws for energy, momentum, and angular momentum.

By replacing $\mathscr{E}$ by $\mathscr{E}_\perp$ in the expression for the energy in the field [eq. (XXI.147a)], we obtain a new expression, U, depending only on the transverse components; to distinguish it from the preceding one we shall call it the *radiation energy*:

$$U \equiv \frac{1}{8\pi} \int (\mathscr{E}_\perp{}^2 + \mathscr{H}^2)\, \mathrm{d}\boldsymbol{r}. \tag{XXI.188}$$

We similarly define the *radiation momentum* $\boldsymbol{X}$ [cf. eq. (XXI.147b)] and the *radiation angular momentum* $\boldsymbol{G}$ [cf. eq. (XXI.151)]:

$$\boldsymbol{X} \equiv \frac{1}{4\pi} \int (\mathscr{E}_\perp \times \mathscr{H})\, \mathrm{d}\boldsymbol{r} \tag{XXI.189}$$

$$\boldsymbol{G} \equiv \frac{1}{4\pi} \int [\boldsymbol{r} \times (\mathscr{E}_\perp \times \mathscr{H})]\, \mathrm{d}\boldsymbol{r}. \tag{XXI.190}$$

U, $\boldsymbol{X}$ and $\boldsymbol{G}$ are respectively the energy, the momentum and the angular momentum of the system in the absence of charge.

They can be put in a form analogous to the one given in section I for the energy, momentum and angular momentum of the scalar field (cf. § 4, 6 and 7). By replacing $\mathscr{H}$ by curl $\boldsymbol{A}$ in expressions (XXI.188–190), by effecting a certain number of integrations by parts

and by using the fact that $\mathscr{E}_\perp$ and $\boldsymbol{A}$ are transverse fields vanishing at infinity, one finds:

$$U = \frac{1}{8\pi} \int \sum_i (\mathscr{E}_{\perp i}^2 + (\nabla A_i)^2)\, d\boldsymbol{r}, \tag{XXI.188'}$$

$$\boldsymbol{X} = \frac{1}{4\pi} \int \sum_i \mathscr{E}_{\perp i} (\nabla A_i)\, d\boldsymbol{r}, \tag{XXI.189'}$$

and

$$\boldsymbol{G} = \boldsymbol{G}^{(0)} + \boldsymbol{G}^{(s)}, \tag{XXI.190'}$$

where:

$$\boldsymbol{G}^{(0)} \equiv \frac{1}{4\pi} \int \sum_i \mathscr{E}_{\perp i} (\boldsymbol{l} A_i)\, d\boldsymbol{r}, \tag{XXI.191a}$$

$$\boldsymbol{G}^{(s)} \equiv \frac{1}{4\pi} \int (\mathscr{E}_\perp \times \boldsymbol{A})\, d\boldsymbol{r}. \tag{XXI.191b}$$

In order to obtain these expressions it is convenient to use the techniques and notations of the tensor calculus, and in particular to use the antisymmetrical tensor of 3-dimensional Euclidean space ε_{ijk} (definition given in note, p. 510) and the identity

$$\varepsilon_{ijk}\, \varepsilon_{klm} = \delta_{il}\, \delta_{jm} - \delta_{im}\, \delta_{jl}.$$

Thus

$$\mathscr{H}^2 = (\text{curl } \boldsymbol{A})^2 = \varepsilon_{ijk}\, \varepsilon_{klm} \frac{\partial A_j}{\partial x_i} \frac{\partial A_m}{\partial x_l} = \frac{\partial A_m}{\partial x_l} \frac{\partial A_m}{\partial x_l} - \frac{\partial A_l}{\partial x_m} \frac{\partial A_m}{\partial x_l}$$

$$= \frac{\partial A_m}{\partial x_l} \frac{\partial A_m}{\partial x_l} - \frac{\partial}{\partial x_l} \left(A_m \frac{\partial A_l}{\partial x_m} \right) + A_m \frac{\partial^2 A_l}{\partial x_m\, \partial x_l}$$

and since $\boldsymbol{A}$ is a transverse field vanishing at infinity:

$$\int \mathscr{H}^2\, d\boldsymbol{r} = \int \frac{\partial A_m}{\partial x_l} \frac{\partial A_m}{\partial x_l}\, d\boldsymbol{r} \equiv \int \sum_m (\nabla A_m)^2\, d\boldsymbol{r}$$

from which one obtains equation (XXI.188'). Using an analogous technique, one finds:

$$(\mathscr{E}_\perp \times \mathscr{H})_i = \varepsilon_{ijk}\, \varepsilon_{klm}\, \mathscr{E}_{\perp j} \frac{\partial A_m}{\partial x_l} = \mathscr{E}_{\perp m} \frac{\partial}{\partial x_i} A_m - \frac{\partial}{\partial x_l} (\mathscr{E}_{\perp l} A_i),$$

from which one obtains equation (XXI.189'). Similarly,

$$[\boldsymbol{r} \times (\mathscr{E}_\perp \times \mathscr{H})]_i = \mathscr{E}_{\perp m} \left(\varepsilon_{ijk}\, x_j \frac{\partial A_m}{\partial x_k} \right) - \frac{\partial}{\partial x_l} (\varepsilon_{ijk}\, x_j\, \mathscr{E}_{\perp l}\, A_k) + \varepsilon_{ijk}\, \mathscr{E}_{\perp j}\, A_k$$

from which one obtains equation (XXI.190') with definitions (XXI.191*a*, *b*) of $\boldsymbol{G}^{(0)}$ and $\boldsymbol{G}^{(s)}$.

The comparison of these expressions with the corresponding expressions for the scalar field is instructive. Let us compare, for example, expressions (XXI.30) and (XXI.188′) for the energy of the field. Differences arise because: (*i*) $\boldsymbol{A}$ is a vector field, (*ii*) its mass is null ($\mu = 0$), (*iii*) the field units adopted here differ by a factor of $\sqrt{4\pi}$ from those used in section I. Apart from this, expression (XXI.188′) appears as a simple generalization of expression (XXI.30) if we suppose, as suggested by the second of equations (XXI.178) and as will be confirmed in what follows, that the momentum canonically conjugate to the field $\boldsymbol{A}$ is $(-\boldsymbol{\mathscr{E}}_\perp)$. The same can be said for the expressions for the momentum [eqs. (XXI.55) and (XXI.189′)] and the angular momentum [eqs. (XXI.61) and (XXI.190′)]. The vector character of the field is particularly apparent in the angular momentum; in addition to the term $\boldsymbol{G}^{(0)}$ obtained by a simple generalization of the expression for the angular momentum of the scalar field, there is a supplementary term $\boldsymbol{G}^{(s)}$; $\boldsymbol{G}^{(0)}$ represents the "orbital angular momentum" and $\boldsymbol{G}^{(s)}$ the "spin" of the field [1]).

We are now in a position to eliminate the longitudinal field from the expressions obtained in § 18 for the total energy, the total momentum and the total angular momentum of the system formed by the electromagnetic field and the charged particles interacting with this field.

Consider first the total momentum of the system. According to the definition given in § 18,

$$\begin{aligned} \boldsymbol{P}_{\text{tot}} &= \sum_n \boldsymbol{\Pi}_n + \boldsymbol{W} \\ &= \sum_n \boldsymbol{\Pi}_n + \boldsymbol{X} + \boldsymbol{X}', \end{aligned} \tag{XXI.192}$$

where

$$\boldsymbol{X}' \equiv \frac{1}{4\pi} \int \boldsymbol{\mathscr{E}}_{\|} \times \boldsymbol{\mathscr{H}}) \, d\boldsymbol{r}.$$

---

[1]) In the present case things are complicated by the transversality condition; among other consequences, it results from this that the amplitudes of the field $\boldsymbol{A}$ at each point in space do not form a set of independent dynamical variables. The particular form of the commutation relations given in Problem XXI.5 is due to these same reasons. Also for the same reasons, the vector operator $\boldsymbol{G}^{(0)}$ and $\boldsymbol{G}^{(s)}$ do not separately verify the commutation relations characteristic of angular momenta.

Replacing $\mathscr{E}_{||}$ and $\mathscr{H}$ by their definitions [eqs. (XXI.177–178)], we have

$$\begin{aligned}\mathscr{E}_{||} \times \mathscr{H} &= -(\operatorname{grad} \varphi \times \mathscr{H}) = \varphi \operatorname{curl} \mathscr{H} - \operatorname{curl} (\varphi \mathscr{H}) \\ &= -\varphi(\triangle \boldsymbol{A}) - \operatorname{curl} (\varphi \mathscr{H}).\end{aligned}$$

Integrating by parts and noting that φ and $\boldsymbol{A}$ vanish at infinity, we obtain:

$$\boldsymbol{X}' = -\frac{1}{4\pi} \int \varphi(\triangle \boldsymbol{A})\, d\boldsymbol{r} = -\frac{1}{4\pi} \int \boldsymbol{A}(\triangle \varphi)\, d\boldsymbol{r}.$$

According to equation (XXI.183), $\Delta\varphi$ is a sum of δ-functions; the integration is easily carried out and gives:

$$\boldsymbol{X}' = \sum_n e_n \boldsymbol{A}(n)$$

where, in accordance with the convention adopted above, $\boldsymbol{A}(n) \equiv \boldsymbol{A}(\boldsymbol{R}_n)$. Substituting the expression found into equation (XXI.192), we are led to *define the momentum* $\boldsymbol{P}_n$ *of the* n^{th} *particle* by the relation [1]:

$$\boldsymbol{P}_n \equiv \boldsymbol{\Pi}_n + e_n \boldsymbol{A}(n). \tag{XXI.193}$$

The total momentum is then written:

$$\boldsymbol{P}_{\text{tot}} = \sum_n \boldsymbol{P}_n + \boldsymbol{X}. \tag{XXI.194}$$

It is the sum of the momenta of the particles and the momentum of the radiation. This sum is a constant of the motion.

One proceeds in an exactly analogous fashion in dealing with the angular momentum. One is led to define the angular momentum $\boldsymbol{L}_n$ of the nth particle by the relation:

$$\boldsymbol{L}_n \equiv \boldsymbol{R}_n \times \boldsymbol{P}_n \tag{XXI.195}$$

$\boldsymbol{P}_n$ being the momentum as defined above. The calculation gives the following expression for the total angular momentum of the system:

$$\boldsymbol{J}_{\text{tot}} = \sum_n \boldsymbol{L}_n + \boldsymbol{G}. \tag{XXI.196}$$

---

[1] Recall that we are in radiation gauge and that the field $\boldsymbol{A}$ is purely transverse. More generally, we can obtain a definition *independent of gauge* by putting:

$$\boldsymbol{P}_n \equiv \boldsymbol{\Pi}_n + e_n \boldsymbol{A}_{\perp}(n).$$

Finally, we consider the total energy of the system. According to the definition given in § 17, it can be put in the form:

$$\sum_n M_n + U + U',$$

where

$$U' \equiv W^0 - U \equiv \frac{1}{8\pi} \int \mathscr{E}_{\parallel}^2 \, d\boldsymbol{r}.$$

Replacing $\mathscr{E}_{\parallel}$ by its definition and effecting an integration by parts, one finds:

$$U' = \frac{1}{8\pi} \int (\text{grad } \varphi)^2 \, d\boldsymbol{r} = -\frac{1}{8\pi} \int \varphi(\triangle\varphi) \, d\boldsymbol{r}.$$

φ is the electrostatic potential created by the electric charges present. If we assume point charges, as we have done up to the present, we find:

$$U' = \frac{1}{2} \sum_n \sum_p \frac{e_n e_p}{|\boldsymbol{R}_n - \boldsymbol{R}_p|}.$$

In this double sum, the "cross terms" ($n \neq p$) represent the Coulomb energy of the set of particles, namely:

$$H_{\text{coul}} \equiv \sum_{n<p} \frac{e_n e_p}{|\boldsymbol{R}_n - \boldsymbol{R}_p|}. \tag{XXI.197}$$

[The symbol $\sum_{n<p}$ denotes the sum over all pairs of particles.] The square terms are all infinite. The difficulty involved here is of the same origin as the one encountered in the calculation of $\boldsymbol{F}_i^{\text{long}}$ at the end of § 22. The ith square term represents the electrostatic energy of the ith particle in its own electric field; this energy, u_i', depends in a critical way on the distribution of charge inside the particle and becomes infinite in the limit when the radius of the particle tends to zero (cf. § 19); to obtain the correct value for the self-energy we must add to this term the contribution of the cohesive forces mentioned in § 19, a contribution not included in the present calculation. In the absence of a self-consistent theory of the internal structure of the particles, we shall assume that each of these "self-energy" terms can be omitted if we make an appropriate modification of the rest mass m_i in the expression for the mass energy M_i. We then obtain the following expression for the total energy of the system:

$$E_{\text{tot}} = \sum_n M_n + H_{\text{coul}} + U. \tag{XXI.198}$$

This is the sum of the mass and Coulomb energies of the particles and of the radiation energy.

24. Hamiltonian for Free Radiation

The equations of motion obtained in § 22 can be put in canonical form. We first treat the free radiation case, that is, the case of the electromagnetic field in the absence of charge.

The dynamical state of the radiation is defined at each instant by giving the transverse field $\boldsymbol{A}$ and its velocity $\partial\boldsymbol{A}/\partial t$. The electric and magnetic fields are related to these two vector fields by equations (XXI.178). Their evolution in the course of time is determined by the equation of motion (XXI.182) which, in the absence of charge and of current, reduces to:

$$\Box \boldsymbol{A} \equiv \ddot{\boldsymbol{A}} - \triangle \boldsymbol{A} = 0 \qquad (\text{div } \boldsymbol{A} = 0). \tag{XXI.199}$$

This is the same equation as we had for the free scalar field [cf. eq. (XXI.1)]. The only difference is that here we have a field of transverse vectors rather than a scalar field, and that the mass term is null. Apart from this, the canonical formalism can be established by proceeding in exactly the same way as for the scalar field in section I. Let $\boldsymbol{T}^{(1)}$, $\boldsymbol{T}^{(2)}$, ..., $\boldsymbol{T}^{(\tau)}$, ..., be a complete set of orthonormal transverse fields [cf. eqs. (XXI.170)] like the set defined in § 21; we further suppose that it is a basis of *real normal coordinates*:

$$\boldsymbol{T}^{(\tau)*} = \boldsymbol{T}^{(\tau)} \qquad (\triangle + k_\tau^2)\, \boldsymbol{T}^{(\tau)} = 0. \tag{XXI.200}$$

The corresponding normal coordinates q_τ, are defined by [cf. eq. (XXI.6)]:

$$q_\tau = \frac{1}{\sqrt{4\pi}} \int (\boldsymbol{T}^{(\tau)} \cdot \boldsymbol{A})\, d\boldsymbol{r}. \tag{XXI.201}$$

These are real functions of time satisfying the equation of motion:

$$\ddot{q}_\tau + k_\tau^2 q_\tau = 0, \tag{XXI.202}$$

characteristic of a harmonic oscillator of frequency k_τ. Such an oscillatory motion is given by the Hamiltonian

$$h_\tau = \tfrac{1}{2}(p_\tau^2 + k_\tau^2 q_\tau^2), \tag{XXI.203}$$

where p_τ is the momentum canonically conjugate to q_τ; p_τ is equal to $\dot{q}_\tau$ and is obtained from the "velocity" of the field, $\dot{\boldsymbol{A}}$, by the relation:

$$p_\tau = \frac{1}{\sqrt{4\pi}} \int (\boldsymbol{T}^{(\tau)} \cdot \dot{\boldsymbol{A}})\, \mathrm{d}\boldsymbol{r}. \tag{XXI.204}$$

The total Hamiltonian of the free radiation is obtained by summing over the Hamiltonians relative to the different eigenvibrations:

$$H_{\mathrm{rad}} = \sum_\tau h_\tau. \tag{XXI.205}$$

The factor $1/\sqrt{4\pi}$ in definitions (XXI.201) and (XXI.204) is due to the units chosen here for the electromagnetic field [1]).

H_{rad} is the radiation energy in terms of the canonical variables. This can be shown as follows. One has [cf. eq. (XXI.178)]:

$$\boldsymbol{A} = \sqrt{4\pi} \sum_\tau q_\tau \boldsymbol{T}^{(\tau)} \tag{XXI.206}$$

$$\mathscr{E}_\perp = -\sqrt{4\pi} \sum_\tau p_\tau \boldsymbol{T}^{(\tau)}. \tag{XXI.207}$$

From the reality and orthonormalization properties of $\boldsymbol{T}^{(\tau)}$, we obtain:

$$\int \mathscr{E}_\perp^2\, \mathrm{d}\boldsymbol{r} = 4\pi \sum_\tau p_\tau^2. \tag{XXI.208}$$

On the other hand, since $\boldsymbol{T}^{(\tau)}$ is a transverse vector obeying equation (XXI.200), we have, by integrating by parts:

$$\begin{aligned} \int \sum_i (\nabla A_i)^2\, \mathrm{d}\boldsymbol{r} &= -\int \sum_i A_i (\Delta A_i)\, \mathrm{d}\boldsymbol{r} \\ &= 4\pi \sum_\tau k_\tau^2 q_\tau^2. \end{aligned} \tag{XXI.209}$$

Substituting (XXI.208) and (XXI.209) in the expression for the energy U [eq. (XXI.188′)], one finds:

$$U = \tfrac{1}{2} \sum_\tau (p_\tau^2 + k_\tau^2 q_\tau^2) = H_{\mathrm{rad}}.$$

25. Hamiltonian for Radiation Coupled to a Set of Particles

In the general case, we can take as canonical variables on the one hand the q_τ and p_τ introduced in the preceding paragraph to describe the dynamical state of the radiation and on the other the variables $\boldsymbol{R}_n$ and $\boldsymbol{P}_n$ giving the position and momentum of each particle. $\boldsymbol{P}_n$

[1]) Cf. note, p. 1009.

is related to the mechanical momentum of the particle by formula (XXI.193).

From the results of § 22 [eq. (XXI.182) giving the motion of $\boldsymbol{A}$, eqs. (XXI.185–187) giving the Lorentz force], one easily obtains the equations of motion in these new variables. In particular, equation (XXI.202) for the free oscillator is replaced by the equation:

$$\ddot{q}_\tau + k_\tau^2 q_\tau = \sqrt{4\pi} \sum_n e_n \boldsymbol{V}_n \cdot \boldsymbol{T}^{(\tau)}(n) \qquad [\boldsymbol{V}_n \equiv \dot{\boldsymbol{R}}_n;\ \boldsymbol{T}^{(\tau)}(n) \equiv \boldsymbol{T}^{(\tau)}(\boldsymbol{R}_n)]. \tag{XXI.210}$$

These equations of motion can be put in the canonical Hamiltonian form with the Hamiltonian:

$$H \equiv H(q_\tau, p_\tau;\ \boldsymbol{R}_n, \boldsymbol{P}_n) = H_{\text{rad}} + H_{\text{coul}} + \sum_n H_n. \tag{XXI.211}$$

H_{rad} is the free radiation Hamiltonian defined by equation (XXI.205), H_{coul} the Coulomb potential given by equation (XXI.197) and H_n the mass energy of the nth particle expressed in terms of the canonical variables, i.e.

$$H_n \equiv \sqrt{m_n^2 + (\boldsymbol{P}_n - e_n \boldsymbol{A}(n))^2}, \tag{XXI.212}$$

where:

$$\boldsymbol{A}(n) = \sqrt{4\pi} \sum_\tau q_\tau \boldsymbol{T}^{(\tau)}(n). \tag{XXI.213}$$

Verification of this important property is not difficult and is left to the reader.

IV. QUANTUM THEORY OF RADIATION

26. Quantization of Free Radiation. Photons

> And God said, "Let there be light": and there was light. And God saw the light, that it was good ... And the evening and the morning were the first day.

The Quantum Theory of electromagnetic radiation can be deduced from the Classical Theory by applying the correspondence principle, as was done for the scalar field.

To this effect, we take as a starting point the Hamiltonian formalism developed in § 24 and 25.

Consider first the free radiation case. To the real dynamical variables q_τ and p_τ of the Classical Theory there correspond the observables obeying the commutation relations [cf. eq. (XXI.11)]:

$$[q_\tau, q_{\tau'}] = [p_\tau, p_{\tau'}] = 0$$
$$[q_\tau, p_{\tau'}] = i\delta_{\tau\tau'}.$$

The Hamiltonian governing the evolution of the system is the operator H_{rad} defined by equations (XXI.203) and (XXI.205). The radiation is therefore a superposition of quantized oscillators in accordance with Planck's law.

The entire discussion of § 3 can be applied here, in particular the part concerning the corpuscular interpretation. The radiation corpuscles are called *photons*. With each vibrational mode τ there is associated a certain type of photon characterized by the "wave function" $\boldsymbol{T}^{(\tau)}$ and whose energy is k_τ. The photons obey Bose–Einstein statistics, which is in excellent agreement with the experimental results concerning the thermodynamical properties of radiation (black body radiation etc.,). The remarks at the end of § 3 concerning the energy of the vacuum and field-fluctuations in the vacuum apply here without change.

The above quantization method is independent of the set of basis vector-fields $\boldsymbol{T}^{(\tau)}$ that is adopted; but these fields are supposed real. However, it can be reformulated so as to be valid also when the fields $\boldsymbol{T}^{(\tau)}$ are complex and the demonstration of § 5 justifying quantization with complex waves can be repeated in its entirety for the present case. The general formulas are easy to write down and will not be given here. We shall treat in some detail the plane-wave fields, and give a brief discussion of spherical-wave fields. Of all possible expansions, the plane-wave expansion is the most frequently used. The spherical-wave expansion or multipole expansion is useful in problems of emission or absorption.

27. Plane Waves. Radiation Momentum

The formulation using the plane-wave expansion is patterned on the one given in § 6 for the scalar field. We shall give the most important formulas, indicating in passing the points of difference with the scalar case.

Consider then a complete set of transverse plane-wave fields. In order to introduce discrete indices, we enclose these fields in a cube of side L, and impose the usual periodicity conditions at the edges of the cube. A field $\boldsymbol{U}^{(k\varpi)}$ belonging to this basis is defined by giving the wave-vector $\boldsymbol{k}$ and the quantum number ϖ labeling its polarization ($\varpi = 1, 2$):

$$\boldsymbol{U}^{(k\varpi)} = L^{-3/2}\, \boldsymbol{\epsilon}^{(\varpi)}\, e^{i\boldsymbol{k}\cdot\boldsymbol{r}}. \tag{XXI.214}$$

The expression for $\boldsymbol{T}^{(k\varpi)}$ introduced in § 21 [eq. (XXI.173*b*)] differs only in the normalization. The vectors $\boldsymbol{\epsilon}^{(1)}$ and $\boldsymbol{\epsilon}^{(2)}$ obey equations (XXI.174), which define them up to a unitary transformation.

With each plane wave $\boldsymbol{U}^{(k\varpi)}$, we associate an annihilation operator $a_{(k\varpi)}$ and a creation operator $a_{(k\varpi)}{}^{\dagger}$. The field $\boldsymbol{A}$ is represented by the expansion [cf. eqs. (XXI.46) and (XXI.206)]:

$$\boldsymbol{A} = \sum_{\boldsymbol{k}} \sum_{\varpi} (2\pi/k)^{\frac{1}{2}} \left(a_{(k\varpi)}\, \boldsymbol{U}^{(k\varpi)} + a_{(k\varpi)}{}^{\dagger}\, \boldsymbol{U}^{(k\varpi)*}\right) \tag{XXI.215}$$

and the field $\boldsymbol{\mathscr{E}}_{\perp}$ by [cf. eqs. (XXI.47) and (XXI.207)]:

$$\boldsymbol{\mathscr{E}}_{\perp} = \sum_{\boldsymbol{k}} \sum_{\varpi} i(2\pi k)^{\frac{1}{2}} \left(a_{(k\varpi)}\, \boldsymbol{U}^{(k\varpi)} - a_{(k\varpi)}{}^{\dagger}\, \boldsymbol{U}^{(k\varpi)*}\right). \tag{XXI.216}$$

The a and $a^{\dagger}$ obey the commutation relations characteristic of annihilation and creation operators:

$$[a_{(k\varpi)}, a_{(k'\varpi')}] = [a_{(k\varpi)}{}^{\dagger}, a_{(k'\varpi')}{}^{\dagger}] = 0 \tag{XXI.217*a*}$$

$$[a_{(k\varpi)}, a_{(k'\varpi')}{}^{\dagger}] = \delta_{\boldsymbol{k}\boldsymbol{k}'}\, \delta_{\varpi\varpi'}. \tag{XXI.217*b*}$$

The operator

$$N_{(k\varpi)} \equiv a_{(k\varpi)}{}^{\dagger}\, a_{(k\varpi)} \tag{XXI.218}$$

represents the number of photons in the state $\boldsymbol{U}^{(k\varpi)}$.

The radiation Hamiltonian, fixed in such a way as to have the vacuum energy equal to zero, is given by

$$H_{\text{rad}} = \sum_{\boldsymbol{k}} \sum_{\varpi} N_{(k\varpi)}\, k. \tag{XXI.219}$$

The momentum $\boldsymbol{X}$ is defined in terms of the operators $\boldsymbol{E}_{\perp}$ and $\boldsymbol{A}$ by expression (XXI.189′). Substituting for these the expansions (XXI.216) and (XXI.215) one finds (Problem XXI.6):

$$\boldsymbol{X} = \sum_{\boldsymbol{k}} \sum_{\varpi} N_{(k\varpi)}\, \boldsymbol{k}. \tag{XXI.220}$$

We can conclude from expressions (XXI.219) and (XXI.220) that

a photon in the state $U^{(k\varpi)}$ is a corpuscle of momentum $\boldsymbol{k}$ and energy k, that is, a corpuscle of mass null and momentum $\boldsymbol{k}$.

The number of states of the photon whose momentum is situated in the interval $(\boldsymbol{k}, \boldsymbol{k}+\delta\boldsymbol{k})$ is $2(L/2\pi)^3\delta\boldsymbol{k}$. The factor 2 here is due to there being two linearly independent polarization states for a given momentum ($\varpi = 1$ or 2). The density of levels $\varrho_L(k)$ of energy k, of momentum in the direction Ω and *of given polarization* ϖ (definition of § 6) is:

$$\varrho_L(k) = \frac{L^3 k^2}{(2\pi)^3} \tag{XXI.221}$$

[cf. eq. (XXI.49)].

28. Polarization

Photons having the same momentum $\boldsymbol{k}$ can differ from one another by their polarization. Classically, the polarization is the direction of vibration of the transverse electric field; it is a real or complex unit vector (defined to within a phase) perpendicular to the direction of propagation $\boldsymbol{k}$. This definition can be extended without difficulty to the present case. To each value of $\boldsymbol{k}$, there correspond two linearly independent polarizations, $\boldsymbol{\epsilon}^{(1)}$ and $\boldsymbol{\epsilon}^{(2)}$, and any other possible polarization of the photon is a (unitary) linear combination of these basis polarizations [1]).

As basis polarizations, one can take two orthogonal *rectilinear polarizations*, that is two real vectors $\boldsymbol{\epsilon}^{(1)}$ and $\boldsymbol{\epsilon}^{(2)}$. They can be chosen for example such that $\boldsymbol{\epsilon}^{(1)}$, $\boldsymbol{\epsilon}^{(2)}$ and $\boldsymbol{k}$ form a right-handed Cartesian triad of real vectors, which fixes the basis to within a rotation about the propagation axis:

$$\boldsymbol{\epsilon}^{(\varpi)*} = \boldsymbol{\epsilon}^{(\varpi)} \qquad \boldsymbol{\epsilon}^{(1)} \times \boldsymbol{\epsilon}^{(2)} = \frac{\boldsymbol{k}}{k}.$$

One can also use the two *circular polarizations* $\mathbf{e}^{(+)}$ and $\mathbf{e}^{(-)}$ as a basis:

$$\begin{aligned} \mathbf{e}^{(+)} &= -\frac{\sqrt{2}}{2}(\boldsymbol{\epsilon}^{(1)} + \mathrm{i}\boldsymbol{\epsilon}^{(2)}) \\ \mathbf{e}^{(-)} &= \frac{\sqrt{2}}{2}(\boldsymbol{\epsilon}^{(1)} - \mathrm{i}\boldsymbol{\epsilon}^{(2)}). \end{aligned} \tag{XXI.222}$$

---

[1]) Here we consider only the pure states, currently called *completely polarized*. The *partially polarized* states are statistical mixtures; they can be represented by a density operator of the (2-dimensional) space of the polarization states.

Note that:

$$\mathbf{e}^{(+)} = -\mathbf{e}^{(-)*}.$$

We shall use the symbol $\frown$ to indicate quantities defined with respect to this basis. We can thus define operators of annihilation $\hat{a}_{(\mathbf{k}\eta)}$ and of creation $\hat{a}_{(\mathbf{k}\eta)}{}^{\dagger}$ of circularly polarized photons: the corresponding plane wave is:

$$\hat{\mathbf{U}}^{(\mathbf{k}\eta)} \equiv L^{-3/2}\,\mathbf{e}^{(\eta)}\,e^{i\mathbf{k}\cdot\mathbf{r}} \qquad (\eta = + \text{ or } -).$$

The interest of circular polarizations is essentially due to the following property:

A circularly polarized photon has a well-defined angular momentum along its direction of propagation; the component of the angular momentum is equal to $+1$ *or* -1 *according as the polarization is right-handed* $(\eta = +)$ *or left-handed* $(\eta = -)$.

To show this, suppose that the radiation is in one of the states $|\mathbf{k}\,\pm\rangle$ defined by

$$|\mathbf{k}\,\pm\rangle \equiv \hat{a}_{(\mathbf{k}\pm)}{}^{\dagger}\,|0\rangle.$$

This state represents a photon of momentum $\mathbf{k}$ and of right $(+)$ or of left $(-)$ circular polarization. It can be shown (Problem (XXI.8) that:

$$(\mathbf{k}\cdot\mathbf{G}^{(0)})\,|\mathbf{k}\,\pm\rangle = 0 \qquad \text{(XXI.223)}$$

$$(\mathbf{k}\cdot\mathbf{G}^{(s)})\,|\mathbf{k}\,\pm\rangle = (\pm\,k)\,|\mathbf{k}\,\pm\rangle, \qquad \text{(XXI.224)}$$

where the vector operators $\mathbf{G}^{(0)}$ and $\mathbf{G}^{(s)}$ are defined respectively by equations (XXI.191*a*) and (XXI.191*b*). According to equation (XXI.190′), the total angular momentum $\mathbf{G}$ is the sum of these two vector operators; denote its component along $\mathbf{k}$ by G_k:

$$G_k \equiv \frac{(\mathbf{k}\cdot\mathbf{G})}{k}.$$

From the two preceding equations, one obtains the desired result:

$$G_k\,|\mathbf{k}\,\pm\rangle = (\pm\,1)\,|\mathbf{k}\,\pm\rangle.$$

Q.E.D.

This property is characteristic of a particle of spin 1. More generally, if a particle of spin 1 has a well-defined momentum $\mathbf{k}$ the component of its orbital angular momentum along $\mathbf{k}$ vanishes, and the component of its spin along $\mathbf{k}$ can take one of the three values $+1$, 0, -1. In

the particular case of a photon, however, the longitudinal plane waves, which correspond to the value 0 for the spin component, must be eliminated.

29. Multipole Expansion. Photons of Determined Angular Momentum and Parity

Rather than expanding in plane waves, we can expand in a series of spherical waves, each of which corresponds to a well-defined value of the angular momentum and parity. We shall limit ourselves here to giving the main results without the demonstrations. The latter are not difficult.

A vector field $\boldsymbol{B}(\boldsymbol{r})$ can be considered as a three-component wave function representing the dynamical state of a particle of spin 1. There exists a certain arbitrariness in the choice of the components of $\boldsymbol{B}$ which constitute the wave function. For example, one could take B_x, B_y and B_z but, for the needs of the present discussion, it is preferable to take the standard components:

$$B_+ = -\frac{\sqrt{2}}{2}(B_x - \mathrm{i}B_y)$$

$$B_0 = B_z$$

$$B_- = \frac{\sqrt{2}}{2}(B_x + \mathrm{i}B_y).$$

The wave function is a function of $\boldsymbol{r}$ and of the index μ labeling the components in question ($\mu = +$, 0 or $-$):

$$\Psi(\boldsymbol{r}, \mu) \equiv B_\mu(\boldsymbol{r}).$$

On this point there is a perfect parallel between particles of spin 1 and particles of spin $\frac{1}{2}$ (cf. § XIII.20–21).

The observables of this spin 1 particle are functions of its position $\boldsymbol{r}$, its momentum $\boldsymbol{p}$ and its spin $\boldsymbol{s}$. In the representation defined here, $\boldsymbol{p} = -\mathrm{i}\boldsymbol{\nabla}$ and the matrices representing the components s_+, s_z, s_- of the spin are given by equations (XIII.28) taken with $j = j' = 1$. It will easily be verified that $\boldsymbol{s}^2 = 2$, a value characteristic of an angular momentum equal to 1.

We can define the orbital angular momentum $\boldsymbol{l}$ of the particle:

$$\boldsymbol{l} = \boldsymbol{r} \times \boldsymbol{p}$$

and its total angular momentum:

$$\boldsymbol{j} \equiv \boldsymbol{l} + \boldsymbol{s}.$$

The components of $\boldsymbol{j}$ are associated with the infinitesimal rotations of the vector field, in accordance with the usual definition of the angular momentum operator.

We similarly define the parity P; it is an operator whose action on the field $\boldsymbol{B}(\boldsymbol{r})$ gives its transform under reflection in the origin [1]):

$$PB_\mu(\boldsymbol{r}) = -B_\mu(-\boldsymbol{r}).$$

We now wish to construct a complete set of vector fields of well-defined angular momentum (jm) satisfying the wave equation:

$$(\triangle + k^2)\boldsymbol{B}(\boldsymbol{r}) = 0.$$

This amounts to finding a common set of eigenfunctions of $\boldsymbol{p}^2$, $\boldsymbol{j}^2$ and j_z. The corresponding eigenvalues are k^2, $j(j+1)$ and m respectively; k can take all values from 0 to $+\infty$ (the spectrum of k can be made discrete by enclosing the field in a sphere of finite radius R), j all integral values from 0 to $+\infty$ and m all integral values from $-j$ to $+j$.

From the properties of addition of angular momenta, one can deduce that $\boldsymbol{p}^2$, $\boldsymbol{l}^2$, $\boldsymbol{j}^2$ and j_z form a complete set of commuting observables and that to each set (kjm) there correspond:

(*i*) if $j \neq 0$, *three* linearly independent states, for which we have respectively:

$$l = j+1, \qquad j, \qquad j-1;$$

(*ii*) if $j = 0$, *one and only one* state for which $l = 1$.

These states have a well-defined parity, namely:

$$P = (-)^{l+1}.$$

Rather than classify the states (kjm) according to the possible values of l, as above, we can classify them according to their transverse or longitudinal character and their parity. The transverse or lon-

[1]) We suppose that $\boldsymbol{B}(\boldsymbol{r})$ is a field of polar vectors, whence the presence of the minus sign in the spatial reflection. To within the sign of the parity, all of the properties given below apply equally well to axial-vector fields or pseudovector fields (cf. note p. 904).

gitudinal character of a field is related to the properties of the operator $(\boldsymbol{s}\cdot\boldsymbol{p})$; in the representation adopted here we have [eq. (XIII.93)]

$$(\boldsymbol{s}\cdot\boldsymbol{p}) \equiv \text{curl}$$
$$(\boldsymbol{s}\cdot\boldsymbol{p})^2 - \boldsymbol{p}^2 \equiv \text{grad div} \qquad (\equiv \text{curl curl} + \triangle).$$

For a longitudinal field $(\boldsymbol{s}\cdot\boldsymbol{p})^2 = 0$; for a transverse field, $(\boldsymbol{s}\cdot\boldsymbol{p})^2 - \boldsymbol{p}^2 = 0$. It can be shown that $\boldsymbol{p}^2$, $(\boldsymbol{s}\cdot\boldsymbol{p})^2$, P, $\boldsymbol{j}^2$ and j_z form a complete set of commuting observables and that to each set (kjm) there correspond:

(*i*) if $j \neq 0$, *three* linearly independent states, for which we have respectively:

$$\begin{array}{lccc} (\boldsymbol{s}\cdot\boldsymbol{p})^2 = & 0 & k^2 & k^2 \\ P = & (-)^j & (-)^j & (-)^{j+1} \end{array}$$

(*ii*) if $j = 0$, *one and only one* state, even and longitudinal:

$$(\boldsymbol{s}\cdot\boldsymbol{p})^2 = 0 \qquad P = +1$$

(this is the p state already mentioned above).

To construct the corresponding basis functions, we return to the functions u_{kjm} introduced in § 7 in connection with the scalar field; the constant in definition (XXI.60) is a certain function of k which we shall suppose chosen in such a way that the u_{kjm} form a complete orthonormal set of scalar functions. To each set (kjm) there corresponds one and only one longitudinal state; the corresponding basis function is:

$$\boldsymbol{\Lambda}_{kjm} \equiv \text{grad}\,\frac{u_{kjm}}{k}$$

(this is a vector-field of parity $(-)^j$). To each set (kjm) with the exception of the cases: $j = m = 0$, there correspond two transverse states of opposite parity; the corresponding basis functions are respectively [1]:

$$\boldsymbol{\Theta}^{(-)}_{kjm} \equiv \frac{\boldsymbol{l}\, u_{kjm}}{\sqrt{j(j+1)}} \qquad (P = (-)^{j+1}) \qquad \text{(XXI.225}a\text{)}$$

$$\boldsymbol{\Theta}^{(+)}_{kjm} \equiv \frac{\nabla \times \boldsymbol{l}\, u_{kjm}}{k\sqrt{j(j+1)}} \qquad (P = (-)^j). \qquad \text{(XXI.225}b\text{)}$$

---

[1]) The parities indicated here are of opposite sign to those given by most authors (in particular Blatt and Weisskopf, *loc. cit.*) since the vector fields introduced here are polar-vector fields while these authors consider axial-vector fields.

Note in passing that:

$$\mathbf{\Theta}^{(\pm)}_{kjm} = \operatorname{curl} \frac{\mathbf{\Theta}^{(\mp)}_{kjm}}{k}.$$

The set of fields $\mathbf{\Theta}^{(\varpi)}_{kjm}$ (k varying from 0 to $+\infty$; $j = 1, 2, \ldots, \infty$; $m = -j, \ldots, +j$; $\varpi = +$ or $-$) form a complete orthonormal set of transverse fields in terms of which we can expand the operators $\boldsymbol{A}$ and $\mathscr{E}_\perp$:

$$\boldsymbol{A} = \sum_{kjm\varpi} (2\pi/k)^{\frac{1}{2}} [a^{(\varpi)}_{kjm} \mathbf{\Theta}^{(\varpi)}_{kjm} + a^{(\varpi)\dagger}_{kjm} \mathbf{\Theta}^{(\varpi)*}_{kjm}] \qquad \text{(XXI.226)}$$

$$\mathscr{E}_\perp = \sum_{kjm\varpi} (2\pi k)^{\frac{1}{2}} \mathrm{i}[a^{(\varpi)}_{kjm} \mathbf{\Theta}^{(\varpi)}_{kjm} - a^{(\varpi)\dagger}_{kjm} \mathbf{\Theta}^{(\varpi)*}_{kjm}]. \qquad \text{(XXI.227)}$$

According to the usual terminology, the terms $\mathbf{\Theta}^{(+)}_{kjm}$ in these expansions represent the *electric* 2^j*-pole* contribution; the terms $\mathbf{\Theta}^{(-)}_{kjm}$ the *magnetic* 2^j*-pole* contribution. All of the work of § 27 in connection with the plane-wave expansion can be repeated here for the multipole expansion. The operators $a^{(\varpi)}_{kjm}$ and $a^{(\varpi)\dagger}_{kjm}$ can be respectively interpreted as annihilation and creation operators of a photon of energy k, angular momentum (jm) and parity $(-)^j\varpi$. The spectra of k, j, m and ϖ were mentioned above; it is to be observed that there is no photon with zero angular momentum.

30. Coupling with an Atomic System

> And God said, "Let there be lights in the firmament of the heaven to divide the day from the night..." and God saw that it was good. And the evening and the morning were the fourth day.

We now consider radiation coupled to a system of particles. The treatment is analogous to that of the scalar field and we shall limit ourselves here to the essential points.

The Hamiltonian of the system is, to within a few obvious modifications, the same as that of the classical theory [cf. eq. (XXI.211)]. The modifications in question depend on the dynamical properties and the spin of the particles considered. We limit ourselves here to the case of an atomic system of Z electrons and a nucleus of charge $-Ze$ (e is the charge of the electron).

We first suppose the nucleus to be absent, and form the Hamil-

tonian of the system of Z electrons + radiation. The total Hamiltonian is then given by equation (XXI.211), where H_{rad} represents the free radiation Hamiltonian, H_{coul} the Coulomb interaction of the electrons, and H_n the Dirac Hamiltonian of the nth electron in the presence of the field $\boldsymbol{A}$, i.e.:

$$H_n = \boldsymbol{\alpha}_n \cdot [\boldsymbol{P}_n - e\boldsymbol{A}(n)] + \beta_n m. \tag{XXI.228}$$

$\boldsymbol{\alpha}_n$ and β_n are the Dirac matrices relative to the nth electron.

In the non-relativistic approximation, H_n gives, after subtraction of the rest mass:

$$(H_n)_{NR} = \frac{(\boldsymbol{P}_n - e\boldsymbol{A}(n))^2}{2m} - \frac{e}{2m}(\boldsymbol{\sigma}_n \cdot \mathscr{H}(n)). \tag{XXI.229}$$

We note that the electron-radiation coupling term in H_n [eq. (XXI.228)] can be written:

$$H_n' = -e \int (\boldsymbol{j}_n(\boldsymbol{r}) \cdot \boldsymbol{A}(\boldsymbol{r}))\, \mathrm{d}\boldsymbol{r} \tag{XXI.230}$$

where $\boldsymbol{j}_n(\boldsymbol{r})$ represents the Dirac current density of the electron at point $\boldsymbol{r}$, that is [cf. eq. (XX.143)]:

$$\boldsymbol{j}_n(\boldsymbol{r}) \equiv \boldsymbol{\alpha}_n \delta(\boldsymbol{R}_n - \boldsymbol{r}). \tag{XXI.231}$$

$e\boldsymbol{j}_n(\boldsymbol{r})$ is the electric current density of the nth electron. By substituting the non-relativistic approximation for $\boldsymbol{j}_n$ [eq. (XX.189–190)] in (XXI.230) one obtains the coupling $(H_n')_{NR}$ in the non-relativistic approximation:

$$(H_n')_{NR} \equiv H_n^{(\mathrm{I})} + H_n^{(\mathrm{II})} + H_n'' \tag{XXI.232}$$

$$H_n^{(\mathrm{I})} = -\frac{e}{2m}(\boldsymbol{P}_n \cdot \boldsymbol{A}(n) + \boldsymbol{A}(n) \cdot \boldsymbol{P}_n) \tag{XXI.233a}$$

$$H_n^{(\mathrm{II})} = -\mathrm{i}\frac{e}{2m}[\boldsymbol{\sigma}_n \cdot ([\boldsymbol{P}_n \times \boldsymbol{A}(n)] + [\boldsymbol{A}(n) \times \boldsymbol{P}_n])] = -\frac{e}{2m}[\boldsymbol{\sigma}_n \cdot \mathscr{H}(n)] \tag{XXI.233b}$$

$$H_n'' = \frac{e^2}{2m}\boldsymbol{A}^2(n). \tag{XXI.233c}$$

This is just the electron-radiation coupling figuring in expression (XXI.229).

To take into account the presence of the atomic nucleus, we must add to the Hamiltonian the Coulomb interaction of the nucleus with

each of the Z electrons, and also the Hamiltonian of the nucleus in the presence of the radiation, which includes a kinetic energy term and a term for the coupling with the radiation [1]). In the (essentially non-relativistic) approximation where the nucleus is treated as infinitely heavy, these last two terms are negligible; the nucleus can then be thought of as fixed at the origin and we need only to add the Z Coulomb terms $-Ze^2/R_n$. The atom is then treated simply as a set of Z electrons, each moving in the external potential $-Ze^2/R$ and coupled to the electromagnetic field. The corresponding Hamiltonian can be put in the form

$$H = H_{\text{rad}} + H_{\text{at}} + H'. \tag{XXI.234}$$

H_{rad} is still the free-radiation Hamiltonian. H_{at} is the Hamiltonian of the atom itself, that is, of the electrons in the Coulomb field of the nucleus and in Coulomb interaction with one another; it is the sum of the Coulomb interaction term H_{coul} and the kinetic energies of the Z electrons. H' is the term coupling the electrons of the atom with the radiation; it is the sum of the Z coupling terms H_n' defined by equation (XXI.230). In what follows we shall treat the motion of the electrons in the non-relativistic approximation; we therefore adopt the approximate form for H_n' given by equations (XXI.232) and (XXI.233), in other words:

$$H' = H^{(I)} + H^{(II)} + H'', \tag{XXI.235}$$

where

$$H^{(I)} = \sum_n H_n^{(I)}, \quad H^{(II)} = \sum_n H_n^{(II)}, \quad H'' = \sum_n H_n''. \tag{XXI.235'}$$

Two approximation methods are generally used to study such a system.

The first of these consists in *treating the electromagnetic radiation classically*. It is justified in all cases when the energy transfer between the atom and the radiation field is so large compared with the energy of the emitted or absorbed photons that the discontinuous character of this transfer can be neglected. It is therefore a good approximation for strong intensities and low frequencies (many photons present).

[1]) Here we suppose that the nucleus can be treated as a point charge. This approximation is legitimate provided that the frequencies which come into play are small compared with the excitation energies of the nucleus.

It is particularly suited to the study of an atom in a static electromagnetic field or in a radioelectric field. The electromagnetic field is then considered as a given, possibly time-dependent, external field, and the problem reduces to the study of an atom in an external field.

The second method consists in *treating the coupling term H' as a small perturbation* as has already been done in the case of the scalar field.

The discussion of § 10 can be repeated here with very little change. The unperturbed Hamiltonian is:

$$H_0 \equiv H_{\text{rad}} + H_{\text{at}}.$$

In order to apply the perturbation method we must use a representation in which H_0 is diagonal. Taking the plane wave represention defined in § 27 for the radiation field, we define the basis vectors of H_0 as the simultaneous eigenvectors of H_{at} and of the operators $N_{(\boldsymbol{k}\varpi)}$ representing the number of photons of given momentum and polarization.

The matrix elements of H' in this representation are easily obtained by using its expression in terms of the $a_{(\boldsymbol{k}\varpi)}$ and the $a_{(\boldsymbol{k}\varpi)}^{\dagger}$. Substituting expansion (XXI.215) for $\boldsymbol{A}$ in expression (XXI.233) one finds

$$H_n^{(\mathrm{I})} = -\frac{e}{m}\sum_s \left(\frac{2\pi}{kL^3}\right)^{\frac{1}{2}} (a_s(\boldsymbol{P}\cdot\boldsymbol{\epsilon})\, e^{i\boldsymbol{k}\cdot\boldsymbol{R}} + \text{Herm. conj.}) \qquad \text{(XXI.236}a\text{)}$$

$$H_n^{(\mathrm{II})} = -\frac{e}{2m}\sum_s \left(\frac{2\pi}{kL^3}\right)^{\frac{1}{2}} (i\, a_s(\boldsymbol{\sigma}\cdot\boldsymbol{k}\times\boldsymbol{\epsilon})\, e^{i\boldsymbol{k}\cdot\boldsymbol{R}} + \text{Herm. conj.}) \qquad \text{(XXI.236}b\text{)}$$

$$\begin{aligned} H_n'' = \frac{e^2}{2m}\sum_s\sum_{s'} \frac{\pi}{\sqrt{kk'L^3}} &([a_s a_{s'}(\boldsymbol{\epsilon}\cdot\boldsymbol{\epsilon}')\, e^{i(\boldsymbol{k}+\boldsymbol{k}')\cdot\boldsymbol{R}} + \text{Herm. conj.}] \\ &+ [a_s^{\dagger} a_{s'}(\boldsymbol{\epsilon}^*\cdot\boldsymbol{\epsilon}')\, e^{i(\boldsymbol{k}'-\boldsymbol{k})\cdot\boldsymbol{R}} + \text{Herm. conj.}]). \end{aligned} \qquad \text{(XXI.237)}$$

[To simplify the writing we have put $s \equiv (\boldsymbol{k}\,\varpi)$ and $s' \equiv (\boldsymbol{k}'\varpi')$, and have represented the vectors $\boldsymbol{R}_n$, $\boldsymbol{P}_n$, $\boldsymbol{\sigma}_n$, $\boldsymbol{\epsilon}^{(\varpi)}$ and $\boldsymbol{\epsilon}^{(\varpi')}$ by $\boldsymbol{R}$, $\boldsymbol{P}$, $\boldsymbol{\sigma}$, $\boldsymbol{\epsilon}$ and $\boldsymbol{\epsilon}'$ respectively.] The desired expression for H' is obtained by substituting these expansions in (XXI.235) and (XXI.235′).

The presence of the coupling term H' has the effect of displacing all of the atomic levels, and of rendering them all unstable with the exception of the ground state (recall that the photon mass is null). This effect can be treated entirely within the framework of the

perturbation method. One obviously encounters the same difficulties as in the scalar case, and this in spite of the fact that the coupling constant e is small ($e^2 = 1/137$). In particular, there is an "ultraviolet catastrophe" related to the fact that H' is a local interaction [cf. eq. (XXI.230) and (XXI.231)]. This catastrophe can be avoided by introducing a suitable cut-off. For the same reasons as given before, one usually places this cut-off at a frequency of the order of m. The qualifications that must be made with regard to this cut-off procedure and the difficulties relating to the necessity of renormalizing the mass and the charge of the particles have already been discussed. All of the considerations of section II concerning level shifts and properties of metastable states (lifetime, line width, formation and decay modes) can be extended without any essential change to the radiation case [1]). Similarly, the few collision problems treated in section II can be treated in the radiation case by the same methods. In concluding this chapter we shall describe two simple applications of the Quantum Theory of Radiation; the calculation of probabilities for the emission of radiation by an atom and the calculation of the cross section for Compton scattering in the low frequency limit.

31. Emission of a Photon by an Atom. Dipole Emission

Consider an atom in one of its excited states $|\lambda\rangle$ of energy E_λ. Such an atom is capable of making transitions to any one of the states $|\alpha\rangle$, $|\beta\rangle$, ... whose energy E_α, E_β, ... is lower than E_λ, with emission of a photon whose energy is equal to the difference between the energies of the initial and final states.

Let us therefore consider the transition:

$$\lambda \to \mu \mathbf{k}\boldsymbol{\epsilon}.$$

μ is the final state of the atom, $\mathbf{k}$ and $\boldsymbol{\epsilon}$ the momentum and polarization of the photon emitted, and

$$k = E_\lambda - E_\mu. \tag{XXI.238}$$

[1]) The radiation case is complicated by the fact that photons are corpuscles of zero mass and spin 1. The zero mass results, in addition to the "ultraviolet catastrophe", in an "infrared catastrophe", that is, in the existence of low-frequency divergences in certain terms in the perturbation expansion. The "infrared catastrophe" is a difficulty of the method itself and can be avoided by improving the method. The "ultraviolet catastrophe", on the other hand, is a difficulty in the theory proper, for which at the present time there is no fully satisfactory solution.

The transition probability per unit time $w_{\lambda\to\mu\mathbf{k}\epsilon}$ is given to the first order in the perturbation calculation by the formula (cf. § 12):

$$w_{\lambda\to\mu\mathbf{k}\epsilon} = 2\pi|\langle\lambda|H'|\mu\mathbf{k}\epsilon\rangle|^2\left(\frac{L^3k^2}{(2\pi)^3}\right). \tag{XXI.239}$$

It will be seen from equation (XXI.235) that, H' is composed of three terms. The contribution from the third term, H'', to the matrix element in the above formula vanishes [1]). The contributions from the first two terms are easily obtained with the aid of expansions (XXI.236), giving

$$\langle\lambda|H^{(\mathrm{I})}|\mu\mathbf{k}\epsilon\rangle = -\frac{e}{m}\left(\frac{2\pi}{kL^3}\right)^{\frac{1}{2}}F^{(\mathrm{I})} \tag{XXI.240a}$$

$$\langle\lambda|H^{(\mathrm{II})}|\mu\mathbf{k}\epsilon\rangle = \frac{ie}{m}\left(\frac{2\pi}{kL^3}\right)^{\frac{1}{2}}F^{(\mathrm{II})}, \tag{XXI.240b}$$

where

$$F^{(\mathrm{I})} \equiv \langle\lambda|\sum_n(\boldsymbol{\epsilon}\cdot\boldsymbol{P}_n)\,e^{i\mathbf{k}\cdot\mathbf{R}_n}|\mu\rangle \tag{XXI.241a}$$

$$F^{(\mathrm{II})} \equiv \langle\lambda|\tfrac{1}{2}\sum_n(\boldsymbol{\epsilon}\cdot\mathbf{k}\times\boldsymbol{\sigma}_n)\,e^{i\mathbf{k}\cdot\mathbf{R}_n}|\mu\rangle. \tag{XXI.241b}$$

Substituting these expressions in (XXI.239), we finally obtain the formula:

$$w_{\lambda\to\mu\mathbf{k}\epsilon} = \frac{e^2k}{2\pi m^2}|F^{(\mathrm{I})} - iF^{(\mathrm{II})}|^2, \tag{XXI.242}$$

where $F^{(I)}$ and $F^{(II)}$ denote the matrix elements defined by relations (XXI.241).

In ordinary atomic transitions, the order of magnitude of the transition probability depends critically on the angular momenta j_λ and j_μ and the parities Π_λ and Π_μ of states $|\lambda\rangle$ and $|\mu\rangle$ respectively.

If $j_\lambda = j_\mu = 0$ the transition probability vanishes since the matrix elements $F^{(I)}$ and $F^{(II)}$ are identically null (Problem XXI.10). *This*

---

[1]) The action of H'' either conserves the number of photons present, or increases or diminishes this number by 2 units. Note that H'' is of the second order in e and can therefore be omitted in a perturbation calculation to the first order in e. In calculating the level shift, which is of the second order (cf. § 11), the contribution from H'' ceases to be negligible. It is also important in photon scattering (cf. § 32).

is the well known $0 \to 0$ *selection rule.* It is related to the fact that there is no photon of zero angular momentum (cf. § 29).

If j_λ or j_μ is different to zero, the transition probability can be evaluated by making the long wavelength approximation. Let R be a length of the order of atomic dimensions; in ordinary atomic transitions,

$$k \lesssim \frac{e^2}{R} \simeq me^4,$$

so that

$$kR \lesssim e^2 \ll 1.$$

Expressions (XXI.241) can therefore be evaluated by replacing the exponentials by the first term in their Taylor expansion having a non-vanishing contribution. The order of this term is mainly determined by the selection rules for angular momentum and parity, a subject that will not be gone into here [1]). All things being equal, the transition probability is greatest in the case of the so-called *electric dipole* transitions, in which case

$$\Delta j \equiv |j_\lambda - j_\mu| \leqslant 1 \qquad \Delta\Pi \equiv \Pi_\lambda \Pi_\mu = -1 \qquad \text{(XXI.243)}$$
$$(j_\lambda + j_\mu = 0 \text{ excepted}).$$

Next come the magnetic dipole ($\Delta j \leqslant 1, \Delta\Pi = 1, j_\lambda + j_\mu \neq 0$) and the electric quadrupole ($\Delta j \leqslant 2, \Delta\Pi = +1, j_\lambda + j_\mu \neq 0$ or 1) transitions which are $(kR)^{-2}$ times slower than the electric dipole transitions.

If the conditions for electric dipole emission, (XXI.243), are fulfilled, we may write:

$$F^{(\mathrm{I})} \simeq \langle\lambda| \sum_n (\boldsymbol{\epsilon} \cdot \boldsymbol{P}_n) |\mu\rangle$$

$$F^{(\mathrm{II})} \simeq \tfrac{1}{2}\, \mathrm{i} \langle\lambda| \sum_n (\boldsymbol{\epsilon} \cdot \boldsymbol{k} \times \boldsymbol{\sigma}_n)\, (\boldsymbol{k} \cdot \boldsymbol{R}_n) |\mu\rangle.$$

---

[1]) The results appear in a more natural way if we use a multipole expansion rather than the expansion in plane waves; due to the selection rules most of the contributions vanish, the others can be evaluated by replacing the multipole fields $\boldsymbol{\Theta}$ by the first term in their expansion in powers of kr. Note [cf. eq. (XXI.225)] that when $kr \ll 1$:

$$\boldsymbol{\Theta}^{(+)}_{kjm} \propto (kr)^{j-1} \qquad \boldsymbol{\Theta}^{(-)}_{kjm} \propto (kr)^{j}.$$

A complete discussion of this method, with reference to radiative transitions in nuclei, can be found in J. Blatt and V. Weisskopf, Ch. XII, *loc. cit.*, note 2, p. 576.

We therefore have roughly: $F^{(I)} \approx 1/R$, $F^{(II)} \approx k^2 R$ i.e.:

$$|F^{(II)}/F^{(I)}| \approx (kR)^2 \ll 1.$$

The contribution of $F^{(II)}$ is negligible compared with that of $F^{(I)}$. In order to calculate $F^{(I)}$, it is convenient to introduce the electric dipole moment of the electrons:

$$\boldsymbol{D} \equiv e \sum_n \boldsymbol{R}_n. \tag{XXI.244}$$

[This is the vector operator denoted by $\mathbf{Q}^{(1)}$ in § XIII.33 and XVII.3; cf. eq. (XVII.35).]

$\boldsymbol{D}$ satisfies the relation:

$$[\boldsymbol{D}, H_{\text{at}}] = ie \sum_n \frac{\boldsymbol{P}_n}{m}.$$

Therefore:

$$\frac{ie}{m} \langle\lambda| \sum_n \boldsymbol{P}_n |\mu\rangle = \langle\lambda| (\boldsymbol{D} H_{\text{at}} - H_{\text{at}} \boldsymbol{D}) |\mu\rangle$$

$$= (E_\mu - E_\lambda) \langle\lambda|\boldsymbol{D}|\mu\rangle = -k\langle\lambda|\boldsymbol{D}|\mu\rangle,$$

whence

$$F^{(\text{I})} = \frac{imk}{e} \langle\lambda| (\boldsymbol{D}\cdot\boldsymbol{\epsilon}) |\mu\rangle$$

and

$$w_{\lambda\to\mu\mathbf{k}\epsilon} = \frac{k^3}{2\pi} |\langle\lambda| (\boldsymbol{D}\cdot\boldsymbol{\epsilon}) |\mu\rangle|^2. \tag{XXI.245}$$

It is convenient at this point to specify the "magnetic quantum numbers" m_λ and m_μ of the initial and final states, and to introduce the reduced matrix element $\langle\lambda\|\boldsymbol{D}\|\mu\rangle$ in accordance with definition (XIII.125). If D_q denotes the qth standard component of the vector operator $\boldsymbol{D}$, the Wigner–Eckart theorem gives:

$$\langle j_\lambda m_\lambda | D_q | \mu j_\mu m_\mu \rangle = (2j_\lambda + 1)^{-\frac{1}{2}} \langle j_\mu 1 m_\mu q | j_\lambda m_\lambda \rangle \langle\lambda\|\boldsymbol{D}\|\mu\rangle.$$

Since $\boldsymbol{\epsilon}\cdot\boldsymbol{D} = \sum_q \epsilon_q D_q$, where ϵ_q ($q = +, 0, -$) denotes the standard components of the polarization vector [definition (XIII.94)], we finally obtain:

$$w_{\lambda\to\mu\mathbf{k}\epsilon} = \left[\frac{k^3 |\langle\lambda\|\boldsymbol{D}\|\mu\rangle|^2}{2\pi}\right] \times \left[\frac{|\sum_q \epsilon_q \langle j_\mu 1 m_\mu q | j_\lambda m_\lambda\rangle|^2}{(2j_\lambda + 1)}\right]. \tag{XXI.246}$$

If the atom is unpolarized in its initial state and if the polarization

in the final state is not observed, we must sum over m_μ and average over m_λ, i.e.:

$$\langle w_{\lambda\to\mu\mathbf{k}\epsilon}\rangle \equiv \frac{1}{2j_\lambda+1}\sum_{m_\lambda m_\mu} w_{\lambda\to\mu\mathbf{k}\epsilon}.$$

This is easily done using the symmetry and orthogonality properties of the C.–G. coefficients [cf. eq. (C.13c) and (C.14a)], and gives:

$$\langle w_{\lambda\to\mu\mathbf{k}\epsilon}\rangle = \frac{k^3|\langle\lambda\|\boldsymbol{D}\|\mu\rangle|^2}{6\pi(2j_\lambda+1)}.$$

As might have been expected, the expression obtained depends neither on the direction of emission nor on the polarization of the photon emitted. The total transition probability $\lambda\to\mu$ per unit time is obtained by summing over the two polarization states and integrating over the angles of emission, i.e. by multiplying the expression by 8π:

$$w^{(\mathrm{tot})}_{\lambda\to\mu} = \frac{4k^3}{3(2j_\lambda+1)}|\langle\lambda\|\boldsymbol{D}\|\mu\rangle|^2. \qquad \text{(XXI.247)}$$

32. Low Energy Compton Scattering. The Thomson Formula

As an example of a collision problem, rather than consider those treated in section II – and which can be solved by the same methods – we shall consider the simpler problem of Compton scattering (the scattering of photons by free electrons) [1]).

The state of the system before collision consists of an electron of given momentum $\boldsymbol{P}_\mathrm{i}$ and a photon whose momentum and polarization are denoted by $\boldsymbol{k}_\mathrm{i}$ and $\boldsymbol{\epsilon}_\mathrm{i}$ respectively. It is represented by a certain ket vector $|i\rangle$

$$|i\rangle \equiv |\boldsymbol{P}_\mathrm{i}\boldsymbol{k}_\mathrm{i}\boldsymbol{\epsilon}_\mathrm{i}\rangle.$$

We wish to calculate the cross section for scattering from this initial state to a certain final state in which the momentum of the electron, the momentum of the photon and its polarization are $\boldsymbol{P}_\mathrm{f}$, $\boldsymbol{k}_\mathrm{f}$ and $\boldsymbol{\epsilon}_\mathrm{f}$ respectively and which we denote by $|f\rangle$:

$$|f\rangle \equiv |\boldsymbol{P}_\mathrm{f}\boldsymbol{k}_\mathrm{f}\boldsymbol{\epsilon}_\mathrm{f}\rangle.$$

Energy and momentum are conserved in the collision.

---

[1]) In the case of a scalar field, the scattering cross section for a field corpuscle by a free particle vanishes.

In order to calculate the cross section for all energies, the electron must be treated as a Dirac particle coupled to the radiation field by the interaction $-e\boldsymbol{\alpha}\cdot\boldsymbol{A}(\boldsymbol{R})$. In this paragraph, we shall limit ourselves to treating the problem in the low-frequency limit [1]).

The electron is supposed initially at rest ($\boldsymbol{P}_i = 0$), and the incident photon sufficiently soft for the energy transferred to the electron to be small compared with its rest mass:

$$k_f \ll m$$

(whence: $k_f \simeq k_i$, $P_f \approx k_f$). One may then treat the electron in the non-relativistic approximation, and take as the Hamiltonian for the system:

$$H = H_0 + H',$$

where

$$H_0 = H_{\text{rad}} + \frac{\boldsymbol{P}^2}{2m}.$$

In accordance with equation (XXI.235) H' is the sum of three terms $H^{(\text{I})}$, $H^{(\text{II})}$ and H'', defined by equations (XXI.236*a*), (XXI.236*b*) and (XXI.237) respectively.

To calculate the collision cross section we apply the formulas of collision theory. It is clear that the reservations made regarding this point in § 14 must also be made here. We shall not return to this question but will suppose, as in § 14, that the formulas given by collision theory are correct providing the electron mass m appearing in them is the mass that is experimentally measured.

However, the straightforward application of the collision theory formalism encounters another difficulty, namely the impossibility of defining and separating out the center-of-mass variables. This difficulty is only an apparent one. In reality, the elimination of the motion of the center of mass plays only an accessory role in non-relativistic collision theory; the resulting simplifications are due to the fact that the total momentum of the system is conserved in the collision. It is possible to reformulate Collision Theory without making this elimination and the conservation of momentum then appears explicitly.

---

[1]) The general case can be treated by the same method (cf. Heitler, *loc. cit.*) and leads to the Klein–Nishina formula. The calculations are long but not difficult. In fact, they can be considerably shortened by using the covariant formulation of the Quantum Theory of Radiation (cf. Jauch and Rōhrlich, *loc. cit.*).

This new formulation will be used here. The only modification concerns the definition of the transition amplitude $T_{i\to f}$. We shall adopt for T the same formal definition:

$$T \equiv H' + H' \frac{1}{E - H + i\varepsilon} H'.$$

Since T is invariant under translation, the matrix element $\langle f|T|i\rangle$ has a δ-function factor assuring the conservation of the total momentum. $T_{i\to f}$ is defined by the relation:

$$\langle f|T|i\rangle = (2\pi)^3\, \delta(\boldsymbol{P}_f + \boldsymbol{k}_f - \boldsymbol{P}_i - \boldsymbol{k}_i)\, T_{i\to f}. \qquad \text{(XXI.248)}$$

The factor $(2\pi)^3$ is introduced here to take into account the normalization adopted for the electron plane wave:

$$\langle \boldsymbol{P}|\boldsymbol{P}'\rangle = (2\pi)^3\, \delta(\boldsymbol{P} - \boldsymbol{P}').$$

The collision cross section is now still given by the general formula (XXI.116).

In order to calculate $\langle f|T|i\rangle$, we expand T into a power series in e and retain only the terms of lowest order, namely the terms in e^2:

$$\langle f|T|i\rangle = \langle f|H''|i\rangle + \langle f|(H^{(I)} + H^{(II)}) \frac{1}{E - H_0 + i\varepsilon} (H^{(I)} + H^{(II)})|i\rangle + O(e^4). \qquad \text{(XXI.249)}$$

It is clear from expressions (XXI.236*a*, *b*) and (XXI.237), that each of the two terms on the right-hand side contains the factor:

$$\int e^{i(\boldsymbol{P}_i + \boldsymbol{k}_i - \boldsymbol{P}_f - \boldsymbol{k}_f)\cdot \boldsymbol{R}}\, d\boldsymbol{R} \equiv (2\pi)^3\, \delta(\boldsymbol{P}_f + \boldsymbol{k}_f - \boldsymbol{P}_i - \boldsymbol{k}_i).$$

This is the above-mentioned conservation-of-momentum factor. From expression (XXI.237),

$$\langle f|H''|i\rangle = \frac{2\pi e^2}{mkL^3} (\boldsymbol{\epsilon}_f^* \cdot \boldsymbol{\epsilon}_i) \int e^{i(\boldsymbol{k}_i + \boldsymbol{P}_i - \boldsymbol{k}_f - \boldsymbol{P}_f)\cdot \boldsymbol{R}}\, d\boldsymbol{R} \qquad \text{(XXI.250)}$$

putting: $k = (k_i k_f)^{\frac{1}{2}}$. Since we are considering only the low-frequency limit, we neglect terms of the order of k/m beside unity and write simply:

$$k_i \simeq k_f \simeq k.$$

It is easily shown that the contribution from the second term on the right-hand side of equation (XXI.249) is an order of k/m higher than

that of the first term. It will therefore be ignored. Thus, in the present approximation ($e^2 \ll 1$, $k \ll m$):

$$T_{i\to f} = \frac{2\pi e^2}{mkL^3}(\boldsymbol{\epsilon}_f^* \cdot \boldsymbol{\epsilon}_i) \tag{XXI.251}$$

whence the scattering cross section [cf. eq. (XXI.116)]:

$$\begin{aligned}\frac{d\sigma_{i\to f}}{d\Omega} &= 2\pi L^3 |T_{i\to f}|^2 \left(\frac{L^3 k^2}{(2\pi)^3}\right) \\ &= \frac{e^4}{m^2} |\boldsymbol{\epsilon}_f^* \cdot \boldsymbol{\epsilon}_i|^2.\end{aligned} \tag{XXI.252}$$

To obtain the scattering cross section for non-polarized photons in the direction $\boldsymbol{k}_f$ we must average over the two possible initial polarization states and sum over the two possible final polarization states. Denoting the scattering angle of the photons by θ ($\cos\theta = \boldsymbol{k}_i \cdot \boldsymbol{k}_f / k^2$), we then obtain:

$$\begin{aligned}\left.\frac{d\sigma_{i\to f}}{d\Omega}\right|_{\text{non pol}} &= \frac{1}{2}\sum_{\boldsymbol{\epsilon}_i \boldsymbol{\epsilon}_f} \frac{d\sigma_{i\to f}}{d\Omega} \\ &= \frac{1}{2}\frac{e^4}{m^2}(1+\cos^2\theta).\end{aligned} \tag{XXI.253}$$

The total scattering cross section for the photons, σ_{tot} is obtained by integrating over the angles of emission. One finds:

$$\sigma_{\text{tot}} = \frac{8\pi}{3}\frac{e^4}{m^2}. \tag{XXI.254}$$

These results are all in good agreement with experiment.

Expressions (XXI.252–254) will be found to be identical with those given by Classical Radiation Theory [1]). In particular, we obtain the *classical Thomson formula* for the total cross section. This identity of the results of the two theories confirms what was said in § I.5 concerning the significance and the limitations of the classical Theory of the Compton effect. The Classical Theory obviously cannot account for the corpuscular character of radiation, nor for the fact that the transfers of momentum and energy between the radiation and the

---

[1]) The constant $\hbar$ does not appear in these expressions. Each of them is proportional to the square of the classical radius of the electron: $r_0 = e^2/mc^2$. Expression (XXI.252), for example, can also be written $r_0^2 \cos^2\alpha$, where α denotes the angle between the initial and final polarizations (supposed rectilinear).

electron take place by discrete quanta; but it correctly predicts the average value of the momentum transferred to the electron per unit time and per unit incident flux [eq. (XXI.254)], and the angular distribution and the polarization of the scattered radiation [eq. (XXI.252)] and also, when the Doppler effect is taken into account, the shift in the wavelength of the scattered radiation [eqs. (I.5) and (I.6)].

The agreement of the two theories is fully realized only in the low-frequency limit where the radiation wavelength, $1/k$, is so large that the scattering process is practically independent of the internal structure of the electron. Since the electron is a quantum object, its dimensions are of the order of $1/m$ and the effects of its structure [1]) are of the order of k/m. In the non-relativistic limit, they are given by the second term in expression (XXI.249). When the incident energy is increased, the quantum nature of the electron becomes more and more evident and one observes increasingly pronounced deviations from the predictions of the Classical Theory.

EXERCISES AND PROBLEMS

1. $\Phi(\boldsymbol{r})$ denoting the real scalar field of section I, we wish to measure the integral of this observable over a finite domain; more generally, we consider the integral:

$$\hat{\Phi} \equiv \int \Phi(\boldsymbol{r})\, P(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r},$$

where $P(\boldsymbol{r})$ is a given real non-negative function whose integral over all space is equal to 1. We put:

$$\chi(\boldsymbol{\varkappa}) \equiv \int \exp\,(\mathrm{i}\boldsymbol{\varkappa}\cdot\boldsymbol{r})\, P(\boldsymbol{r})\, \mathrm{d}\boldsymbol{r} \qquad (\chi(0) = 1).$$

Show that the statistical distribution of $\hat{\Phi}$ in the vacuum state is a Gaussian and that:

$$(\triangle\hat{\Phi})^2 = \int \frac{|\chi(\boldsymbol{\varkappa})|^2}{16\pi^3(\mu^2 + \varkappa^2)^{\frac{1}{2}}}\, \mathrm{d}\boldsymbol{\varkappa}.$$

In particular, if the weight function $P(\boldsymbol{r})$ has appreciable values only in a domain of linear dimensions a, and $\mu a \ll 1$, one has: $\triangle\hat{\Phi} \approx 1/a$.

Treat the same problem with the electromagnetic field and show that in the vacuum state: $\triangle\mathscr{E}_\perp \approx \triangle\mathscr{H} \approx 1/a^2$.

---

[1]) In the classical theory, structural effects are of the order of ke^2/m, the ratio of the classical radius to the wavelength of the radiation.

2. Show that the correction δM defined in § 11 is given to the second order by the expression on the right-hand side of eq. (XXI.76).

3. Consider the elastic scattering problem defined in § 14; suppose the target non-polarized and denote the scattering angle by θ: $\cos\theta = \boldsymbol{k}_i \cdot \boldsymbol{k}_f / k^2$. Show that, in the second-order Born approximation and in the long-wavelength limit ($k\langle R\rangle \ll 1$), the angular distribution of elastically scattered corpuscles is in $\cos^2\theta$ and that the cross section tends to zero in the limit when $k \to 0$ (this property is characteristic of scalar coupling).

4. We consider the resonant scattering problem of § 15 and follow the notation employed there. Suppose that the initial state $|\alpha\rangle$ of the atomic system has a zero angular momentum, and that the resonant state $|\lambda\rangle$ has an angular momentum equal to $\boldsymbol{l}$. Show that, if we neglect the contribution $A^{(\mathrm{pot})}$, the transition amplitude T for the elastic scattering process $\boldsymbol{k}_i \to \boldsymbol{k}_f$ is given by the formula:

$$T = \frac{\Pi}{\omega L^3} \frac{2l+1}{k} P_l(\cos\theta) \frac{\Gamma_{\lambda\to\alpha}}{(E - E_\lambda - \delta E_\lambda) + \frac{1}{2}\mathrm{i}\Gamma_\lambda}$$

in which θ denotes the scattering angle and $\Gamma_{\lambda\to\alpha}$ the probability of the radiative transition $\lambda \to \alpha$ [cf. eq. (XXI.82)] and that consequently the elastic scattering cross section $\mathrm{d}\sigma_{\mathrm{el}}/\mathrm{d}\Omega$ is given by the formula:

$$\frac{\mathrm{d}\sigma_{\mathrm{el}}}{\mathrm{d}\Omega} = \frac{(2l+1)^2}{k^2} P_l^2(\cos\theta) \frac{\Gamma_{\lambda\to\alpha}^2}{4(E - E_\lambda - \delta E_\lambda)^2 + \Gamma_\lambda^2}.$$

Show that these two expressions are a simple generalization of those obtained in the resonance scattering of a particle in a central potential [eqs. (X.64) and (X.65)] and that the discussion and the calculations of § X.15 and X.16 concerning the notion of resonance and of a metastable state apply in the present case.

5. $A_j(\boldsymbol{r})$ and $\mathscr{E}_i(\boldsymbol{r}')$ denoting respectively the jth and the ith components of the potential $\boldsymbol{A}$ (in radiation gauge) and of the transverse electric field $\mathscr{E}_\perp$ prove the commutation relation:

$$[\mathscr{E}_i(\boldsymbol{r}), A_j(\boldsymbol{r}')] = 4\pi\mathrm{i}\theta_{ij}(\boldsymbol{r} - \boldsymbol{r}'),$$

where

$$\theta_{ij}(\boldsymbol{r} - \boldsymbol{r}') \equiv \delta_{ij}\,\delta(\boldsymbol{r} - \boldsymbol{r}') - \frac{\partial^2}{\partial x_i \partial x_j'} \left(\frac{1}{4\pi\,|\boldsymbol{r} - \boldsymbol{r}'|}\right)$$

[$\theta_{ij}(\boldsymbol{r} - \boldsymbol{r}')$ is the projector onto the subspace of transverse vector fields].

6. Prove formula (XXI.220) giving the total momentum of the free radiation field.

7. Show that the intrinsic part $\mathbf{G}^{(s)}$ of the angular momentum of the radiation satisfies the formula:

$$\mathbf{G}^{(s)} = -\mathrm{i} \sum_{\mathbf{k}\varpi\varpi'} a_{(\mathbf{k}\varpi)}{}^{\dagger}\, a_{(\mathbf{k}\varpi')} \left(\boldsymbol{\epsilon}^{(\varpi)*} \times \boldsymbol{\epsilon}^{(\varpi')}\right)$$

(notations of § 27); in particular, if $N_{\mathbf{k}+}$ and $N_{\mathbf{k}-}$ denote the respective number of right-handed and left-handed photons of momentum $\mathbf{k}$ [definition (XXI.222)], we have:

$$\mathbf{G}^{(s)} = \sum_{\mathbf{k}} (N_{\mathbf{k}+} - N_{\mathbf{k}-})\, \mathbf{k}/k.$$

8. Prove relations (XXI.223) and (XXI.224).

9. [1]) Let $u^{LM}(\mathbf{r})$ be a scalar field transforming under rotation like the spherical harmonic $Y_L{}^M(\theta, \varphi)$, and $\mathbf{V}(\mathbf{r}, \mathbf{p})$ a vector (polar or axial) formed from the vectors $\mathbf{r}$ and $\mathbf{p}$. Show that the vector fields $\mathbf{U}^{LM}(\mathbf{r})$ defined by

$$\mathbf{U}^{LM} \equiv \mathbf{V}\,(\mathbf{r}, -\mathrm{i}\nabla)\, u^{LM}(\mathbf{r}),$$

where $\mathbf{V}$ is an operator acting on the function u^{LM}, represent a state of angular momentum (LM).

[In virtue of this theorem, the vector fields $\boldsymbol{\Lambda}_{kjm}$ and $\boldsymbol{\Theta}_{kjm}$ defined in § XXI.29 represent states of angular momentum $(j\, m)$.]

[2]) In the same way as a particle of spin 1 has a 3-component wave function forming a vector field, a particle of integer spin s has a $(2s + 1)$ component wave function forming an irreducible tensor field of order s. Let $X_{\mu}{}^{(s)}(\mathbf{r}, \mathbf{p})$ be the μth component of an irreducible tensor of order s formed with the vectors $\mathbf{r}$ and $\mathbf{p}$. Show that the tensor field $\mathbf{U}^{LM}(\mathbf{r})$ whose μth component is given by

$$U^{LM}(\mathbf{r}) = X_{\mu}{}^{(s)}\,(\mathbf{r}, -\mathrm{i}\nabla)\; U^{LM}(\mathbf{r}) \qquad (\mu = -s, \ldots, +s)$$

represents a state of angular momentum (LM) of this particle of spin s.

10. Using the explicit formulas for the matrix elements eqs. (241*a*, *b*), show that the probability of emission of a photon vanishes if the spins of the initial and final states are both null ($0 \to 0$ *selection rule*).

[N.B. This property is obvious if the multipole expansion is used instead of the plane-wave expansion.]

11. Calculate the probability per unit time for the radiative transition $2p \to 1s$ in the hydrogen atom (one finds: $w_{2p\to 1s} = 6.25 \times 10^8\ \mathrm{sec}^{-1}$).

12. Show that in the long-wavelength approximation, the probability of a *magnetic dipole* radiative transition $\lambda \to \mu\mathbf{k}\boldsymbol{\epsilon}$ (notation of § 30) is given by the formula [cf. eq. (XXI.245)]

$$w_{\lambda\to\mu\mathbf{k}\boldsymbol{\epsilon}} = \frac{k}{2\pi} \left|\langle\lambda|[\mathbf{M}\cdot(\boldsymbol{\epsilon}\times\mathbf{k})]|\mu\rangle\right|^2,$$

where $\boldsymbol{M}$ is the total magnetic moment of the atomic system namely:

$$\boldsymbol{M} = \frac{e}{2m} \sum_i (\boldsymbol{l}_i + 2\boldsymbol{s}_i).$$

[N.B. This formula correctly gives the radiative transition probability when $\Delta\Pi = +1$ and $j_\lambda + j_\mu \leqslant 1$; if $j_\lambda + j_\mu \geqslant 2$, the electric quadrupole contribution is not null and in general cannot be neglected.]

From this deduce that the total magnetic dipole transition probability is given by the formula [cf. eq. (XXI.247)]:

$$w_{\lambda \to \mu}^{(\text{tot})} = \frac{4k^3}{3(2j_\lambda + 1)} |\langle \lambda \| \boldsymbol{M} \| \mu \rangle|^2.$$

Show that in the case of the transition $2s \to 1s$ of the hydrogen atom, this quantity vanishes in the non-relativistic approximation (i.e. if one uses the wave function of the Schrödinger theory).

13. Calculate the cross section for the photoelectric effect for the ground state of the hydrogen atom. Suppose that: $me^4 \ll k \ll m$ ($m \equiv$ mass of the electron, $k \equiv$ momentum of the incident photon, $\hbar = c = 1$); consequently, the energy transferred to the photoelectron is large enough for its final state to be treated as a plane wave, but its velocity v is small enough for it to be treated as a non-relativistic particle. Since the magnetic interaction (the term $H^{(II)}$), gives a negligible contribution, one finds:

$$\frac{d\sigma}{d\Omega} = 4\sqrt{2}\, e^8 \left(\frac{m}{k}\right)^{7/2} \times \frac{e^4}{m^2} \frac{\cos^2 \varphi}{(1 - v \cos \theta)^4}$$

φ and θ denoting respectively the angles made by the direction of emission with those of the polarization (supposed rectilinear) and the momentum of the incident photon. One obtains the results of the semi-classical theory of the effect (cf. Problem XVII.2) by putting $v = 0$ in this formula.

APPENDIX C [1])

VECTOR ADDITION COEFFICIENTS AND ROTATION MATRICES

PLAN OF THE APPENDIX

I. Clebsch–Gordon (C.–G.) Coefficients and "3j" Symbols

1. Angular Momentum. Notations and Conventions.
2. Definition and Notations.
3. Principal Properties.
4. Methods of Calculation.
5. Special Values and Tables.

II. Racah Coefficients and "6j" Symbols

6. Definition and Notations.
7. Principal Properties.
8. Racah Formula and Tables.

III. "9j" Symbols

9. Definition and Principal Properties.

IV. Rotation Matrices

10. Rotations. Rotation Operators. $R^{(J)}$ Matrices.
11. General Properties of the $R^{(J)}$ Matrices.
12. Calculation of the Matrix Elements $R^{(J)}_{MM'}$.
13. Integral Values of J ($J = l$) and Rotation of Spherical Harmonics.

V. Irreducible Tensor Operators

14. Definition and Principal Properties.
15. Tensor Products of Irreducible Tensor Operators.

[1]) This Appendix was prepared in collaboration with J. Horowitz. For a more complete treatment of these questions, see notably A. E. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton University Press, 1957), which, in addition, gives a list of the principal tables now published. See also U. Fano and G. Racah, *Irreducible Tensorial Sets* (Academic Press Inc., New York, 1959).

1. Angular Momentum. Notations and Conventions

The following notations and conventions concerning the angular momentum will be adopted throughout the Appendix.

Units: $\hbar = 1$.

Angular momentum components: $\boldsymbol{J}$, angular momentum operator, having cartesian coordinates J_x, J_y and J_z.

$$J_\pm = J_x \pm \mathrm{i}J_y. \tag{C.1}$$

Commutation relations

$$[J_x, J_y] = \mathrm{i}J_z \qquad [J_y, J_z] = \mathrm{i}J_x \qquad [J_z, J_x] = \mathrm{i}J_y \tag{C.2}$$

$$[J_z, J_\pm] = \pm J_\pm \qquad [J_+, J_-] = 2J_z. \tag{C.3}$$

Basis vectors of a standard representation $\{\boldsymbol{J}^2 J_z\}$: $|\tau JM\rangle$

$$\boldsymbol{J}^2|\tau JM\rangle = J(J+1)\,|\tau JM\rangle \tag{C.4}$$

$$J_z|\tau JM\rangle = M|\tau JM\rangle \tag{C.5}$$

$$J_\pm|\tau JM\rangle = \sqrt{J(J+1)-M(M\pm 1)}\,|\tau J\ M\pm 1\rangle \tag{C.6}$$

$$\langle \tau JM|\tau' J' M'\rangle = \delta_{\tau\tau'}\,\delta_{JJ'}\,\delta_{MM'} \tag{C.7}$$

(J integral or half-integral $\geqslant 0$, $M = -J, -J+1, \ldots, +J$).

τ denotes the quantum numbers that must be added to J and M to form a complete set; in the rest of the Appendix it will be omitted when not needed.

I. CLEBSCH–GORDON (C.–G.) COEFFICIENTS AND "3j" SYMBOLS

2. Definition and Notations

$\boldsymbol{j}_1$, $\boldsymbol{j}_2$ angular momenta of quantum systems 1 and 2 respectively. $\boldsymbol{J}$, angular momentum of the total system of 1 and 2 taken together:

$$\boldsymbol{J} = \boldsymbol{j}_1 + \boldsymbol{j}_2. \tag{C.8}$$

The tensor product of the $(2j_1+1)$ vectors of system 1

$$|j_1 m_1\rangle \qquad (j_1 \text{ fixed}, m_1 = -j_1, \ldots, +j_1)$$

by the $(2j_2+1)$ vectors of system 2

$$|j_2 m_2\rangle \qquad (j_2 \text{ fixed}, m_2 = -j_2, \ldots, +j_2)$$

gives the $(2j_1+1)(2j_2+1)$ simultaneous eigenvectors of $\boldsymbol{j}_1^2$, $\boldsymbol{j}_2^2$, j_{1z}, j_{2z}, the vectors

$$|j_1 j_2 m_1 m_2\rangle \equiv |j_1 m_1\rangle |j_2 m_2\rangle \tag{C.9}$$

from which we can obtain, by a unitary transformation, the $(2j_1+1)$ $(2j_2+1)$ simultaneous eigenvectors of $\boldsymbol{j}_1^2$, $\boldsymbol{j}_2^2$, $\boldsymbol{J}^2$, J_z, the vectors

$$|j_1 j_2 J M\rangle \tag{C.10}$$
$$(J = |j_1 - j_2|, \ldots, j_1 + j_2;\ M = -J, \ldots, +J).$$

Definition

The Clebsch–Gordon [1]), or vector addition, coefficients

$$\langle j_1 j_2 m_1 m_2 | J M\rangle$$

are the coefficients of that unitary transformation:

$$|j_1 j_2 J M\rangle = \sum_{m_1 m_2} |j_1 j_2 m_1 m_2\rangle \langle j_1 j_2 m_1 m_2 | J M\rangle. \tag{C.11}$$

Phase convention [2])

We complete the definition of the vectors in (C.9) and (C.10) by fixing their relative phases as follows:

(*i*) the $|j_1 m_1\rangle$, the $|j_2 m_2\rangle$ and the $|j_1 j_2 J M\rangle$ obey relations (C.6);

(*ii*) $\langle j_1 j_2 j_1 (j_1 - J)\, |J J\rangle$ real > 0.

---

[1]) There are many symbols employed in the literature to denote the Clebsch–Gordon coefficients. We note in particular:

$\langle j_1 j_2 m_1 m_2 | j_1 j_2 J M\rangle$ [Condon and Shortly, *Theory of Atomic Spectra* (University Press, Cambridge, 4th ed., 1957)].

$C_{j_1 j_2}(J M; m_1 m_2)$ [Blatt and Weisskopf, *Theoretical Nuclear Physics* (Wiley, New York, 1952)].

$S^{j_1 j_2}_{j m_1 m_2}$ [E. P. Wigner, *Group Theory and its Application to the Quantum Mechanics of Atomic Spectra* (English translation; Academic Press, New York, 1958)].

[2]) This convention is adopted by most authors, in particular by Wigner, Condon and Shortly, Blatt and Weisskopf, *op. cit.* note 1 of this page, and by Racah, Phys. Rev. **62** (1942) 437.

"$3j$" *symbol* (Wigner) [1]):

$$\begin{pmatrix} j_1 & j_2 & J \\ m_1 & m_2 & -M \end{pmatrix} \equiv \frac{(-)^{j_1-j_2+M}}{\sqrt{2J+1}} \langle j_1 j_2 m_1 m_2 | JM \rangle. \tag{C.12}$$

3. Principal Properties

Reality. They are all real:

$$\langle j_1 j_2 m_1 m_2 | JM \rangle^* = \langle j_1 j_2 m_1 m_2 | JM \rangle.$$

Selection rules

(*i*) $m_1 + m_2 = M$;

(*ii*) $|j_1 - j_2| \leqslant J \leqslant j_1 + j_2$ ("triangular inequalities").

If these two conditions are not met, $\langle j_1 j_2 m_1 m_2 | JM \rangle = 0$.

Symmetries:

$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix}$ is:

(*i*) invariant in a circular permutation of the three columns;

(*ii*) multiplied by $(-)^{j_1+j_2+j_3}$ in a permutation of two columns;

(*iii*) multiplied by $(-)^{j_1+j_2+j_3}$ when we simultaneously change the signs of m_1, m_2, and m_3.

Consequences:

$$\langle j_1 j_2 m_1 m_2 | JM \rangle = (-)^{j_1+j_2-J} \langle j_2 j_1 m_2 m_1 | JM \rangle \tag{C.13a}$$

$$= (-)^{j_1-J+m_2} \sqrt{\frac{2J+1}{2j_1+1}} \langle J\, j_2\, M\, -m_2 | j_1\, m_1 \rangle \tag{C.13b}$$

$$= (-)^{j_2-J-m_1} \sqrt{\frac{2J+1}{2j_2+1}} \langle j_1 J\, -m_1\, M | j_2\, m_2 \rangle \tag{C.13c}$$

$$= (-)^{j_1+j_2-J} \langle j_1 j_2\, -m_1\, -m_2 | J\, -M \rangle. \tag{C.13d}$$

---

[1]) Racah, *op. cit.*, note 2, p. 1055, employs the symbol:

$$V(abc, \alpha\beta\gamma) \equiv (-)^{a-b-c} \begin{pmatrix} a & b & c \\ \alpha & \beta & \gamma \end{pmatrix} \equiv \frac{(-)^{c-\gamma}}{\sqrt{2c+1}} \langle ab\alpha\beta | c\, -\gamma \rangle.$$

Orthogonality relations

$$\sum_{m_1=-j_1}^{+j_1}\sum_{m_2=-j_2}^{+j_2}\langle j_1 j_2 m_1 m_2|JM\rangle\langle j_1 j_2 m_1 m_2|J'M'\rangle = \delta_{JJ'}\delta_{MM'} \qquad (C.14a)$$
$$(|j_1-j_2|\leqslant J\leqslant j_1+j_2;\quad -J\leqslant M\leqslant J)$$

$$\sum_{J=|j_1-j_2|}^{j_1+j_2}\sum_{M=-J}^{+J}\langle j_1 j_2 m_1 m_2|JM\rangle\langle j_1 j_2 m_1' m_2'|JM\rangle = \delta_{m_1 m_1'}\delta_{m_2 m_2'} \qquad (C.14b)$$
$$(-j_1\leqslant m_1\leqslant +j_1;\quad -j_2\leqslant m_2\leqslant +j_2)$$

$$\sum_{m_1=-j_1}^{+j_1}\sum_{m_2=-j_2}^{+j_2}\begin{pmatrix} j_1 & j_2 & j_3\\ m_1 & m_2 & m_3\end{pmatrix}\begin{pmatrix} j_1 & j_2 & j_3'\\ m_1 & m_2 & m_3'\end{pmatrix} = \frac{1}{2j_3+1}\delta_{j_3 j_3'}\delta_{m_3 m_3'} \qquad (C.15a)$$

$$\sum_{j_3=|j_1-j_2|}^{j_1+j_2}\sum_{m_3=-j_3}^{+j_3}(2j_3+1)\begin{pmatrix} j_1 & j_2 & j_3\\ m_1 & m_2 & m_3\end{pmatrix}\begin{pmatrix} j_1 & j_2 & j_3\\ m_1' & m_2' & m_3\end{pmatrix} = \delta_{m_1 m_1'}\delta_{m_2 m_2'}. \qquad (C.15b)$$

Composition relation for the spherical harmonics

$$\int Y_{l_1}^{m_1}(\Omega)\,Y_{l_2}^{m_2}(\Omega)\,Y_{l_3}^{m_3}(\Omega)\,\mathrm{d}\Omega = \left[\frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi}\right]^{\frac{1}{2}}\begin{pmatrix} l_1 & l_2 & l_3\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} l_1 & l_2 & l_3\\ m_1 & m_2 & m_3\end{pmatrix} \qquad (C.16)$$

whence:

$$Y_{l_1}^{m_1}(\Omega)\,Y_{l_2}^{m_2}(\Omega) = \sum_{L=|l_1-l_2|}^{l_1+l_2}\sum_{M=-L}^{L}\left[\frac{(2l_1+1)(2l_2+1)}{4\pi(2L+1)}\right]^{\frac{1}{2}}\langle l_1 l_2 0\,0|L0\rangle\langle l_1 l_2 m_1 m_2|LM\rangle\,Y_L^M(\Omega) \qquad (C.17a)$$

$$= \sum_{L=|l_1-l_2|}^{l_1+l_2}\sum_{M=-L}^{L}(-)^M\left[\frac{(2l_1+1)(2l_2+1)(2L+1)}{4\pi}\right]^{\frac{1}{2}}\begin{pmatrix} l_1 & l_2 & L\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} l_1 & l_2 & L\\ m_1 & m_2 & M\end{pmatrix}Y_L^{-M}(\Omega). \qquad (C.17b)$$

4. Methods of Calculation

Recursion relations

Relating the C.–G. whose arguments differ at most by:

(*i*) $\Delta J=0\qquad \Delta M=+1$

$$\begin{aligned}&\sqrt{J(J+1)-M(M+1)}\,\langle j_1 j_2 m_1 m_2|JM\rangle\\ &=\sqrt{j_1(j_1+1)-m_1(m_1+1)}\,\langle j_1 j_2\, m_1+1\, m_2|J\,M+1\rangle\\ &\quad+\sqrt{j_2(j_2+1)-m_2(m_2+1)}\,\langle j_1 j_2 m_1\, m_2+1|J\,M+1\rangle\end{aligned} \qquad (C.18)$$

(*ii*) $\Delta J=0\qquad \Delta M=-1$

$$\begin{aligned}&\sqrt{J(J+1)-M(M-1)}\,\langle j_1 j_2 m_1 m_2|JM\rangle\\ &=\sqrt{j_1(j_1+1)-m_1(m_1-1)}\,\langle j_1 j_2\, m_1-1\, m_2|J\,M-1\rangle\\ &\quad+\sqrt{j_2(j_2+1)-m_2(m_2-1)}\,\langle j_1 j_2 m_1\, m_2-1|J\,M-1\rangle\end{aligned} \qquad (C.19)$$

(*iii*) $\Delta J = \pm 1 \qquad \Delta M = 0$

$$A_0\langle j_1 j_2 m_1 m_2 | JM\rangle = A_+\langle j_1 j_2 m_1 m_2 | J+1\ M\rangle + A_-\langle j_1 j_2 m_1 m_2 | J-1\ M\rangle \tag{C.20}$$

with

$$A_0 = m_1 - m_2 + M\,\frac{j_2(j_2+1)-j_1(j_1+1)}{J(J+1)} \qquad (M = m_1+m_2)$$

$$A_+ = f(J+1)$$

$$A_- = f(J)$$

$$f(x) = \sqrt{x^2 - M^2}\left[\frac{[(j_1+j_2+1)^2 - x^2][x^2-(j_1-j_2)^2]}{4x^2(2x-1)\,(2x+1)}\right]^{\frac{1}{2}}$$

The Racah formula

$$\begin{pmatrix} a & b & c \\ \alpha & \beta & \gamma \end{pmatrix} = (-)^{a-b-\gamma}\sqrt{\Delta(abc)}\,\sqrt{(a+\alpha)!\,(a-\alpha)!\,(b+\beta)!\,(b-\beta)!\,(c+\gamma)!\,(c-\gamma)!}$$
$$\times \sum_t (-)^t [t!(c-b+t+\alpha)!\,(c-a+t-\beta)!\,(a+b-c-t)!\,(a-t-\alpha)!\,(b-t+\beta)!]^{-1} \tag{C.21}$$

$$(\alpha+\beta+\gamma = 0, \qquad |a-b| \leqslant c \leqslant a+b)$$

with

$$\Delta(abc) \equiv \frac{(a+b-c)!\,(b+c-a)!\,(c+a-b)!}{(a+b+c+1)!} \tag{C.22}$$

$\sum_t$ extends over all integral values of t for which the factorials have a meaning, i.e. for which the arguments of the factorials are positive or null ($0! = 1$). The number of terms in this sum is $\nu+1$, where ν is the smallest of the nine numbers:

$$\begin{array}{lll} a \pm \alpha & b \pm \beta & c \pm \gamma \\ a+b-c & b+c-a & c+a-b. \end{array}$$

5. Special Values and Tables

Special values

(*i*) J and M taking their maximum value:

$$\langle j_1 j_2 j_1 j_2 | j_1+j_2\ j_1+j_2\rangle = 1;$$

(*ii*) one of the j null: $\langle j0m0|jm\rangle = 1$ or

$$\begin{pmatrix} j & j & 0 \\ m & -m & 0 \end{pmatrix} = \frac{(-)^{j-m}}{\sqrt{2j+1}};$$

(*iii*) $m_1 = m_2 = m_3 = 0$:

if $l_1+l_2+l_3$ is odd,

$$\begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix} = 0; \tag{C.23a}$$

if $2p \equiv l_1+l_2+l_3$ is even

$$\begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix} = (-)^p \sqrt{\Delta(l_1\, l_2\, l_3)}\, \frac{p!}{(p-l_1)!\,(p-l_2)!\,(p-l_3)!} \tag{C.23b}$$

(l_1, l_2, l_3 integers $\geqslant 0$ verifying the "triangular inequalities").

Special cases of the Racah formula

The formulae below, or those obtained from them by use of the symmetry relations, give the C.–G. in the following special cases:

(*i*) $m_1 = \pm j_1$ or $m_2 = \pm j_2$ or $M = \pm J$:

$$\langle j_1 j_2 m_1 m_2 | JJ\rangle = \langle j_2 j_1\, -m_2\, -m_1 | J\, -J\rangle$$

$$= (-)^{j_1-m_1} \sqrt{\frac{(2J+1)!\,(j_1+j_2-J)!}{(j_1+j_2+J+1)!\,(J+j_1-j_2)!\,(J+j_2-j_1)!}} \sqrt{\frac{(j_1+m_1)!\,(j_2+m_2)!}{(j_1-m_1)!\,(j_2-m_2)!}} \tag{C.24}$$

$$(m_1+m_2=J);$$

(*ii*) one of the j is the sum of the two others:

if $J = j_1+j_2$

$$\langle j_1 j_2 m_1 m_2 | JM\rangle = \sqrt{\frac{(2j_1)!\,(2j_2)!}{(2J)!}} \sqrt{\frac{(J+M)!\,(J-M)!}{(j_1+m_1)!\,(j_1-m_1)!\,(j_2+m_2)!\,(j_2-m_2)!}} \tag{C.25}$$

$$\langle J\, j_2\, M\, -m_2 | j_1 m_1\rangle = (-)^{j_2-m_2} \sqrt{\frac{2j_1+1}{2J+1}}\, \langle j_1 j_2 m_1 m_2 | JM\rangle. \tag{C.26}$$

Tables of the "3j" symbols

The following tables give expressions for the symbol

$$\begin{pmatrix} j & s & (j+e) \\ m & \mu & (-m\, -\mu) \end{pmatrix}$$

as a function of j and m for

$$s = 0,\ \tfrac{1}{2},\ 1,\ \tfrac{3}{2},\ 2$$

$$0 \leqslant e \leqslant s \qquad 0 \leqslant \mu \leqslant s.$$

With these, and with the aid of the symmetry relations, we can easily calculate any of the C.–G. for which one of the j is equal to 0, $\frac{1}{2}$, 1, $\frac{3}{2}$ or 2.

The tabulated function is the function $\Phi^s_{\mu e}(jm)$ defined by

$$\begin{pmatrix} j & s & (j+e) \\ m & \mu & (-m-\mu) \end{pmatrix} = (-)^{j+m} \sqrt{\frac{(2j+e-s)!}{(2j+e+s+1)!}} \frac{(j+m+\mu+e)!}{(j+m)!} \Phi^s_{\mu e} \quad \text{(C.27)}$$

$s = 0$ $\quad \Phi^0_{00} = 1$

$s = \frac{1}{2}$ $\quad \Phi^{\frac{1}{2}}_{\frac{1}{2}\frac{1}{2}} = 1$

$s = 1$

Table of $\Phi^1_{\mu e}$

	$e = 0$	$e = 1$
$\mu = 0$	$-2m$	$-\sqrt{2(j-m+1)}$
$\mu = 1$	$-\sqrt{2(j-m)}$	1

$s = {}^3/_2$

Table of $\Phi^{3/2}_{\mu e}$

	$e = \frac{1}{2}$	$e = {}^3/_2$
$\mu = \frac{1}{2}$	$j-3m$	$-\sqrt{3(j-m+1)}$
$\mu = {}^3/_2$	$-\sqrt{3(j-m)}$	1

$s = 2$

Table of $\Phi^2_{\mu e}$

	$e = 0$	$e = 1$	$e = 2$
$\mu = 0$	$2[3m^2 - j(j+1)]$	$2m\sqrt{6(j-m+1)}$	$\sqrt{6(j-m+2)(j-m+1)}$
$\mu = 1$	$(2m+1)\sqrt{6(j-m)}$	$2(j-2m)$	$-2\sqrt{j-m+1}$
$\mu = 2$	$\sqrt{6(j-m)(j-m-1)}$	$-2\sqrt{(j-m)}$	1

II. RACAH COEFFICIENTS AND "6j" SYMBOLS

6. Definition and Notations

Coupling of three angular momenta

$\boldsymbol{J}$, total angular momentum of a system formed of three separate component systems, of angular momenta $\boldsymbol{j}$, $\boldsymbol{j}'$, $\boldsymbol{j}''$ respectively:

$$\boldsymbol{J} = \boldsymbol{j}+\boldsymbol{j}'+\boldsymbol{j}''.$$

In the $(2j+1)(2j'+1)(2j''+1)$-dimensional space spanned by the vectors $|mm'm''\rangle \equiv |jm\rangle|j'm'\rangle|j''m''\rangle$ (j, j', j'' given; m, m', m'' variable), the subspace of angular momentum (JM) is usually of more than one dimension [$2J$ and $2M$, integral numbers even or odd like $2(j+j'+j'')$; min $|j \pm j' \pm j''| \leqslant J \leqslant j+j'+j''$; $-J \leqslant M \leqslant +J$].

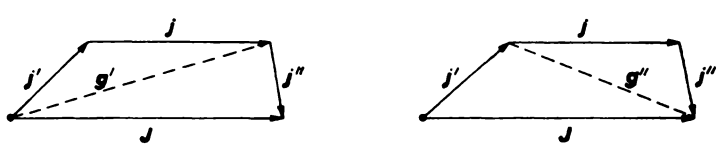
The following two coupling schemes permit the construction of two (in general different) systems of basis vectors for this subspace [1]):

a) $\boldsymbol{j}'+\boldsymbol{j}=\boldsymbol{g}', \boldsymbol{g}'+\boldsymbol{j}''=\boldsymbol{J}$ (Fig. C.1*a*) vectors $|(j'j)g', j''; JM\rangle$

$$|(j'j)g', j''; JM\rangle = \sum_{\substack{mm'm'' \\ \mu'}} |mm'm''\rangle \langle j'jm'm|g'\mu'\rangle \langle g'j''\mu'm''|JM\rangle$$

b) $\boldsymbol{j}+\boldsymbol{j}''=\boldsymbol{g}'', \boldsymbol{j}'+\boldsymbol{g}''=\boldsymbol{J}$ (Fig. C.1*b*) vectors $|j', (jj'')g''; JM\rangle$

$$|j', (jj'')g''; JM\rangle = \sum_{\substack{mm'm'' \\ \mu''}} |mm'm''\rangle \langle jj''mm''|g''\mu''\rangle \langle j'g''m'\mu''|JM\rangle.$$



a) $\boldsymbol{j}' + \boldsymbol{j} = \boldsymbol{g}', \boldsymbol{g}' + \boldsymbol{j}'' = \boldsymbol{J}$ *b)* $\boldsymbol{j} + \boldsymbol{j}'' = \boldsymbol{g}'', \boldsymbol{j}' + \boldsymbol{g}'' = \boldsymbol{J}$

Fig. C.1. $\boldsymbol{J} = \boldsymbol{j} + \boldsymbol{j}' + \boldsymbol{j}''$. Different coupling schemes.

We can pass from one system to the other by a certain unitary transformation.

---

[1]) The order in which the different vectors are coupled is important. By changing that order, for example by interchanging the order of j and j'' in the coupling scheme (*a*), we may modify the sign of the resulting vector.

Definition of the Wigner "6j" symbols $\begin{Bmatrix} j_1 & j_2 & j_3 \\ J_1 & J_2 & J_3 \end{Bmatrix}$

They are related to the coefficients of this unitary transformation by the defining relation

$$\langle j', (j\,j'')g''; JM | (j'\,j)g', j''; J'M' \rangle = \delta_{JJ'}\delta_{MM'}\sqrt{(2g'+1)(2g''+1)}\,(-)^{j+j'+j''+J}\begin{Bmatrix} j' & j & g' \\ j'' & J & g'' \end{Bmatrix} \tag{C.28}$$

i.e.

$$|(j'\,j)g', j''; JM\rangle = \sum_{g''} |j', (j\,j'')g''; JM\rangle \sqrt{(2g'+1)(2g''+1)}\,(-)^{j+j'+j''+J}\begin{Bmatrix} j' & j & g' \\ j'' & J & g'' \end{Bmatrix} \tag{C.29}$$

The Racah W coefficients [1])

Racah uses certain W coefficients equal to the "6j" to within a sign:

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ J_1 & J_2 & J_3 \end{Bmatrix} = (-)^{j_1+j_2+J_1+J_2}\,W(j_1 j_2 J_2 J_1; j_3 J_3). \tag{C.30}$$

Associated coefficients

In expressions for angular distributions one sometimes encounters the coefficients $Z(LJL'J'; j\lambda)$ and $F_\lambda(LL'J_1J_2)$, for which there are detailed tables, and which are defined respectively by

$$Z(L\,J\,L'\,J'; j\,\lambda) = (-)^{J+J'+\frac{1}{2}(L'-L+\lambda)}\sqrt{(2L+1)(2L'+1)(2J+1)(2J'+1)(2\lambda+1)} \times \begin{pmatrix} L & L' & \lambda \\ 0 & 0 & 0 \end{pmatrix} \begin{Bmatrix} L & L' & \lambda \\ J' & J & j \end{Bmatrix}. \tag{C.31a}$$

$$F_\lambda(L\,L'\,J_1J_2) = (-)^{J_1+J_2-1}\sqrt{(2L+1)(2L'+1)(2J_2+1)(2\lambda+1)} \times \begin{pmatrix} L & L' & \lambda \\ 1 & -1 & 0 \end{pmatrix} \begin{Bmatrix} L & L' & \lambda \\ J_2 & J_2 & J_1 \end{Bmatrix}. \tag{C.31b}$$

The $F_\lambda(LL'J_1J_2)$ are involved notably when the reaction products include particles of spin 1, and the polarization is not observed.

---

[1]) Racah, *op. cit.* The sign factors disappear from the two preceding relations when we substitute the W for the "6j". However, the "6j" have simpler symmetry relations (cf. § 7).

Associated tetrahedron for the "6j"

In order to avoid certain confusions in the manipulation of the W coefficients and the "6j" symbols, it is useful to associate with each

$$\begin{Bmatrix} j' & j & g' \\ j'' & J & g'' \end{Bmatrix}$$

a certain tetrahedron in which each edge represents one of the six angular momenta of the symbol (Fig. C.2). In this representation,

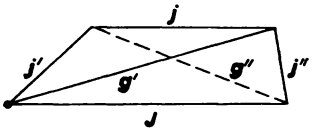


Fig. C.2. Tetrahedron associated with the symbol $\begin{Bmatrix} j' & j & g' \\ j'' & J & g'' \end{Bmatrix}$.

each pair of opposite edges is associated with the two angular momenta from a given column, and also the three angular momenta of the first line correspond to the edges of one of the faces of the tetrahedron. In addition, to each face correspond three angular momenta, of which the one is obtained by vector addition of the other two, as follows from the definition of the "6j".

7. Principal Properties

Reality: the "6j" are all real.

Selection rules

In order to have $\begin{Bmatrix} j_1 & j_2 & j_3 \\ J_1 & J_2 & J_3 \end{Bmatrix} \neq 0$,

the three angular momenta represented by each face of the tetrahedron must be such that it is possible for one of them to be the sum of the two others. In other words, it is necessary that the elements of each of the triads

$$(j_1 j_2 j_3) \quad (j_1 J_2 J_3) \quad (J_1 j_2 J_3) \quad (J_1 J_2 j_3)$$

(*i*) satisfy the triangular inequalities;

(*ii*) have an integral sum.

(*N.B.* either the six j are integral, or the three j of a same face are integral, or two j corresponding to opposite edges are integral.)

Symmetry relations. The "6j" symbol is invariant

(*i*) in a permutation of its columns,

e.g.: $$\begin{Bmatrix} j_1 & j_2 & j_3 \\ J_1 & J_2 & J_3 \end{Bmatrix} = \begin{Bmatrix} j_2 & j_1 & j_3 \\ J_2 & J_1 & J_3 \end{Bmatrix};$$

(*ii*) in an exchange of two elements of the first line with the corresponding elements in the second line,

e.g.: $$\begin{Bmatrix} j_1 & j_2 & j_3 \\ J_1 & J_2 & J_3 \end{Bmatrix} = \begin{Bmatrix} J_1 & J_2 & j_3 \\ j_1 & j_2 & J_3 \end{Bmatrix}.$$

In other words, to define a "6j" symbol it is sufficient to give the six angular momenta and their relative positions on the associated tetrahedron.

Fundamental relations in terms of the C.–G.

The following relations relate the symbol $\begin{Bmatrix} j_1 & j_2 & j_3 \\ J_1 & J_2 & J_3 \end{Bmatrix}$ to the C.–G. coefficients formed with the triads $(j_1 j_2 j_3)$, $(j_1 J_2 J_3)$, $(J_1 j_2 J_3)$, $(J_1 J_2 j_3)$ relative to each of the four faces of the associated tetrahedron:

$$\sum_{\substack{M_1 M_2 M_3 \\ m_1 m_2}} (-)^{J_1+J_2+J_3+M_1+M_2+M_3} \times \begin{pmatrix} J_1 & J_2 & j_3 \\ M_1 & -M_2 & m_3 \end{pmatrix} \begin{pmatrix} J_2 & J_3 & j_1 \\ M_2 & -M_3 & m_1 \end{pmatrix} \begin{pmatrix} J_3 & J_1 & j_2 \\ M_3 & -M_1 & m_2 \end{pmatrix} \begin{pmatrix} j_1 & j_2 & j_3' \\ m_1 & m_2 & m_3' \end{pmatrix} = \delta_{j_3 j_3'} \delta_{m_3 m_3'} \frac{1}{2j_3+1} \begin{Bmatrix} j_1 & j_2 & j_3 \\ J_1 & J_2 & J_3 \end{Bmatrix}. \tag{C.32}$$

(*N.B.* This sum is actually only over two indices, the M and the m being related.)

$$\sum_{M_1 M_2 M_3} (-)^{J_1+J_2+J_3+M_1+M_2+M_3} \begin{pmatrix} J_1 & J_2 & j_3 \\ M_1 & -M_2 & m_3 \end{pmatrix} \begin{pmatrix} J_2 & J_3 & j_1 \\ M_2 & -M_3 & m_1 \end{pmatrix} \begin{pmatrix} J_3 & J_1 & j_2 \\ M_3 & -M_1 & m_2 \end{pmatrix} = \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \begin{Bmatrix} j_1 & j_2 & j_3 \\ J_1 & J_2 & J_3 \end{Bmatrix}. \tag{C.33}$$

(*N.B.* This sum is actually only over one index.)

$$\sum_{g m_g} (-)^{g+m_g} (2g+1) \begin{Bmatrix} j_1 & J_1 & g \\ J_2 & j_2 & f \end{Bmatrix} \begin{pmatrix} j_1 & J_1 & g \\ m_1 & M_1 & -m_g \end{pmatrix} \begin{pmatrix} j_2 & J_2 & g \\ m_2 & M_2 & m_g \end{pmatrix} = (-)^{j_2+J_1+f+g} \sum_{m_f} (-)^{f+m_f} \begin{pmatrix} j_1 & j_2 & f \\ m_1 & m_2 & -M_f \end{pmatrix} \begin{pmatrix} J_1 & J_2 & f \\ M_1 & M_2 & m_f \end{pmatrix}. \tag{C.34}$$

(*N.B.* The sums over m_g and m_f have each just one term.)

The Racah–Elliot relations and the orthogonality relation

$$\sum_x (-)^{2x}(2x+1)\begin{Bmatrix} a & b & x \\ a & b & f \end{Bmatrix} = 1. \tag{C.35a}$$

$$\sum_x (-)^{a+b+x}(2x+1)\begin{Bmatrix} a & b & x \\ b & a & f \end{Bmatrix} = \delta_{f_0}\sqrt{(2a+1)(2b+1)}. \tag{C.35b}$$

$$\sum_x (2x+1)\begin{Bmatrix} a & b & x \\ c & d & f \end{Bmatrix}\begin{Bmatrix} c & d & x \\ a & b & g \end{Bmatrix} = \delta_{fg}\frac{1}{2f+1}. \tag{C.35c}$$

$$\sum_x (-)^{f+g+x}(2x+1)\begin{Bmatrix} a & b & x \\ c & d & f \end{Bmatrix}\begin{Bmatrix} c & d & x \\ b & a & g \end{Bmatrix} = \begin{Bmatrix} a & d & f \\ b & c & g \end{Bmatrix}. \tag{C.35d}$$

$$\sum_x (-)^{a+b+c+d+e+f+g+h+x+j}(2x+1)\begin{Bmatrix} a & b & x \\ c & d & g \end{Bmatrix}\begin{Bmatrix} c & d & x \\ e & f & h \end{Bmatrix}\begin{Bmatrix} e & f & x \\ b & a & j \end{Bmatrix} = \begin{Bmatrix} g & h & j \\ e & a & d \end{Bmatrix}\begin{Bmatrix} g & h & j \\ f & b & c \end{Bmatrix}. \tag{C.35e}$$

8. Racah Formula and Tables

The Racah formula

$$\begin{Bmatrix} j_1 & j_2 & j_3 \\ J_1 & J_2 & J_3 \end{Bmatrix} = [\Delta(j_1 j_2 j_3)\,\Delta(j_1 J_2 J_3)\,\Delta(J_1 j_2 J_3)\,\Delta(J_1 J_2 j_3)]^{\frac{1}{2}} \times \sum_t \frac{(-)^t (t+1)!}{(t-j_1-j_2-j_3)!\,(t-j_1-J_2-J_3)!\,(t-J_1-j_2-J_3)!\,(t-J_1-J_2-j_3)!} \times (j_1+j_2+J_1+J_2-t)!\,(j_2+j_3+J_2+J_3-t)!\,(j_3+j_1+J_3+J_1-t)! \tag{C.36}$$

Same definition of $\Delta(abc)$ and same summation convention as for (C.22). The number of terms in $\sum_t$ is equal to $1+\sigma$ where σ is the smallest of the twelve numbers

$$\begin{array}{llll} j_1+j_2-j_3 & j_1+J_2-J_3 & J_1+j_2-J_3 & J_1+J_2-j_3 \\ j_2+j_3-j_1 & J_2+J_3-j_1 & j_2+J_3-J_1 & J_2+j_3-J_1 \\ j_3+j_1-j_2 & J_3+j_1-J_2 & J_3+J_1-j_2 & j_3+J_1-J_2 \end{array}$$

Special cases

(*i*) one of the j is null

$$\begin{Bmatrix} j & j' & 0 \\ J & J' & g \end{Bmatrix} = (-)^{j+J+g}\frac{\delta_{jj'}\delta_{JJ'}}{\sqrt{(2j+1)(2J+1)}} \qquad (|j-J| \leqslant g \leqslant j+J) \tag{C.37}$$

(*ii*) one of the j is equal to $\boxed{\tfrac{1}{2}}$

$$\begin{Bmatrix} j & j+\frac{1}{2} & \frac{1}{2} \\ J & J+\frac{1}{2} & g+\frac{1}{2} \end{Bmatrix} = (-)^{1+g+j+J}\left[\frac{(1+g+j-J)\,(1+g-j+J)}{(2j+1)\,(2j+2)\,(2J+1)\,(2J+2)}\right]^{\frac{1}{2}} \quad (C.38)$$
$$(|j-J| \leqslant g \leqslant j+J).$$

$$\begin{Bmatrix} j & j+\frac{1}{2} & \frac{1}{2} \\ J+\frac{1}{2} & J & g \end{Bmatrix} = (-)^{1+g+j+J}\left[\frac{(1-g+j+J)\,(2+g+j+J)}{(2j+1)\,(2j+2)\,(2J+1)\,(2J+2)}\right]^{\frac{1}{2}} \quad (C.39)$$
$$(|j-J| \leqslant g \leqslant j+J).$$

III. "9j" SYMBOLS

9. Definition and Principal Properties

Coupling of four angular momenta and definition of the "9j":

$\boldsymbol{J}$, total angular momentum of a system formed of four separate component systems with respective angular momenta $\boldsymbol{j}_1$, $\boldsymbol{j}_2$, $\boldsymbol{j}_3$, $\boldsymbol{j}_4$:

$$\boldsymbol{J} = \boldsymbol{j}_1 + \boldsymbol{j}_2 + \boldsymbol{j}_3 + \boldsymbol{j}_4.$$

In the $\prod_{i=1}^{4}(2j_i+1)$ dimensional space spanned by the vectors

$$|m_1 m_2 m_3 m_4\rangle \equiv \prod_{i=1}^{4} |j_i m_i\rangle \qquad (j_i \text{ given, } m_i \text{ variable}),$$

the following two coupling schemes lead to two distinct systems of basis vectors for the subspace of angular momentum (JM):

a) $\boldsymbol{j}_1+\boldsymbol{j}_2=\boldsymbol{J}_{12}$ $\boldsymbol{j}_3+\boldsymbol{j}_4=\boldsymbol{J}_{34}$ $\boldsymbol{J}_{12}+\boldsymbol{J}_{34}=\boldsymbol{J}$
vectors $|(j_1 j_2)J_{12}, j_3 j_4)(J_{34}); JM\rangle$;

b) $\boldsymbol{j}_1+\boldsymbol{j}_3=\boldsymbol{J}_{13}$ $\boldsymbol{j}_2+\boldsymbol{j}_4=\boldsymbol{J}_{24}$ $\boldsymbol{J}_{13}+\boldsymbol{J}_{24}=\boldsymbol{J}$
vectors $|(j_1 j_3)J_{13}, (j_2 j_4)J_{24}; JM\rangle$.

The Wigner "9j" symbols are to within a constant the coefficients of the unitary transformation that takes us from one basis to the other. By definition

$$\langle (j_1 j_2)J_{12}, (j_3 j_4)J_{34}; JM | (j_1 j_3)J_{13}, (j_2 j_4)J_{24}; J'M'\rangle$$
$$= \delta_{JJ'}\,\delta_{MM'}\sqrt{(2J_{12}+1)\,(2J_{34}+1)\,(2J_{13}+1)\,(2J_{24}+1)}\begin{Bmatrix} j_1 & j_2 & J_{12} \\ j_3 & j_4 & J_{34} \\ J_{13} & J_{24} & J \end{Bmatrix}. \quad (C.40a$$

One can also define the "9j" by

$$\langle [j_1 j_2) J_{12}, j_3] J_{123} j_4; JM | [(j_4 j_2) J_{42}, j_3] J_{423} j_1; J'M' \rangle \tag{C.40b}$$
$$= (-)^{J_{423}+j_1-J_{123}-j_4} \delta_{JJ'} \delta_{MM'} \sqrt{(2J_{12}+1)(2J_{123}+1)(2J_{42}+1)(2J_{423}+1)} \begin{Bmatrix} j_2 & J_{12} & j_1 \\ J_{42} & j_3 & J_{423} \\ j_4 & J_{123} & J \end{Bmatrix}.$$

The following notation is also employed

$$\times \begin{pmatrix} j_1 & j_2 & J_{12} \\ j_3 & j_4 & J_{34} \\ J_{13} & J_{24} & J \end{pmatrix} = \begin{Bmatrix} j_1 & j_2 & J_{12} \\ j_3 & j_4 & J_{34} \\ J_{13} & J_{24} & J \end{Bmatrix}.$$

A relation in terms of the "3j" symbols

$$\begin{pmatrix} J_{13} & J_{24} & J \\ M_{13} & M_{24} & M \end{pmatrix} \begin{Bmatrix} j_1 & j_2 & J_{12} \\ j_3 & j_4 & J_{34} \\ J_{13} & J_{24} & J \end{Bmatrix} = \sum_{\substack{m_1 m_2 m_3 m_4 \\ M_{13} M_{24}}} \begin{pmatrix} j_1 & j_2 & J_{12} \\ m_1 & m_2 & M_{12} \end{pmatrix} \tag{C.40c}$$
$$\times \begin{pmatrix} j_3 & j_4 & J_{34} \\ m_3 & m_4 & M_{34} \end{pmatrix} \begin{pmatrix} j_1 & j_3 & J_{13} \\ m_1 & m_3 & M_{13} \end{pmatrix} \begin{pmatrix} j_2 & j_4 & J_{24} \\ m_2 & m_4 & M_{24} \end{pmatrix} \begin{pmatrix} J_{12} & J_{34} & J \\ M_{12} & M_{34} & M \end{pmatrix}.$$

Symmetry relations

$$\begin{Bmatrix} J_1 & J_2 & J_3 \\ J_4 & J_5 & J_6 \\ J_7 & J_8 & J_9 \end{Bmatrix} \quad \text{is}$$

(i) multiplied by $(-)^R$, where $R = \sum_{i=1}^{9} J_i$, in the exchange of two lines or two columns;

(ii) invariant in a reflection through one of the diagonals.

Orthogonality relation

$$\sum_{J_{13} J_{24}} (2J_{13}+1)(2J_{24}+1) \begin{Bmatrix} j_1 & j_2 & J_{12} \\ j_3 & j_4 & J_{34} \\ J_{13} & J_{24} & J \end{Bmatrix} \begin{Bmatrix} j_1 & j_2 & J_{12}' \\ j_3 & j_4 & J_{34}' \\ J_{13} & J_{24} & J \end{Bmatrix} = \frac{\delta_{J_{12} J_{12}'} \delta_{J_{34} J_{34}'}}{(2J_{12}+1)(2J_{34}+1)}.$$

A relation in terms of the "6j" symbols

$$\begin{Bmatrix} j_1 & j_2 & J_{12} \\ j_3 & j_4 & J_{34} \\ J_{13} & J_{24} & J \end{Bmatrix} = \sum_g (-)^{2g} (2g+1) \begin{Bmatrix} j_1 & j_2 & J_{12} \\ J_{34} & J & g \end{Bmatrix} \begin{Bmatrix} j_3 & j_4 & J_{34} \\ j_2 & g & J_{24} \end{Bmatrix} \begin{Bmatrix} J_{13} & J_{24} & J \\ g & j_1 & j_3 \end{Bmatrix}. \tag{C.41}$$

Case when one of the j is null

$$\begin{Bmatrix} j_1 & j_2 & f \\ j_3 & j_4 & f' \\ g & g' & 0 \end{Bmatrix} = \delta_{ff'}\,\delta_{gg'}\,\frac{(-)^{j_2+j_3+f+g}}{\sqrt{(2f+1)\,(2g+1)}}\begin{Bmatrix} j_1 & j_2 & f \\ j_4 & j_3 & g \end{Bmatrix}. \tag{C.42}$$

IV. ROTATION MATRICES

10. Rotations. Rotation Operators. $R^{(J)}$ Matrices

$$\begin{aligned} \mathscr{R}_{\boldsymbol{u}}(\varphi) &\equiv \text{Rotation through an angle } \varphi \text{ about the axis } \boldsymbol{u} \\ \mathscr{R}(\alpha\beta\gamma) &\equiv \text{Rotation with Euler angles } (\alpha, \beta, \gamma) \\ &\equiv \mathscr{R}_Z(\gamma)\,\mathscr{R}_{\boldsymbol{u}}(\beta)\,\mathscr{R}_z(\alpha) \end{aligned} \tag{C.43}$$

The operations on the right-hand side are to be performed in the order from *right to left* (cf. Fig. C.3).

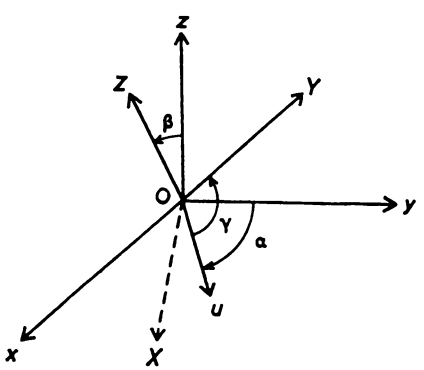


Fig. C.3. The Euler angles

$\alpha = (Oy, Ou), \qquad \beta = (Oz, OZ), \qquad \gamma = (Ou, OY).$

Associated matrix: matrix for the transformation of the coordinates of vectors (denoted by the same letter $\mathscr{R}$ as the rotation itself).

If $\boldsymbol{V}(V_1, V_2, V_3)$ is any vector, and $\boldsymbol{V}'(V_1', V_2', V_3')$ its transform in the rotation $\mathscr{R}$, we have:

$$V_i' = \mathscr{R}_{ij} V_j, \quad \mathscr{R}_{ij} \text{ elements of the associated matrix.}$$

[*Consequence.* $\boldsymbol{A}_j$ being the transform of the unit vector $\boldsymbol{a}_j$ along

the jth axis ($j=1$, 2 or 3) we have: $\boldsymbol{A}_j \equiv \mathscr{R}[\boldsymbol{a}_j] = \boldsymbol{a}_i \mathscr{R}_{ij}$, $\mathscr{R}_{ij} = \boldsymbol{a}_i \cdot \boldsymbol{A}_j]$

$$\mathscr{R}^* = \mathscr{R} \qquad \tilde{\mathscr{R}} = \mathscr{R}^{-1} \qquad \det \mathscr{R} = 1 \tag{C.44}$$

$$\mathscr{R}(\alpha\beta\gamma) \equiv \begin{pmatrix} \cos\gamma\cos\beta\cos\alpha - \sin\gamma\sin\alpha & -\sin\gamma\cos\beta\cos\alpha - \cos\gamma\sin\alpha & \sin\beta\cos\alpha \\ \cos\gamma\cos\beta\sin\alpha + \sin\gamma\cos\alpha & -\sin\gamma\cos\beta\sin\alpha + \cos\gamma\cos\alpha & \sin\beta\sin\alpha \\ -\cos\gamma\sin\beta & \sin\gamma\sin\beta & \cos\beta \end{pmatrix}. \tag{C.45}$$

Rotation operator

Unitary operator R, which applied to a ket $|\rangle$ gives its transform in the rotation $\mathscr{R}$:

$$\mathscr{R}[|\rangle] \equiv R|\rangle \qquad R^\dagger R = RR^\dagger = 1. \tag{C.46}$$

If Q is an observable of the quantum system:

$$\mathscr{R}[Q] \equiv RQR^{-1}. \tag{C.47}$$

If $\boldsymbol{B} \equiv (B_x, B_y, B_z)$ is a vector operator attached to the system ($B_i \equiv \boldsymbol{B} \cdot \boldsymbol{a}_i$)

$$\mathscr{R}[B_i] \equiv R B_i R^{-1} = (\boldsymbol{B} \cdot \boldsymbol{A}_i) = \tilde{\mathscr{R}}_{ij} B_j. \tag{C.48}$$

N.B. It is the inverse of $\mathscr{R}$ and not $\mathscr{R}$ itself that occurs here.

The application of (C.48) to the transformation of the components of the angular momentum $\boldsymbol{J}$ in the rotation $\mathscr{R}(\alpha\beta\gamma)$ gives

$$RJ_\pm R^{-1} = e^{\mp i\gamma}\left[\frac{1+\cos\beta}{2} e^{\mp i\alpha} J_\pm - \frac{1-\cos\beta}{2} e^{\pm i\alpha} J_\mp - \sin\beta J_z\right] \tag{C.49a}$$

$$RJ_z R^{-1} = \tfrac{1}{2}\sin\beta\,(e^{-i\alpha}J_+ + e^{i\alpha}J_-) + \cos\beta\, J_z. \tag{C.49b}$$

Expression in terms of the total angular momentum $\boldsymbol{J}$

Infinitesimal rotation:

$$R_u(\varepsilon) = 1 - i\varepsilon(\boldsymbol{J}\cdot\boldsymbol{u}) \qquad (\varepsilon \ll 1). \tag{C.50}$$

Finite rotations:

$$R_u(\varphi) = e^{-i\varphi(\boldsymbol{J}\cdot\boldsymbol{u})} \tag{C.51}$$

$$R(\alpha\beta\gamma) = e^{-i\alpha J_z} e^{-i\beta J_y} e^{-i\gamma J_z}. \tag{C.52}$$

Correspondence between rotations and rotation operators

There is a one-to-one correspondence between infinitesimal rotations and R operators infinitely close to 1, but the same is not necessarily true for the finite rotations.

To any finite rotation $\mathscr{R}$, there correspond in general two rotations, R' and R'', related by the equation

$$R'' = DR'$$

where D is the operator defined by:

$$D = \begin{cases} +1 & \text{if } J \text{ is integral} \\ -1 & \text{if } J \text{ is half-integral.} \end{cases}$$

In order to have $R' = R''$, it is necessary that the space of state vectors be made up exclusively of states of integral J.

$$R_u(2\pi) = D$$
$$R_u(4\pi) = 1.$$

Let $(\alpha\beta\gamma)$ and $(\alpha_1\beta_1\gamma_1)$ be two sets of Euler angles defining the same rotation [eq. (XIII.42)]:

$$R(\alpha_1\beta_1\gamma_1) = D^{n_\alpha+n_\beta+n_\gamma} R(\alpha\beta\gamma). \tag{C.53}$$

Rotation matrices $R^{(J)}(\alpha\beta\gamma)$

Matrix of $(2J+1)$ dimensions with elements:

$$\begin{aligned} R^{(J)}_{MM'}(\alpha\beta\gamma) &\equiv \langle JM|R(\alpha\beta\gamma)|JM'\rangle \\ &\equiv \langle JM|\, e^{-i\alpha J_z} e^{-i\beta J_y} e^{-i\gamma J_z}\, |JM'\rangle. \end{aligned} \tag{C.54}$$

The vectors $|JM\rangle$ (J fixed, $M = -J, \ldots, +J$) are $(2J+1)$ eigenvectors of J^2 and J_z, obtained one from another by relations (C.6).

Matrix $r^{(J)}(\beta)$

$$\begin{aligned} r^{(J)}(\beta) &\equiv R^{(J)}(0, \beta, 0) \\ r^{(J)}_{MM'}(\beta) &\equiv \langle JM|\, e^{-i\beta J_y}\, |JM'\rangle \end{aligned} \tag{C.55}$$

$$R^{(J)}_{MM'}(\alpha\beta\gamma) = e^{-i\alpha M}\, r^{(J)}_{MM'}(\beta)\, e^{-i\gamma M'}. \tag{C.56}$$

11. General Properties of the $R^{(J)}$ Matrices

Inverse:

$$[R^{(J)}(\alpha\beta\gamma)]^{-1} = R^{(J)}(-\gamma, -\beta, -\alpha). \tag{C.57}$$

Determinant:

$$\det R^{(J)} = 1. \tag{C.58}$$

Uniformity:

$$R_u{}^{(J)}(2\pi) = (-)^{2J}. \tag{C.59}$$

To each set of Euler angles there corresponds a single $R^{(J)}$ matrix.

To each rotation $\mathscr{R}$ there corresponds a single $R^{(J)}$ matrix if J is integral, and two matrices of which one is the negative of the other if J is half-integral.

Reality: $r^{(J)}(\beta)$ is a real matrix.

Rotations through an angle π about the axes

Notation: X, Y, $Z \equiv$ Rotation operators for rotations through $+\pi$ about Ox, Oy, Oz respectively.

$$X^2 = Y^2 = Z^2 = XYZ = (-)^{2J}. \tag{C.60}$$

$$X|JM\rangle = \mathrm{e}^{-\imath\pi J}|J\,-M\rangle, \quad \text{whence } X^{(J)}_{MM'} = \mathrm{e}^{-\imath\pi J}\,\delta_{M\,-M'}. \tag{C.61}$$

$$Y|JM\rangle = (-)^{J-M}|J\,-M\rangle, \quad \text{whence } Y^{(J)}_{MM'} = (-)^{J+M}\,\delta_{M\,-M'}. \tag{C.62}$$

$$Z|JM\rangle = \mathrm{e}^{-\imath\pi M}|JM\rangle, \quad \text{whence } Z^{(J)}_{MM'} = \mathrm{e}^{-\imath\pi M}\,\delta_{MM'}. \tag{C.63}$$

Transformation of the angular momentum operator in the rotation Y:

$$YJ_xY^\dagger = -J_x \quad YJ_yY^\dagger = J_y \quad YJ_zY^\dagger = -J_z \quad YJ_\pm Y^\dagger = -J_\mp$$

whence:

$$YR^{(J)}Y^\dagger = R^{(J)*}. \tag{C.64}$$

Symmetry relations [deduced from (C.62) and (C.64)]

$$r^{(J)}_{MM'} = (-)^{M-M'}\,r^{(J)}_{-M,\,-M'}. \tag{C.65}$$

$$R^{(J)*}_{MM'} = (-)^{M-M'}\,R^{(J)}_{-M,\,-M'}. \tag{C.66}$$

Unitarity and orthogonality relations

$R^{(J)\dagger} = R^{(J)-1}$ whence the unitarity relations

$$\sum_M R^{(J)}_{MM'}\,R^{(J)*}_{MM''} = \delta_{M'M''} \qquad \sum_M R^{(J)}_{M'M}\,R^{(J)*}_{M''M} = \delta_{M'M''} \tag{C.67}$$

and, taking into account the symmetry relations, the orthogonality relations

$$\begin{aligned} \sum_M (-)^{J+M}\,R^{(J)}_{MM'}\,R^{(J)}_{-M\,-M''} &= (-)^{J+M'}\,\delta_{M'M''} \\ \sum_M (-)^{J+M}\,R^{(J)}_{M'M}\,R^{(J)}_{-M''\,-M} &= (-)^{J+M'}\,\delta_{M'M''}. \end{aligned} \tag{C.68}$$

Composition and reduction formulae

In the following formulae the matrices $R^{(j_1)}$, $R^{(j_2)}$, $R^{(J)}$ are relative to the same Euler angles.

Reduction of the tensor product $R^{(j_1)} \otimes R^{(j_2)}$:

$$R^{(j_1)}_{m_1 m_1'} R^{(j_2)}_{m_2 m_2'} = \sum_{J=|j_1-j_2|}^{j_1+j_2} \sum_{M, M'=-J}^{+J} \langle j_1 j_2 m_1 m_2 | JM \rangle R^{(J)}_{MM'} \langle j_1 j_2 m_1' m_2' | JM' \rangle. \quad \text{(C.69)}$$

Composition formula:

$$R^{(J)}_{MM'} = \sum_{m_1, m_1'=-j_1}^{+j_1} \sum_{m_2, m_2'=-j_2}^{+j_2} \langle j_1 j_2 m_1 m_2 | JM \rangle R^{(j_1)}_{m_1 m_1'} R^{(j_2)}_{m_2 m_2'} \langle j_1 j_2 m_1' m_2' | JM' \rangle. \quad \text{(C.70)}$$

In particular, if $J = j_1 + j_2$

$$\begin{aligned} R^{(J)}_{JJ} &= R^{(j_1)}_{j_1 j_1} R^{(j_2)}_{j_2 j_2} \\ R^{(J)}_{J,-J} &= R^{(j_1)}_{j_1,-j_1} R^{(j_2)}_{j_2,-j_2}. \end{aligned} \quad \text{(C.71)}$$

12. Calculation of the Matrix Elements $R^{(J)}_{MM'}$

Principal methods of calculation

When the matrix $r^{(J)}(\beta)$ is known, we can easily deduce the matrix $R^{(J)}(\alpha\beta\gamma)$ from (C.56).

$r^{(J)}$ is real, unitary, and has the symmetry property (C.65):

$$r^{(J)}_{MM'}(\beta) = r^{(J)}_{M',M}(-\beta) = (-)^{M-M'} r^{(J)}_{-M,-M'}(\beta).$$

Thus we need only to know the matrix elements corresponding to $M > 0$ and $M > M'$ to deduce all of them.

As for these we may:

a) calculate them directly by the Wigner formula given below;

b) obtain them by application of formula (C.70) for the composition of matrix elements $r^{(j_1)}$, $r^{(j_2)}$ for smaller angular momenta; in particular all of the $r^{(J)}$ can be obtained one after the other from $r^{(\frac{1}{2})}$;

c) obtain them one from another using the recursion relations that follow from equations (C.49).

The Wigner formula

Put:

$$\xi \equiv \cos \tfrac{1}{2}\beta, \qquad \eta \equiv \sin \tfrac{1}{2}\beta.$$

We have:

$$r^{(J)}_{MM'} = \sum_t (-)^t \frac{\sqrt{(J+M)!\,(J-M)!\,(J+M')!\,(J-M')!}}{(J+M-t)!\,(J-M'-t)!\,t!\,(t-M+M')!} \times \xi^{2J+M-M'-2t} \eta^{2t-M+M'} \quad \text{(C.72)}$$

Same summation convention as for (C.21). The number of terms in $\sum_t$ is equal to $1+\tau$ where τ is the smallest of the four numbers $J \pm M$, $J \pm M'$.

With respect to the variables ξ and η, $r^{(J)}_{MM'}$ is a homogeneous polynomial of degree $2J$.

Special cases of the Wigner formula

$$r^{(J)}_{MJ} = (-)^{J-M} r^{(J)}_{JM} = r^{(J)}_{-J-M} = (-)^{J-M} r^{(J)}_{-M-J}$$
$$= \sqrt{\frac{(2J)!}{(J+M)!\,(J-M)!}}\, \xi^{J+M}\, \eta^{J-M} \qquad \text{(C.73)}$$
$$r^{(J)}_{JJ} = r^{(J)}_{-J-J} = \xi^{2J} \qquad r^{(J)}_{J-J} = (-)^{2J} r^{(J)}_{J-J} = \eta^{2J}.$$

Case $J = \frac{1}{2}$

$$R^{\frac{1}{2}}(\alpha\beta\gamma) = \begin{pmatrix} e^{-\frac{1}{2}i\alpha} \cos\tfrac{1}{2}\beta\, e^{-\frac{1}{2}i\gamma} & -e^{-\frac{1}{2}i\alpha} \sin\tfrac{1}{2}\beta\, e^{+\frac{1}{2}i\gamma} \\ e^{+\frac{1}{2}i\alpha} \sin\tfrac{1}{2}\beta\, e^{-\frac{1}{2}i\gamma} & e^{+\frac{1}{2}i\alpha} \cos\tfrac{1}{2}\beta\, e^{+\frac{1}{2}i\gamma} \end{pmatrix} \qquad \text{(C.74)}$$

(in this expression, the successive lines correspond to $M = \frac{1}{2}, -\frac{1}{2}$; the columns are arranged in the same order from left to right).

13. Integral Values of J ($J = l$) and Rotation of the Spherical Harmonics

Case $l = 1$

$$r^{(1)}(\beta) = \begin{pmatrix} \frac{1}{2}(1+\cos\beta) & -\frac{1}{2}\sqrt{2}\sin\beta & \frac{1}{2}(1-\cos\beta) \\ \frac{1}{2}\sqrt{2}\sin\beta & \cos\beta & -\frac{1}{2}\sqrt{2}\sin\beta \\ \frac{1}{2}(1-\cos\beta) & \frac{1}{2}\sqrt{2}\sin\beta & \frac{1}{2}(1+\cos\beta) \end{pmatrix}. \qquad \text{(C.75)}$$

(In this expression, the successive lines correspond to $M = +1, 0, -1$; the columns are arranged in the same order from left to right.)

Rotation of the spherical harmonics

$\omega \equiv (\theta, \varphi)$ polar angles of a unit vector $\mathbf{v}$ with respect to the coordinate system $Oxyz$ ($v_1 = \sin\theta\cos\varphi$, $v_2 = \sin\theta\sin\varphi$, $v_3 = \cos\theta$).

$\Omega \equiv (\Theta, \Phi)$ polar angles of the same vector $\mathbf{v}$ with respect to the system $OXYZ$.

$(\alpha\beta\gamma)$ Euler angles of the rotation taking $Oxyz$ into $OXYZ$ [1]).

[1]) To remove the arbitrariness in the choice of the Euler angles, we add the supplementary conditions

$$0 \leqslant \alpha < 2\pi \qquad 0 \leqslant \beta \leqslant \pi \qquad -\pi < \gamma \leqslant +\pi$$

(with this choice, $OuzZ$ is a right-handed system). Then:

(*i*) the polar coordinates of OZ with respect to $Oxyz$ are (β, α);

(*ii*) the polar coordinates of Oz with respect to $OXYZ$ are $(\beta, \pi - \gamma)$.

Θ and Φ are well-defined functions of θ and φ in which α, β, γ appear as parameters.

$\omega_1 \equiv (\theta_1, \varphi_1)$ polar angles with respect to $Oxyz$ of the vector $\mathbf{v}_1 \equiv \mathscr{R}^{-1}\mathbf{v}$, which the rotation $\mathscr{R}(\alpha\beta\gamma)$ transforms into $\mathbf{v}$. We have: $\omega_1 = \Omega$.

The spherical harmonic $Y_l^m(\omega)$ represents a certain ket vector, $|lm\rangle$, in the representation $\{\omega\}$:

$$Y_l^m(\omega) = \langle\omega|lm\rangle.$$

The rotation $\mathscr{R}(\alpha\beta\gamma)$ transforms this ket into a ket $R(\alpha\beta\gamma)|lm\rangle$ whose component in the direction $\mathbf{v}$ is equal to the component of $|lm\rangle$ in the direction $\mathbf{v}_1$:

$$Y_l^m(\Omega) \equiv \langle\omega_1|lm\rangle = \langle\omega|R|lm\rangle$$

whence the formula for the *rotation of spherical harmonics*:

$$Y_l^m(\Omega) = \sum_{m'=-l}^{+l} Y_l^{m'}(\omega)\, R^{(l)}_{m',m}(\alpha\beta\gamma). \tag{C.76}$$

Scalar product and addition theorem

Let $\mathbf{v}$, $\mathbf{v}'$ be two unit vectors, ω, ω' their respective polar angles with respect to $Oxyz$ and Ω, Ω' the polar angles of the same two vectors with respect to $OXYZ$. Relation (C.76) and the unitarity relation for $R^{(l)}$ give

$$\sum_m Y_l^m(\Omega)\, Y_l^{m*}(\Omega') = \sum_m Y_l^m(\omega)\, Y_l^{m*}(\omega'). \tag{C.77}$$

In particular, if $\mathbf{v}'$ is directed along OZ, we have the addition theorem

$$\sqrt{\frac{2l+1}{4\pi}}\, Y_l^0(\Omega) \equiv \frac{2l+1}{4\pi}\, P_l(\cos\Theta) = \sum_{m=-l}^{+l} Y_l^m(\theta, \varphi)\, Y_l^{m*}(\beta, \alpha). \tag{C.78}$$

$r^{(l)}_{mm'}$ *in terms of* $\cos\beta$ *and* $\sin\beta$

For integral l

(*i*) if $(-)^{m+m'} = +1$, $r^{(l)}_{mm'} =$ polynomial of degree l in $\cos\beta$;

(*ii*) if $(-)^{m+m'} = -1$, $r^{(l)}_{mm'} = \sin\beta \times$ (polynomial of degree $(l-1)$ in $\cos\beta$).

In particular

$$r^{(l)}_{ml}(\beta) = \sqrt{\frac{(2l)!}{(l+m)!\,(l-m)!}}\, \frac{1}{2^l}\, (1+\cos\beta)^m \sin^{l-m}\beta. \tag{C.79}$$

$R^{(l)}_{mm'}$ *when* m *or* $m'=0$

$$R^{(l)}_{m0}(\alpha\beta\gamma) = \sqrt{\frac{4\pi}{2l+1}}\, Y_l^{m*}(\beta\alpha). \tag{C.80a}$$

$$R^{(l)}_{0m}(\alpha\beta\gamma) = (-)^m \sqrt{\frac{4\pi}{2l+1}}\, Y_l^{m*}(\beta\gamma). \tag{C.80b}$$

$$R^{(l)}_{00}(\alpha\beta\gamma) = P_l(\cos\beta). \tag{C.80c}$$

V. IRREDUCIBLE TENSOR OPERATORS

14. Definition and Principal Properties

Definition

Tensor operator ≡ set of operators that transform linearly one into another under rotation.

Irreducible tensor operator:

The $(2k+1)$ operators $T_q^{(k)}$ $(q=-k, \ldots, +k)$ are by definition the *standard components of an irreducible tensor operator* of order k, $\boldsymbol{T}^{(k)}$, if they transform under rotation according to the formula

$$RT_q^{(k)}R^{-1} = \sum_{q'=-k}^{+k} T_{q'}^{(k)} R^{(k)}_{q'q}. \tag{C.81}$$

Vector operator ≡ irreducible tensor operator of order 1. If V_x, V_y, V_z are its components along $Oxyz$, its standard components are

$$V_+ = -\tfrac{1}{2}\sqrt{2}(V_x+\mathrm{i}V_y), \quad V_0 = V_z, \quad V_- = \tfrac{1}{2}\sqrt{2}(V_x-\mathrm{i}V_y).$$

Scalar operator ≡ irreducible tensor operator of order 0.

Commutation relations with $\boldsymbol{J}$

$$[J_\pm, T_q^{(k)}] = \sqrt{k(k+1)-q(q\pm1)}\; T_{q\pm1}^{(k)}. \tag{C.82a}$$

$$[J_z, T_q^{(k)}] = qT_q^{(k)}. \tag{C.82b}$$

Hermitean conjugate [1]):

$$\boldsymbol{S}^{(k)} = \boldsymbol{T}^{(k)\dagger} \quad \text{if} \quad S_q^{(k)} = (-)^q T^{(k)\dagger}_{-q}. \tag{C.83}$$

[1]) We adopt here the definition of Racah (*loc. cit.*). With this definition the spherical harmonics Y_l^m of a given order l form a Hermitean tensor operator. Note that the tensor product (defined below) of commuting, irreducible, Hermitean, tensor operators is not always Hermitean. In order for it to be so, we must modify the definition of Hermitean conjugation by replacing $(-)^q$ by $(-)^{k+q}$. With this new definition, the spherical harmonics of odd order are anti-Hermitean.

N.B. Our definition of Hermitean conjugation also applies to tensor operators of half-integral order if we let $(-)^q$ signify $\mathrm{e}^{\mathrm{i}\pi q}$.

Fundamental property (Wigner–Eckart)

$$\langle \tau JM|T_q^{(k)}|\tau' J' M'\rangle = \frac{(-)^{2k}}{\sqrt{2J+1}} \langle \tau J\|\boldsymbol{T}^{(k)}\|\tau' J'\rangle \langle J' k M' q|JM\rangle$$
$$= (-)^{J-M} \langle \tau J\|\boldsymbol{T}^{(k)}\|\tau' J'\rangle \begin{pmatrix} J & k & J' \\ -M & q & M' \end{pmatrix}. \tag{C.84}$$

$\langle \tau J\|\boldsymbol{T}^{(k)}\|\tau' J'\rangle$ is by definition the *reduced matrix element* [1]).

We have the conjugation relation (k integral):

$$\langle \tau J\|\boldsymbol{T}^{(k)}\|\tau' J'\rangle^* = (-)^{J'-J} \langle \tau' J'\|\boldsymbol{T}^{(k)\dagger}\|\tau J\rangle. \tag{C.85}$$

Special tensor operators

identity operator:

$$\langle \alpha J\|\alpha' J'\rangle = \delta_{\alpha\alpha'} \delta_{JJ'} \sqrt{2J+1}$$

total angular momentum:

$$\langle \alpha J\|\boldsymbol{J}\|\alpha' J'\rangle = \delta_{\alpha\alpha'} \delta_{JJ'} \sqrt{J(J+1)\,(2J+1)}.$$

15. Tensor Products of Irreducible Tensor Operators [2])

Definitions

Let $\boldsymbol{T}^{(k_1)}$, $\boldsymbol{U}^{(k_2)}$, be two irreducible tensor operators of order k_1, k_2 respectively. By definition:

$\boldsymbol{T}^{(k_1)} \otimes \boldsymbol{U}^{(k_2)} \equiv$ set of the $(2k_1+1)(2k_2+1)$ operators $T^{(k_1)}_{q_1} U^{(k_2)}_{q_2}$ (not necessarily linearly independent). It is a (reducible) tensor operator.

$\boldsymbol{V}^{(K)} \equiv [\boldsymbol{T}^{(k_1)} \otimes \boldsymbol{U}^{(k_2)}]_K$, *tensor product of order* K, is the irreducible tensor operator of order K, of components

$$V^{(K)}_Q = \sum_{q_1 q_2} \langle k_1 k_2 q_1 q_2|KQ\rangle\, T^{(k_1)}_{q_1}\, U^{(k_2)}_{q_2} \tag{C.86}$$

(one necessarily has $|k_1 - k_2| \leqslant K \leqslant k_1 + k_2$).

If $k_1 = k_2 = k$, we define the scalar product [3]):

$$S \equiv (\boldsymbol{T}^{(k)} \cdot \boldsymbol{U}^{(k)}) = \sum_q (-)^q\, T^{(k)}_q\, U^{(k)}_{-q}. \tag{C.87}$$

N.B. $\quad S = (-)^k \sqrt{2k+1}\; V_0{}^{(0)}.$

---

[1]) The definition adopted here is that of Racah (*loc. cit.*).

[2]) In all of the following formulas we limit ourselves to tensor operators of integral order.

[3]) With this definition, the scalar product of two vector operators $(\boldsymbol{V} \cdot \boldsymbol{W})$ is given, in agreement with the usual definition, by $V_x W_x + V_y W_y + V_z W_z$.

Expressions for the reduced matrix elements

Suppose we have a quantum system made up of two component systems, 1 and 2, of angular momentum $\boldsymbol{J}_1$ and $\boldsymbol{J}_2$ respectively ($\boldsymbol{J}=\boldsymbol{J}_1+\boldsymbol{J}_2$).

$|\tau_1 J_1 M_1\rangle$ basis vectors of system 1.

$|\tau_2 J_2 M_2\rangle$ basis vectors of system 2.

$\boldsymbol{T}^{(k_1)}$, $\boldsymbol{U}^{(k_2)}$ irreducible tensor operators, acting exclusively on the variables of systems 1 and 2 respectively.

$\boldsymbol{V}^{(K)}$ tensor product of order K according to definition (C.86).

In the standard representation $\{\tau_1\tau_2\boldsymbol{J}_1^2\boldsymbol{J}_2^2\boldsymbol{J}^2 J_z\}$ the reduced matrix elements of $\boldsymbol{V}^{(K)}$ are given by the composition formula

$$\langle\tau_1\tau_2 J_1 J_2 J\|\boldsymbol{V}^{(K)}\|\tau_1'\tau_2' J_1' J_2' J'\rangle = \sqrt{(2J+1)(2K+1)(2J'+1)} \times \begin{Bmatrix} J_1' & J_2' & J' \\ k_1 & k_2 & K \\ J_1 & J_2 & J \end{Bmatrix} \langle\tau_1 J_1\|T^{(k_1)}\|\tau_1' J_1'\rangle \langle\tau_2 J_2\|U^{(k_2)}\|\tau_2' J_2'\rangle. \tag{C.88}$$

Particular cases where the "9J" symbol reduces to a "6J" symbol

$\boldsymbol{U}=1 \qquad K=k_1=k.$

$$\langle\tau_1\tau_2 J_1 J_2 J\|\boldsymbol{T}^{(k)}\|\tau_1'\tau_2' J_1' J_2' J'\rangle = \delta_{\tau_2\tau_2'}\delta_{J_2 J_2'}\langle\tau_1 J_1\|\boldsymbol{T}^{(k)}\|\tau_1' J_1'\rangle \times (-)^{J'+J_1+J_2+k}\sqrt{(2J+1)(2J'+1)} \begin{Bmatrix} J_1 & k & J_1' \\ J' & J_2 & J \end{Bmatrix} \tag{C.89}$$

$\boldsymbol{T}=1 \qquad K=k_2=k.$

$$\langle\tau_1\tau_2 J_1 J_2 J\|\boldsymbol{U}^{(k)}\|\tau_1'\tau_2' J_1' J_2' J'\rangle = \delta_{\tau_1\tau_1'}\delta_{J_1 J_1'}\langle\tau_2 J_2\|\boldsymbol{U}^{(k)}\|\tau_2' J_2'\rangle \times (-)^{J+J_1+J_2'+k}\sqrt{(2J+1)(2J'+1)} \begin{Bmatrix} J_2 & k & J_2' \\ J' & J_1 & J \end{Bmatrix} \tag{C.90}$$

$K=0 \qquad k_1=k_2=k.$

$$\langle\tau_1\tau_2 J_1 J_2 J M|(\boldsymbol{T}^{(k)}\cdot\boldsymbol{U}^{(k)})|\tau_1'\tau_2' J_1' J_2' J' M'\rangle = \delta_{JJ'}\delta_{MM'}(-)^{J+J_2+J_1'} \begin{Bmatrix} J_1 & k & J_1' \\ J_2' & J & J_2 \end{Bmatrix} \langle\tau_1 J_1\|\boldsymbol{T}^{(k)}\|\tau_1' J_1'\rangle \langle\tau_2 J_2\|\boldsymbol{U}^{(k)}\|\tau_2' J_2'\rangle. \tag{C.91}$$

In particular, when $\boldsymbol{T}^{(k)}=\boldsymbol{J}_1$, eq. (C.89) gives:

$$\langle J_1 J_2\, J+1\|\boldsymbol{J}_1\|J_1 J_2 J\rangle = \frac{1}{2}\sqrt{\frac{[(S+1)^2-(J+1)^2][(J+1)^2-d^2]}{J+1}} \tag{C.92}$$

$$\begin{aligned}\langle J_1 J_2 J \,\|\boldsymbol{J}_1\|\, J_1 J_2 J\rangle &= \tfrac{1}{2}[d(S+1)+J(J+1)]\sqrt{\frac{2J+1}{J(J+1)}} \\ &= \tfrac{1}{2}[J_1(J_1+1)+J(J+1)-J_2(J_2+1)]\sqrt{\frac{2J+1}{J(J+1)}}\end{aligned} \tag{C.93}$$

$$\langle J_1 J_2\, J-1 \,\|\boldsymbol{J}_1\|\, J_1 J_2 J\rangle = -\frac{1}{2}\sqrt{\frac{[(S+1)^2-J^2]\,[J^2-d^2]}{J+1}} \tag{C.94}$$

with

$$S = J_1+J_2 \qquad d = J_1-J_2 \qquad (|d| \leqslant J \leqslant S).$$

APPENDIX D

ELEMENTS OF GROUP THEORY

PLAN OF THE APPENDIX

1. Introduction.

I. Generalities Concerning Groups

2. Definitions.
3. Classes of a Group.
4. Subgroups of a Group.
5. Isomorphism, Homomorphism.

II. Linear Representations of a Group

6. Definitions.
7. Operations on Representation Spaces. Reducibility.
8. Fundamental Theorems.
9. Applications to Quantum Mechanics.

III. Finite Groups

10. Generalities.
11. Representations.
12. Irreducible Components of a Representation.
13. Construction of the Irreducible Invariant Subspaces.

IV. Permutations (Group $\mathscr{S}_n$)

14. Generalities. Cycles. Classes.
15. Partitions.
16. Young "Symmetrizers". Construction of Irreducible Representations.
17. Principal Properties of the Irreducible Representations of $\mathscr{S}_n$.
18. System of n Fermions of Spin $\frac{1}{2}$.

1. Introduction

It frequently occurs that the equations of motion for systems studied in Quantum Mechanics are invariant with respect to certain groups of transformations or, more generally, that the observables dealt with have particularly simple transformation properties with respect to these groups. The methods of Group Theory allow one to sort out the consequences and simplifications that result from the existence of these symmetries.

In actual fact, with a strong dose of intuition and a certain address in the manipulation of operators, one can often exploit these symmetry properties without explicit recourse to the Theory of Groups. A good many excellent physicists prefer to leave the matter there, even if it means having to rediscover from time to time for each particular problem "well-known" results from Group Theory of which they have need. However, in some fields of Physics, like the study of polyatomic molecules or crystals, the required dose of intuition and address is too strong, and the conscientious and avowed use of Group Theory cannot be avoided. When less complicated symmetries are involved, the recourse to Group Theory, while not indispensible, often permits one better to state the problems and automatically provides certain elements of their solution.

The object of this appendix is to familiarize the reader with the Theory of Groups and to provide an introduction to more complete works on the subject [1]).

It consists of a discussion of the principal concepts and the properties most often used in Quantum Mechanics. The demonstrations have nearly all been omitted. Most of them, and in particular all those of sections I and II, are very simple.

I. GENERALITIES CONCERNING GROUPS

2. Definitions

GROUP

A set $\mathcal{G}$ of operations $a, b, c, \ldots$, forms a group if:

[1]) The reader is referred in particular to E. P. Wigner, *loc. cit.* note, p. 1055 and (in German) B. L. van de Waerden, *Die Gruppentheoretische Methode in der Quantenmechanik* (Edwards, Ann Arbor, Michigan, 1944). On continuous groups, see G. Racah, *Princeton Lecture Notes on Group Theory and Spectroscopy* (1951).

(*i*) the *product* of any two of them also belongs to the set [1] [2]:

$$\text{if } a \in \mathcal{G} \text{ and } b \in \mathcal{G}, \text{ then } ab \in \mathcal{G};$$

(*ii*) one of them, I, is the *unit element*:

$$I \in \mathcal{G} \text{ such that, for any } a \in \mathcal{G},\ Ia = aI = a;$$

(*iii*) each of them, a, has a *reciprocal* (or inverse), a^{-1}, belonging to the set:

$$\text{if } a \in \mathcal{G}, \text{ then } a^{-1} \in \mathcal{G} \text{ such that } a^{-1}a = a\,a^{-1} = I;$$

(*iv*) the product is *associative*:

$$(ab)c = a(bc).$$

FINITE GROUP

Group having a finite number N of elements. N is the *order* of the group.

Examples. The group of spatial reflections is a finite group of order 2; its two elements are the identity I and the reflection through the origin s; $s^2 = I$. The group of permutations of n objects, $\mathcal{S}_n$, is a finite group of order $n!$.

CONTINUOUS GROUP

Group having an infinite number of elements depending on one or several continuous parameters.

Examples. Group of rotations in ordinary space $\mathcal{R}_3$; group of translations in ordinary space.

ABELIAN GROUP

Group whose elements all *commute*:

$$ab = ba \text{ for any } a \text{ and } b \in \mathcal{G}$$

Examples. Spatial reflections; spatial translations; rotations about Oz.

3. Classes of a Group

CONJUGATE ELEMENTS

Two elements a and b of the group $\mathcal{G}$ are conjugate to one another

[1] By convention, $a \in \mathcal{G}$ means: the element a belongs to the set $\mathcal{G}$.

[2] The operation denoted by ab consists in first applying b, and then applying a to the result obtained. ab and ba may be different.

if there exists an element x of $\mathcal{G}$ such that

$$b = x\,a\,x^{-1}.$$

(*N.B.* x is not unique.)

If b is conjugate to a, a is conjugate to b: conjugation is a reciprocal correspondence. Moreover, if two elements b and c are conjugate to a third element a, they are conjugate to each other.

Class

The set of conjugates of a given element a of the group $\mathcal{G}$ is by definition a class of $\mathcal{G}$. The element a itself belongs to the set.

The class of elements conjugate to b and the class of elements conjugate to a are identical if b is conjugate to a, and otherwise have no common element. Each element of $\mathcal{G}$ belongs to a well-defined class: the entire group can be subdivided into classes.

If a given element of $\mathcal{G}$ commutes with all the others, it forms a class by itself. In particular, the identity I forms a class by itself.

Example. The set $\mathcal{R}(\varphi)$ of rotations having the same angle φ and differing only in the direction of the axis of rotation constitutes a class of the group $\mathcal{R}_3$; to each angle $\varphi (0 \leqslant \varphi < \pi)$ there corresponds a class of this group.

4. Subgroups of a Group

Definition

$\mathcal{H}$ is a subgroup of $\mathcal{G}$ if it is a group whose elements all belong to the group $\mathcal{G}$.

Example. The rotations about Oz form a subgroup of $\mathcal{R}_3$; the translations parallel to Oz form a subgroup of the spatial-translation group.

Cosets

If x is an element of $\mathcal{G}$ we can form from each element h of the subgroup $\mathcal{H}$ the element xh. We shall denote the set of these elements by $x\mathcal{H}$. There is a one-to-one correspondence between the elements of $\mathcal{H}$ and the elements of $x\mathcal{H}$.

There are two cases to consider:

(*a*) if $x \in \mathcal{H}$, $x\mathcal{H}$ is the subgroup $\mathcal{H}$ itself;

(*b*) if $x \notin \mathcal{H}$, $x\mathcal{H}$ is not a group, and is called the left coset of the subgroup $\mathcal{H}$.

The right coset $\mathscr{H}x$ is similarly defined. In what follows we shall consider only left cosets; right cosets obviously have the same properties.

Two cosets $x_1\mathscr{H}$, $x_2\mathscr{H}$ *either* contain the same elements *or* have no element in common, according as $x_2^{-1}x_1$ does or does not belong to $\mathscr{H}$.

Any element of $\mathscr{G}$ belongs either to $\mathscr{H}$ or to one of its cosets. The subgroup $\mathscr{H}$ and its various cosets constitute the whole group $\mathscr{G}$.

Conjugate subgroups of $\mathscr{H}$

If $\mathscr{H}$ is a subgroup of $\mathscr{G}$, and x an element of $\mathscr{G}$ but not of $\mathscr{H}$, then $x\mathscr{H}x^{-1}$ is also a subgroup of $\mathscr{G}$, called *conjugate* to $\mathscr{H}$.

(*N.B.* If $x \in \mathscr{H}$, $x\mathscr{H}x^{-1}$ is the group $\mathscr{H}$ itself.)

The conjugate subgroups of $\mathscr{H}$ are not necessarily different from $\mathscr{H}$ nor necessarily different from one another.

Invariant subgroup, factor (or quotient) group

$\mathscr{H}$ is an invariant subgroup of $\mathscr{G}$ if it is identical with all its conjugates:

$$\mathscr{H} = x\mathscr{H}x^{-1} \text{ for any } x \in \mathscr{G}.$$

Equivalent definition. A subgroup of $\mathscr{G}$ is invariant if its elements are all those of one or several classes of $\mathscr{G}$.

N.B. This second definition is very useful when we wish to find all the invariant subgroups of a given group.

If $\mathscr{H}$ is an invariant subgroup, $x\mathscr{H}$ and $y\mathscr{H}$ two of its cosets, the product of an element of $x\mathscr{H}$ by an element of $y\mathscr{H}$ belongs to the coset $yx\mathscr{H}$:

$$(y\mathscr{H})\,(x\mathscr{H}) = (yx\mathscr{H}).$$

(*N.B.* If $\mathscr{H}$ is an invariant subgroup, $x\mathscr{H} = \mathscr{H}x$.)

The set formed by the invariant subgroup and its cosets form a group having $\mathscr{H}$ for unit element: the *factor* (or quotient) *group* $\mathscr{G}/\mathscr{H}$.

Example. The group $\mathscr{A}_n$ of even permutations of n objects is an invariant subgroup of $\mathscr{S}_n$. It has one, and only one, coset, the set of odd permutations: the quotient group therefore has two elements.

Simple, semi-simple group

A group is simple if its only invariant subgroup is the identity element.

Example. Group of spatial rotations.

A group is semi-simple if its only invariant Abelian subgroup is the identity element.

Example. The group $\mathscr{S}_n$.

5. Isomorphism, Homomorphism

Isomorphism

Two groups $\mathscr{G}$ and $\hat{\mathscr{G}}$ are isomorphic when there exists a one-to-one correspondence between their elements that conserves the law of multiplication, i.e.:

(*i*) to each element g_i of $\mathscr{G}$ there corresponds one, and only one, element $\hat{g}_i$ of $\hat{\mathscr{G}}$, and conversely;

(*ii*) if $g_i g_j = g_k$, then $\hat{g}_i \hat{g}_j = \hat{g}_k$.

Examples. The symmetries of the equilateral triangle form a group isomorphic to $\mathscr{S}_3$; the symmetries of the regular tetrahedron form a group isomorphic to $\mathscr{S}_4$.

Homomorphism

If the correspondence between the elements of $\mathscr{G}$ and $\hat{\mathscr{G}}$ is not one-to-one then these two groups are homomorphic.

More precisely, $\mathscr{G}$ is homomorphic to $\hat{\mathscr{G}}$ if:

(*i*) to each element g_i of $\mathscr{G}$ there corresponds one, and only one, element $\hat{g}_i$ of $\hat{\mathscr{G}}$, and to each element of $\hat{\mathscr{G}}$ at least one (and perhaps more than one) element of $\mathscr{G}$;

(*ii*) $g_i g_j = g_k$ implies $\hat{g}_i \hat{g}_j = \hat{g}_k$.

If $\mathscr{G}$ has an invariant subgroup $\mathscr{H}$, it is homomorphic to the factor group $\mathscr{G}/\mathscr{H}$.

If $\mathscr{G}$ is homomorphic to $\hat{\mathscr{G}}$, the set $\mathscr{H}$ of elements of $\mathscr{G}$ homomorphic to the identity $\hat{I}$ form an invariant subgroup of $\mathscr{G}$, the set of elements of $\mathscr{G}$ homomorphic to a given element of $\hat{\mathscr{G}}$ different to $\hat{I}$ form a coset of this subgroup; the quotient group $\mathscr{G}/\mathscr{H}$ is isomorphic to $\hat{\mathscr{G}}$.

II. LINEAR REPRESENTATIONS OF A GROUP

6. Definitions

Groups of linear substitutions

The product of square matrices is an associative operation. If a

set of $n \times n$ matrices satisfy the axioms (*i*), (*ii*) and (*iii*) defining a group, then they form a certain group **G**.

Each matrix represents a certain linear operator G of an n-dimensional vector space $\mathscr{E}_n$, and therefore defines a linear transformation of the vectors of that space: if $|1\rangle$, $|2\rangle$, ..., $|n\rangle$ are the n basis vectors of $\mathscr{E}_n$ (this basis is not necessarily orthonormal), the transformation of each of them is given by the equation:

$$G|k\rangle = \sum_j |j\rangle G_{jk}.$$

The type of group just defined – which will be denoted in boldface roman letters – is called a *group of* (n-dimensional) *linear substitutions.*

Representation of a group

By definition, *a* (linear) *representation of a group* $\mathscr{G}$ *is a group of linear substitutions to which* $\mathscr{G}$ *is homomorphic.*

Let **G** be such a group and $\mathscr{E}$ the vector space in which its matrices operate. By definition, $\mathscr{E}$ is the *representation space*, and the number n of its dimensions *the degree* (or dimension) *of the representation.*

If $\mathscr{G}$ is isomorphic to **G**, the representation is said to be *faithful.* If not, the elements of $\mathscr{G}$ homomorphic to the matrix 1 form an invariant subgroup $\mathscr{H}$ and **G** is a faithful representation of the quotient group $\mathscr{G}/\mathscr{H}$.

Representations of degree 1

Any group has at least one representation of the first degree, *its identical representation*, in which each element of the group is represented by the number 1.

For there to be any more of them, it must have invariant subgroups whose quotient groups are Abelian; all other first degree representations of the group are the representations of these Abelian quotient groups (cf. the case of $\mathscr{S}_n$, § 17).

Unitary representation

A representation **G** is unitary if all of its matrices are unitary.

Equivalent representations

Two representations **G**, **G′** are equivalent if they have the same

number of dimensions and if each matrix $G'(g)$ of the one results from a certain linear transformation T of the matrix $G(g)$ of the other representing the same operation of the group $\mathscr{G}$:

$$G'(g) = T\,G(g)\,T^{-1} \text{ for any } g \in \mathscr{G},$$

or again

$$\mathbf{G}' = T\,\mathbf{G}\,T^{-1}.$$

We write symbolically

$$\mathbf{G}' \simeq \mathbf{G}.$$

If we identify the representation spaces $\mathscr{E}$ and $\mathscr{E}'$, the passage from $\mathbf{G}$ to the equivalent representation $\mathbf{G}'$ corresponds to taking a new set of basis vectors in that space.

Conjugate representations

Two representations $\mathbf{G}$, $\mathbf{G}^*$, whose matrices $G(g)$, $G^*(g)$ are complex conjugates one of the other, are by definition conjugate representations.

A representation $\mathbf{G}$ is self-conjugate if it is equivalent to its conjugate: $\mathbf{G} \simeq \mathbf{G}^*$.

Characters

By definition, the trace of the matrix $G(g)$ representing the operation g in the representation $\mathbf{G}$ of $\mathscr{G}$ is the character χ of g in that representation

$$\chi(g) \equiv \mathrm{Tr}\, G(g).$$

It follows from the properties of the trace that two operations belonging to the same class have the same character: *character is a function of class.*

For the same reason, *two equivalent representations have the same set of characters*:

$$\chi'(g) = \chi(g) \quad \text{if} \quad G'(g) = T\,G(g)\,T^{-1}.$$

We write symbolically:

$$\boldsymbol{\chi}' = \boldsymbol{\chi} \qquad \text{if} \qquad \mathbf{G}' \simeq \mathbf{G}.$$

N.B. If G is self-conjugate, its characters are real.

7. Operations on Representation Spaces. Reducibility [1])

Direct sum

Let $\mathbf{G}^a$, $\mathbf{G}^b$ be two representations of the same group $\mathcal{G}$; n_a, n_b their respective degrees; and $\mathscr{E}_a$, $\mathscr{E}_b$ their respective representation spaces.

If $|a1\rangle, |a2\rangle, \ldots, |an_a\rangle$ are the basis vectors of $\mathscr{E}_a$, $|b1\rangle, \ldots, |bn_b\rangle$, those of $\mathscr{E}_b$, the linear substitutions representing the operation g in these two representations are then defined by the laws of transformation of the basis vectors:

$$g[|a\varkappa\rangle] = \sum_{\lambda} |a\lambda\rangle\, G^a_{\lambda\varkappa}(g) \qquad g[|b\mu\rangle] = \sum_{\nu} |b\nu\rangle\, G^b_{\nu\mu}(g). \tag{D.1}$$

By definition, the direct sum $\mathscr{E}_a + \mathscr{E}_b$ of the spaces $\mathscr{E}_a$ and $\mathscr{E}_b$ is the space spanned by the $n_a + n_b$ vectors $|a1\rangle, \ldots, |an_a\rangle, \ldots, |bn_b\rangle$.

The matrices of this new space can be put into the form

$$M \equiv \begin{pmatrix} M_{aa} & M_{ab} \\ M_{ba} & M_{bb} \end{pmatrix}.$$

M_{aa} is an $n_a \times n_a$ matrix transforming the vectors of the space $\mathscr{E}_a$ into vectors of $\mathscr{E}_a$; M_{bb} an $n_b \times n_b$ matrix transforming the vectors of $\mathscr{E}_b$ into vectors of $\mathscr{E}_b$, M_{ab} an $n_a \times n_b$ matrix transforming the vectors of $\mathscr{E}_b$ into vectors of $\mathscr{E}_a$, and M_{ba} an $n_b \times n_a$ matrix having the inverse property. In particular, if A is a matrix of $\mathscr{E}_a$, B a matrix of $\mathscr{E}_b$, we can form by direct addition the matrix $A + B$ of the form (blockdiagonal form)

$$A + B = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}.$$

Note that:

$$\mathrm{Tr}\,(A + B) = \mathrm{Tr}\,A + \mathrm{Tr}\,B, \qquad \det(A + B) = \det A \det B. \tag{D.2}$$

The operation of direct addition of matrices conserves the identity and the law of matrix multiplication:

$$1_{(a+b)} = 1_{(a)} + 1_{(b)} \quad (A_1 + B_1)(A_2 + B_2) = A_1 A_2 + B_1 B_2.$$

---

[1]) The properties set forth in this paragraph are general properties of sets of matrices in one-to-one correspondence. They remain valid even if these sets of matrices do not form a group.

It follows that the set $\mathbf{G}^{a+b}$ of the matrices $G^a(g) + G^b(g)$ constitutes a representation of the group $\mathcal{G}$. The operation g is there represented by a linear substitution that is defined by the laws of transformation (D.1) of the $n_a + n_b$ basis vectors of $\mathcal{E}_a + \mathcal{E}_b$. For the characters of that representation we have: $\chi^{a+b}(g) = \chi^a(g) + \chi^b(g)$; in other words

$$\boldsymbol{\chi}^{a+b} = \boldsymbol{\chi}^a + \boldsymbol{\chi}^b. \tag{D.3}$$

TENSOR PRODUCT (OR KRONECKER PRODUCT OR DIRECT PRODUCT)

The operation of taking the tensor product of spaces or of matrices has already been defined (Ch. VII).

In taking the tensor product of $\mathcal{E}_a$ by $\mathcal{E}_b$, we form the $n_a n_b$-dimensional space $\mathcal{E}_a \otimes \mathcal{E}_b$ whose basis vectors are $|ab\varkappa\mu\rangle \equiv |a\varkappa\rangle |b\mu\rangle$ ($\varkappa = 1, 2, \ldots, n_a$; $\mu = 1, 2, \ldots, n_b$). The matrices $G^a(g) \otimes G^b(g)$ formed by tensor multiplication of the matrices representing g in $\mathbf{G}^a$ and $\mathbf{G}^b$ form a representation $\mathbf{G}^{ab} \equiv \mathbf{G}^a \otimes \mathbf{G}^b$ of degree $n_a n_b$ of the group $\mathcal{G}$ in which a given operation of the group is defined by the law:

$$g[|ab\,\varkappa\mu\rangle] = \sum_{\lambda\nu} |ab\,\lambda\nu\rangle\, G^a_{\lambda\varkappa}(g)\, G^b_{\nu\mu}(g). \tag{D.4}$$

The characters of this representation are given by the relations:

$$\chi^{ab}(g) = \chi^a(g)\, \chi^b(g). \tag{D.5}$$

REDUCIBILITY

An *invariant subspace* of the space $\mathcal{E}$ of a representation $\mathbf{G}$ is a subspace of $\mathcal{E}$ whose vectors are transformed linearly one into another by the matrices of $\mathbf{G}$.

The representation $\mathbf{G}$ is said to be:

(*i*) *irreducible* if $\mathcal{E}$ contains no invariant subspace (other than itself or the null space);

(*ii*) *reducible* if it is not irreducible:

$$\mathcal{E} = \mathcal{E}_1 + \mathcal{E}_2 \quad (\mathcal{E}_1, \mathcal{E}_2 \neq 0), \quad \mathcal{E}_1 \text{ invariant subspace.}$$

In the second case, if $\mathcal{E}_2$ is also invariant, $\mathbf{G}$ is said to be *decomposable*. One can then, with a suitable linear transformation, transform the basis vectors of $\mathcal{E}$ into vectors situated either in $\mathcal{E}_1$ or in $\mathcal{E}_2$. The equivalent representation obtained is the direct sum of a representation $\mathbf{G}_1$ in $\mathcal{E}_1$ and a representation $\mathbf{G}_2$ in $\mathcal{E}_2$:

$$\mathbf{G} \simeq \mathbf{G}_1 + \mathbf{G}_2.$$

$\mathbf{G}_1$ and $\mathbf{G}_2$ are called the *components* of $\mathbf{G}$.

A representation $\mathbf{G}$ is *completely reducible* if it can be put into the form of a (direct) sum of irreducible components:

$$\mathbf{G} \simeq \mathbf{G}^{(1)} + \mathbf{G}^{(2)} + \ldots \quad (\mathbf{G}^{(1)}, \mathbf{G}^{(2)}, \ldots \text{ irreducible}).$$

Any unitary representation is either irreducible or completely reducible.

The rotation matrices $R^{(j)}$ (Appendix C, section 4) relative to a given value of j form an irreducible unitary representation $\mathbf{D}^{(j)}$ of a group. Strictly speaking, $\mathbf{D}^{(j)}$ represents the rotation group $\mathscr{R}_3$ only if j is an integer.

For any j, the $\mathbf{D}^{(j)}$ are the irreducible representations of a group whose infinitesimal transformations are the same as those of $\mathscr{R}_3$ ("covering group" of $\mathscr{R}_3$), the group $\mathscr{U}_2$ of two-dimensional, unimodular, unitary, linear substitutions, group of which $\mathscr{R}_3$ is a factor group. When j is half-integral, $\mathbf{D}^{(j)}$ is a faithful representation of $\mathscr{U}_2$ and there corresponds to each element of $\mathscr{R}_3$ two matrices of $\mathbf{D}^{(j)}$ of opposite sign.

All the irreducible representations of an Abelian group are of degree 1.

Homomorphic mapping of one representation space onto another

A *linear mapping of* $\mathscr{E}_a$ *into* $\mathscr{E}_b$ is a linear correspondence in which each vector $|a\rangle$ of $\mathscr{E}_a$ corresponds to one, and only one vector $|b\rangle$ of $\mathscr{E}_b$. The correspondence is *homomorphic* if it conserves the transformation properties of the vectors in the various operations of the group, that is, if the correspondence

$$|a\rangle \to |b\rangle$$

implies that

$$g[|a\rangle] \to g[|b\rangle] \text{ for any } g \in \mathscr{G}$$

A mapping of $\mathscr{E}_a$ into $\mathscr{E}_b$ is completely determined if we know the $n_b \times n_a$ matrix defining the vector of $\mathscr{E}_b$ corresponding to each basis vector of $\mathscr{E}_a$:

$$|a\varkappa\rangle \to \sum_{\mu} |b\mu\rangle S_{\mu\varkappa}. \tag{D.6}$$

If the mapping is homomorphic, we have the matrix equation

$$S\,G^a(g) = G^b(g)\,S \quad \text{for any } g \in \mathscr{G},$$

i.e.

$$S\,\mathbf{G}^a = \mathbf{G}^b S. \tag{D.7}$$

When the ensemble of vectors $|b\rangle$ corresponding to the vectors of $\mathscr{E}_a$ span the whole of $\mathscr{E}_b$ (this supposes $n_a \geqslant n_b$), we have *a mapping*

of $\mathscr{E}_a$ *onto* $\mathscr{E}_b$ (the *whole* of E_b understood). In this case, all of the $n_b \times n_b$ determinants contained in S are different to zero.

When the correspondence is also one-to-one ($n_a = n_b$), the matrix S is non-singular: $\det S \neq 0$. In this case, the homomorphic mapping of $\mathscr{E}_a$ onto $\mathscr{E}_b$ is called an *isomorphic correspondence* and we obviously have:

$$\mathbf{G}^a \simeq \mathbf{G}^b.$$

8. Fundamental Theorems

The applications of Group Theory in quantum mechanics are based essentially on the following theorems:

SCHUR'S LEMMA

If $\mathbf{G}^a$ *and* $\mathbf{G}^b$ *are two irreducible representations of the same group, and if there exists a homomorphic mapping of the space of one into the space of the other, then the matrix* S *that defines that mapping* [S satisfies equations (D.6) and (D.7)] *has the following properties*:

a) if $\mathbf{G}^a$ *and* $\mathbf{G}^b$ *are inequivalent, necessarily* $S = 0$;

b) if $\mathbf{G}^a \simeq \mathbf{G}^b$ *either* $S = 0$, *or* $\det S \neq 0$;

c) if $\mathbf{G}^a \equiv \mathbf{G}^b$, *$S$ is a multiple of the unit matrix*: $S = c\mathbf{1}$ (c, constant).

COROLLARY

If a (square) matrix S commutes with all the matrices of an *irreducible* representation $\mathbf{G}$ of a group, then it is a multiple of the unit matrix:

$$\text{if } [S, \mathbf{G}] = 0, \text{ necessarily } S = c\mathbf{1}$$

COMPLETELY REDUCIBLE REPRESENTATIONS. UNIQUENESS THEOREMS

Suppose that we have two decompositions into *irreducible* parts of a completely reducible representation $\mathbf{G}$:

$$\mathbf{G} \simeq \mathbf{G}_1 + \mathbf{G}_2 + \ldots + \mathbf{G}_p \qquad \mathbf{G} \simeq \mathbf{G}_1' + \mathbf{G}_2' + \ldots + \mathbf{G}_{p'}'.$$

It can be shown that $p = p'$, and that there is a one-to-one correspondence between each term of the first decomposition and a term in the second which is equivalent to it. In other words:

UNIQUENESS THEOREM. *If* $\mathbf{G}$ *is completely reducible, its decomposition into irreducible parts is unique to within an equivalence.*

Henceforth, unless otherwise specified, two equivalent representations will not be thought of as different. The same irreducible representation can then figure several times in the representation $\mathbf{G}$. Denote by

$$\mathbf{G}^{(1)}, \mathbf{G}^{(2)}, \ldots, \mathbf{G}^{(j)}, \ldots$$

the sequence of *irreducible representations of the group* $\mathcal{G}$. According to the uniqueness theorem, any completely reducible representation $\mathbf{G}$ obeys the equivalence law

$$\mathbf{G} \simeq \sum_j n_j \mathbf{G}^{(j)} \tag{D.8}$$

and the sequence of integers ($\geqslant 0$) $n_1, n_2, \ldots, n_j, \ldots$ is uniquely-defined. Similarly, the set of characters of $\mathbf{G}$ obeys the equation

$$\chi = \sum_j n_j \chi^{(j)}. \tag{D.9}$$

The uniqueness theorem is completed by the two following theorems:

THEOREM II. *If* $\mathbf{G}$ *is completely reducible, so is any component* $\mathbf{G}_1$ *of* $\mathbf{G}$, *and its decomposition in irreducible parts is the sum of a certain number of irreducible components of* $\mathbf{G}$.

Thus, if $\mathbf{G} \simeq \mathbf{G}_1 + \mathbf{G}_2$, and if $\mathbf{G}$ can be decomposed according to (D.8), then

$$\mathbf{G}_1 \simeq \sum_j n_j{}^1 \mathbf{G}^{(j)} \quad \text{with} \quad n_j{}^1 \leqslant n_j \qquad (j = 1, 2, \ldots).$$

THEOREM III. *If* $\mathbf{G}$ *is completely reducible and if there exists a homomorphic mapping of its space* $\mathscr{E}$ *onto the space* $\mathscr{E}_1$ *of another representation* $\mathbf{G}_1$ *of the same group, then* $\mathbf{G}_1$ *is a component of* $\mathbf{G}$.

Theorem III applies in particular when to each basis vector $|\varkappa\rangle$ of the space $\mathscr{E}$ we can make correspond a vector $|\hat{\varkappa}\rangle$ of the space $\mathscr{E}_1$, the $|\hat{\varkappa}\rangle$ spanning the whole of $\mathscr{E}_1$ without necessarily being linearly independent, and the $|\hat{\varkappa}\rangle$ transforming linearly one into another by the same matrix formulae as the $|\varkappa\rangle$, i.e.

$$g[|\hat{\varkappa}\rangle] = \sum_\lambda |\hat{\lambda}\rangle G_{\lambda\varkappa}(g).$$

In this case, it is obvious that the correspondence $|\varkappa\rangle \to |\hat{\varkappa}\rangle$ establishes a homomorphic mapping of $\mathscr{E}$ onto $\mathscr{E}_1$.

In particular:

COROLLARY III. *If the spaces spanned by the vectors* $|a\varkappa\rangle$ ($\varkappa$ variable)

and $|b\mu\rangle$ (μ variable) *are associated respectively with the representations* $\mathbf{G}^a$ *and* $\mathbf{G}^b$, *and if the tensor product* $\mathbf{G}^{ab}$ *of these two representations is completely reducible, then the representation* $\mathbf{G}'$ *defined in the space spanned by the product vectors* $|a\varkappa\rangle\,|b\mu\rangle$ *is a component of* $\mathbf{G}^{ab}$.

N.B. The vectors $|a\varkappa\rangle\,|b\mu\rangle$ are not necessarily linearly independent. If they are, then $\mathbf{G}' \simeq \mathbf{G}^{ab}$.

9. Applications to Quantum Mechanics

The groups that appear in quantum theory are groups of transformations in state-vector space. They are nearly always unitary linear transformations and we limit the discussion to transformations of this type in what follows.

We denote the state-vector space by $\mathscr{E}$ and a set of unitary operators $G_1, G_2, \ldots$ forming a group by $\mathscr{G}$.

Let $|u\rangle$ be a vector of $\mathscr{E}$. $|u\rangle$ and the set of vectors $G_1|u\rangle, G_2|u\rangle, \ldots$, formed by applying the operators of the group to $|u\rangle$ do not necessarily span the whole of $\mathscr{E}$, but a certain subspace $\mathscr{E}_u$ of $\mathscr{E}$. $\mathscr{E}_u$ is an invariant subspace with respect to the transformations of the group, and is associated with a certain unitary (and therefore completely-reducible) representation $\mathbf{G}^u$ of the group $\mathscr{G}$. We say that the vector $|u\rangle$ transforms according to $\mathbf{G}^u$.

Similarly, if Q is a linear operator of $\mathscr{E}$, Q and the set of operators $G_1QG_1^{-1}, G_2QG_2^{-1}, \ldots$, obtained by application of the transformations of the group, span a vector space $\mathscr{E}_Q$ (the vectors of this space are operators) whose elements transform linearly one into another under the operations of the group; $\mathscr{E}_Q$ is associated with a representation $\mathbf{G}^Q$ of $\mathscr{G}$, which, in all the cases to be examined, will be either completely reducible, or else irreducible. By definition, Q transforms according to $\mathbf{G}^Q$.

By definition the operator Q is *invariant with respect to* $\mathscr{G}$ if it transforms according to the identical representation: Q then commutes with all the operators of the group. More generally, the operator Q is a *component of an irreducible tensor operator of the group* $\mathscr{G}$ if it transforms according to an irreducible representation $\mathbf{G}^{(j)}$ of $\mathscr{G}$.

Irreducible subspaces $\mathscr{E}(\tau j)$. Representation suited to the group $\mathscr{G}$

The space $\mathscr{E}$ is the direct sum of irreducible, invariant subspaces $\mathscr{E}(\tau j)$. Each of these is associated with a certain irreducible representa-

tion $\mathbf{G}^{(j)}$ of $\mathcal{G}$; the index τ distinguishes between those associated with the same irreducible representation. We denote by d_j the degree of $\mathbf{G}^{(j)}$.

Let $|\tau j\mu\rangle$ (μ variable) be d_j vectors forming a basis in $\mathcal{E}(\tau j)$. Since $\mathbf{G}^{(j)}$ is defined only to within an equivalence, the choice of this basis is completely arbitrary. It is convenient to make a choice once and for all by defining a *standard basis* for $\mathcal{E}(\tau j)$ in which each operation of the group is represented by a well-defined matrix $G^{(j)}_{\lambda\mu}$:

$$G|\tau j\mu\rangle = \sum_{\lambda=1}^{d_j} |\tau j\lambda\rangle\, G^{(j)}_{\lambda\mu}. \tag{D.10}$$

In what follows, we shall suppose that this standard choice has always been made.

The set of all the vectors $|\tau j\mu\rangle$ (τ, j, μ variable) forms a system of basis vectors for the space $\mathcal{E}$. We shall henceforth give the name of *standard representation suited to the group* $\mathcal{G}$ to the representation $\{\tau j\mu\}$. The term representation is here to be taken in its usual quantum mechanical sense. The transformation properties of ket vectors and operators under the operations of the group $\mathcal{G}$ lead to particularly simple properties of their components in this representation. These properties are summarized by the following two theorems.

COMPONENTS OF KET VECTORS AND OPERATORS IN THE $\{\tau j\mu\}$ REPRESENTATION

THEOREM A. *If the vectors* $|u1\rangle, \ldots, |uv\rangle, \ldots$ *transform linearly one into another like the basis vectors of the unitary representation* $\mathbf{G}^a$, i.e. if

$$G|uv\rangle = \sum_{\varrho} |u\varrho\rangle\, G^a_{\varrho\nu},$$

then their components $\langle \tau j\mu|uv\rangle$ *have the following properties*:

1° If $\mathbf{G}^{(j)}$ DOES NOT APPEAR *in the decomposition*

$$\mathbf{G}^a \simeq \sum_k n_k\, \mathbf{G}^{(k)} \tag{D.11}$$

of $\mathbf{G}^a$ *into irreducible parts, then necessarily* $\langle \tau j\mu|uv\rangle = 0$.

2° If $\mathbf{G}^{(j)}$ DOES APPEAR in (D.11), *and if* $\langle av|\sigma k\chi\rangle$ *is the matrix that effects this decomposition, then*

$$\langle \tau j\mu|uv\rangle = \sum_{\sigma=1}^{n_j} u^j_{\tau\sigma}\, \langle av|\sigma j\mu\rangle^*, \tag{D.12}$$

where the $u^j_{\tau\sigma}$ *are* n_j *constants independent of* μ *and* ν.

THEOREM B. *If the operators $Q_1, Q_2, \ldots, Q_\nu, \ldots$ transform linearly one into another like the basis vectors of the unitary representation* $\mathbf{G}^a$, i.e. if

$$G\,Q_\nu\,G^{-1} = \sum_\varrho Q_\varrho\,G^a_{\varrho\nu};$$

then the matrix elements $\langle \tau_1 j_1 \mu_1 | Q_\nu | \tau_2 j_2 \mu_2 \rangle$ *have the following properties*:

1° If $\mathbf{G}^{(j_1)}$ DOES NOT APPEAR *in the decomposition*

$$\mathbf{G}^a \otimes \mathbf{G}^{(j_2)} \simeq \sum_k n^{aj_1}_k\,\mathbf{G}^{(k)} \tag{D.13}$$

of the tensor product $\mathbf{G}^a \otimes \mathbf{G}^{(j_2)}$ *into irreducible parts, then necessarily*

$$\langle \tau_1 j_1 \mu_1 | Q_\nu | \tau_2 j_2 \mu_2 \rangle = 0.$$

2° If $\mathbf{G}^{(j_1)}$ DOES APPEAR, *and if* $\langle a j_2 \nu \mu_2 | \sigma k \chi \rangle$ *is the matrix that effects this decomposition, then*

$$\langle \tau_1 j_1 \mu_1 | Q_\nu | \tau_2 j_2 \mu_2 \rangle = \sum_{\sigma=1}^{n^{aj_2}_{j_1}} \langle \tau_1 j_1 \| Q \| \tau_2 j_2 \rangle_\sigma \langle a j_2 \nu \mu_2 | \sigma j_1 \mu_1 \rangle^*, \tag{D.14}$$

where the $\langle \tau_1 j_1 \| Q \| \tau_2 j_2 \rangle_\sigma$ *are* $n^{aj_2}_{j_1}$ *constants independent of* μ_1, μ_2 *and* ν.

Important remark:

The elements $\langle a\nu | \sigma k \chi \rangle$, $\langle a j_2 \nu \mu_2 | \sigma k \chi \rangle$ of the unitary matrices involved in these theorems satisfy the defining equations

$$\sum_{\nu\nu'} \langle a\nu | \sigma k \chi \rangle^*\, G^a_{\nu\nu'}\, \langle a\nu' | \sigma' k' \chi' \rangle = \delta_{\sigma\sigma'}\,\delta_{kk'}\,G^{(k)}_{\chi\chi'}, \tag{D.15a}$$

$$\sum_{\substack{\nu\nu' \\ \mu\mu'}} \langle a j \nu \mu | \sigma k \chi \rangle^*\, G^a_{\nu\nu'}\, G^{(j)}_{\mu\mu'}\, \langle a j \nu' \mu' | \sigma' k' \chi' \rangle = \delta_{\sigma\sigma'}\,\delta_{kk'}\,G^{(k)}_{\chi\chi'}. \tag{D.15b}$$

They are completely determined by giving the representations $\mathbf{G}^a$, $\mathbf{G}^a \otimes \mathbf{G}^{(j)}$. They therefore depend *only* on the way in which the vectors $|u\nu\rangle$ or the operators Q_ν, transform under the operations of the group $\mathcal{G}$.

Proof of Theorem A

Let us apply the unitary transformation $\langle a\nu | \sigma k \chi \rangle$ to the basis vectors $|u1\rangle, |u2\rangle, \ldots, |u\nu\rangle, \ldots$. We thereby obtain a new set of basis vectors for the space spanned by the latter:

$$|u \sigma k \chi\rangle = \sum_\nu |u\nu\rangle \langle a\nu | \sigma k \chi \rangle. \tag{D.16}$$

The unitarity relations give

$$|u\nu\rangle = \sum_{\sigma k\chi} |u\sigma k\chi\rangle \langle a\nu|\sigma k\chi\rangle^*. \tag{D.17}$$

It follows from the very definition of this unitary transformation that the d_k vectors $|u\sigma k\chi\rangle$ (σ, k fixed, χ variable) span a certain subspace $\mathscr{E}_u(\sigma k)$ and form a standard basis for the representation $\mathbf{G}^{(k)}$ in this subspace. Expressing them in terms of the vectors of the basis $\{\tau j\mu\}$ we have

$$|u\sigma k\chi\rangle = \sum_{\tau j\mu} |\tau j\mu\rangle \langle \tau j\mu|u\sigma k\chi\rangle.$$

The d_j-line and d_k-column matrix $\langle \tau j\mu|u\sigma k\chi\rangle$ (μ and χ variable, the rest fixed) effects a homomorphic mapping of $\mathscr{E}_u(\sigma k)$ into $\mathscr{E}(\tau j)$. Since the representations of these subspaces are either inequivalent or equal, we have, from Schur's lemma

$$\langle \tau j\mu|u\sigma k\chi\rangle = \delta_{jk}\, \delta_{\mu\chi}\, u^j_{\tau\sigma}, \tag{D.18}$$

where $u^j_{\tau\sigma}$ is a constant independent of μ. With this result, the projection of both sides of Eq. (D.17) onto $|\tau j\mu\rangle$ gives

$$\langle \tau j\mu|u\nu\rangle = \sum_{k\chi} \delta_{jk}\, \delta_{\mu\chi} \Big(\sum_{\sigma} u^j_{\tau\sigma} \langle a\nu|\sigma k\chi\rangle^*\Big).$$

Both parts of Theorem A are contained in this equation. Q.E.D.

Proof of Theorem B

Consider the vectors $Q_\nu|\tau_2 j_2\mu_2\rangle$ (ν and μ_2 variable, τ_2 and j_2 fixed). They transform linearly one into another like the basis vectors of the $\mathbf{G}^a \otimes \mathbf{G}^{(j_2)}$ representation. It does not necessarily follow that the space they span is associated with this representation, for they are not necessarily linearly independent; however, according to Corollary III, the representation in question, if not $\mathbf{G}^a \otimes \mathbf{G}^{(j_2)}$ itself, is in any case one of its components. In spite of the possibility of a linear dependence among these vectors, we can deal with them in exactly the same way as we dealt with the $|u1\rangle$, ..., $|u\nu\rangle$, ... of theorem A.

We define the vectors

$$|q\tau_2 j_2 \sigma k\chi\rangle = \sum_{\nu\mu_2} Q_\nu|\tau_2 j_2 \mu_2\rangle \langle aj_2\nu\mu_2|\sigma k\chi\rangle. \tag{D.16'}$$

From the unitarity relations for the matrix $\langle aj_2\nu\mu_2|\sigma k\chi\rangle$,

$$Q_\nu|\tau_2 j_2 \mu_2\rangle = \sum_{\sigma k\chi} |q\,\tau_2 j_2 \sigma k\chi\rangle \langle a\, j_2\, \nu\, \mu_2|\sigma k\chi\rangle^*. \tag{D.17'}$$

The d_k vectors $|q\tau_2 j_2 \sigma k\chi\rangle$ (χ variable, the rest fixed), are either all null, or form a standard basis for the representation $\mathbf{G}^{(k)}$. In the latter case, we may apply theorem A. Thus in all circumstances,

$$\langle \tau j\mu | q\,\tau_2 j_2\,\sigma k\chi\rangle = \delta_{jk}\,\delta_{\mu\chi}\,\langle \tau j \| Q \| \tau_2 j_2\rangle_\sigma,$$

where $\langle \tau j \| Q \| \tau_2 j_2\rangle_\sigma$ is a constant independent of μ. Hence the projection of both sides of (D.17′) onto $|\tau_1 j_1 \mu_1\rangle$ gives

$$\langle \tau_1 j_1 \mu_1 | Q_\nu | \tau_2 j_2 \mu_2\rangle = \sum_{k\chi} \delta_{j_1 k}\,\delta_{\mu_1\chi}\,(\sum_\sigma \langle \tau_1 j_1 \| Q \| \tau_2 j_2\rangle_\sigma\,\langle a\, j_2\, \nu\mu_2 | \sigma k\chi\rangle^*),$$

whence theorem B. Q.E.D.

Selection rule

If $|u\rangle$, Q, $|v\rangle$ *transform respectively like* $\mathbf{G}^u$, $\mathbf{G}^Q$, $\mathbf{G}^v$ *and if none of the irreducible components of* $\mathbf{G}^v$ *appear in the decomposition of* $\mathbf{G}^Q \otimes \mathbf{G}^u$, *then necessarily*

$$\langle v|Q|u\rangle = 0.$$

This rule is widely used. When $\mathbf{G}^u$ and $\mathbf{G}^v$ are irreducible representations it follows directly from 1° of theorem B, but it applies equally well when none of $\mathbf{G}^u$, $\mathbf{G}^Q$, $\mathbf{G}^v$ are irreducible. To show this, we note that $Q|u\rangle$ transforms (Corollary III) according to the representation $\mathbf{G}^Q \otimes \mathbf{G}^u$ or one of its components, and we apply theorem A to the vector $Q|u\rangle$.

A particular case: the rotation group

Theorems A and B apply notably to the rotation group.

The problem of the addition of two angular momenta is a particular application of theorem A. The $(2j_1+1)(2j_2+1)$ vectors $|\alpha j_1 j_2 m_1 m_2\rangle$ introduced in § XIII.25 transform like the basis vectors of the representation $\mathbf{D}^{(j_1)} \otimes \mathbf{D}^{(j_2)}$ formed by the tensor product of the irreducible representations $\mathbf{D}^{(j)}$, $\mathbf{D}^{(j_2)}$ of the rotation group [1]). According to the discussion of Section V of Chapter XIII, the decomposition of this representation into irreducible parts is given by:

$$\mathbf{D}^{(j_1)} \otimes \mathbf{D}^{(j_2)} \simeq \sum_{J=|j_1-j_2|}^{j_1+j_2} \mathbf{D}^{(J)} \qquad \text{(D.19)}$$

[1]) To speak here of the rotation group $\mathscr{R}_3$ is in fact an abuse of language. What we have, in fact, is the group $\mathscr{U}_2$, "covering group" of $\mathscr{R}_3$.

and the elements of the unitary matrix that realizes this decomposition are the Clebsch–Gordon coefficients $\langle j_1 j_2 m_1 m_2 | JM \rangle$. Since each irreducible representation appears at most once in (D.19), i.e.

$$(n^{j_1 j_2}_{J} = 1 \text{ for } J = j_1 + j_2, j_1 + j_2 - 1, \ldots, |j_1 - j_2|),$$

the sum on the right-hand side of (D.12) contains only one term in this case. The components of $|\alpha j_1 j_2 m_1 m_2\rangle$ in each subspace $\mathscr{E}(\tau J)$ are therefore known to within a constant (independent of M).

Similarly, the Wigner–Eckart theorem (§ XIII.32) results from the application of Theorem B to the components $T_q^{(k)}$ of a tensor operator irreducible with respect to rotations, i.e. to $(2k+1)$ operators transforming like the basis vectors of the $\mathbf{D}^{(k)}$ representation.

The matrix element $\langle \tau_1 j_1 m_1 | T_q^{(k)} | \tau_2 j_2 m_2 \rangle$ is given by expression (D.14). Since each component $\mathbf{D}^{(J)}$ of $\mathbf{D}^{(k)} \otimes \mathbf{D}^{(j_2)}$ appears only once in the decomposition of this representation into irreducible parts ($n^{kj_2}_{J} = 1$ for $J = k + j_2, \ldots, |k - j_2|$), the sum on the right-hand side of (D.14) has only one term in this case.

N.B. The definition of the reduced matrix element $\langle \tau_1 j_1 \| \mathbf{T}^{(k)} \| \tau_2 j_2 \rangle$ adopted in Chapter XIII differs by the factor $\sqrt{2j_1+1}$ from that adopted here.

Invariant observables. $\mathscr{G}$ degeneracy

If Q is an *observable invariant* with respect to $\mathscr{G}$:

$$[Q, G] = 0 \text{ for any } g \in \mathscr{G}$$

The d_j vectors $Q|\tau j \mu\rangle$ (τj fixed, μ variable) transform like the basis vectors of the $\mathbf{G}^{(J)}$ representation. Theorem B (or equally theorem A or Schur's lemma) gives in this case

$$\langle \tau j \mu | Q | \tau' j' \mu' \rangle = \delta_{jj'} \delta_{\mu\mu'} Q^{(j)}_{\tau\tau'}. \tag{D.20}$$

Q is thus represented by a particularly simple matrix in a standard representation suited to the group $\mathscr{G}$.

In such a representation, the eigenvalue problem for Q reduces to the diagonalization of the Hermitean matrices $Q^{(J)}$ whose elements $Q^{(j)}_{\tau\tau'}$ depend only on the indices τ and τ'. To each value of j we thus have corresponding a certain number of eigenvalues of Q, namely the eigenvalues $q_1^{(j)}, q_2^{(j)}, \ldots, q_i^{(j)}, \ldots$ of the matrix $Q^{(j)}$; *each non-degenerate eigenvalue of this matrix is a d_j-fold degenerate eigenvalue of Q*; each p-fold degenerate eigenvalue of $Q^{(j)}$ is a pd_j-fold degenerate eigenvalue of Q.

We note in particular the following two properties:

1° If Q is invariant with respect to $\mathscr{G}$, the subspace of each of its eigenvalues is invariant with respect to $\mathscr{G}$.

2° If Q is an observable invariant with respect to $\mathscr{G}$, defined in a *finite* space whose vectors transform into one another according to the $\mathbf{G}$ representation, and if the decomposition of $\mathbf{G}$ into irreducible parts is given by

$$\mathbf{G} \simeq \sum_k n_k \mathbf{G}^{(k)},$$

then the number of distinct eigenvalues of Q is *at most* equal to $\sum_k n_k$.

Irreducible tensor operators

If an operator Q transforms according to $\mathbf{G}^{\varrho}$, it is always possible to put it into the form of a sum in which each term transforms according to one of the irreducible components of $\mathbf{G}^{\varrho}$. The irreducible tensor operators merit particular attention.

By definition, the components of an irreducible tensor operator $\mathbf{T}^{(k)}$ of order k transform linearly one into another according to the irreducible representation $\mathbf{G}^{(k)}$. That operator therefore defines a d_k-dimensional representation space for $\mathbf{G}^{(k)}$. In particular it has d_k *standard components* $T_\chi^{(k)}$ (k fixed, χ variable) that form a standard basis in this representation space; by definition [cf. relation (D.10)]

$$G\, T_\chi^{(k)}\, G^{-1} = \sum_\varrho T_\varrho^{(k)}\, G_{\varrho\chi}^{(k)}. \tag{D.21}$$

Let

$$\mathbf{G}^{(g)} \otimes \mathbf{G}^{(h)} \simeq \sum_l n_l^{gh}\, \mathbf{G}^{(l)} \tag{D.22}$$

be the decomposition into irreducible parts of the tensor product of the irreducible representations $\mathbf{G}^{(g)}$, $\mathbf{G}^{(h)}$ and let $\langle gh\gamma\eta|\sigma l\lambda\rangle$ ($\sigma = 1, \ldots, n_l^{gh}$) be the elements of the unitary matrix effecting this decomposition. The standard representation of the components $T_\chi^{(k)}$ has, according to theorem B, the noteworthy properties:

$$\langle \tau_1 j_1 \mu_1 | T_\chi^{(k)} | \tau_2 j_2 \mu_2 \rangle = \begin{cases} 0 & \text{if } n_{j_1}^{kj_2} = 0 \\ \sum_{\sigma=1}^{n_{j_1}^{kj_2}} \langle \tau_1 j_1 \| T^{(k)} \| \tau_2 j_2 \rangle_\sigma \langle k j_2 \chi \mu_2 | \sigma j_1 \mu_1 \rangle^* & \text{if } n_{j_1}^{kj_2} \neq 0. \end{cases} \tag{D.23}$$

CONCLUSION

It results from the above considerations that we can fully exploit the transformation properties of ket vectors and operators of Quantum Mechanics in the operations of a given group $\mathscr{G}$ if we know how to:

(*i*) write down all the irreducible representations of this group (to within an equivalence) and construct the matrices of each of these representations corresponding to a standard choice for the basis vectors of that representation;

(*ii*) decompose tensor products of these representations into irreducible parts and determine the matrix which effects the decomposition of each of these products (i.e. determine the coefficients n_l^{gh} of equation (D.22) and the "Clebsch–Gordon coefficients" $\langle gh\gamma\eta|\sigma l\lambda\rangle$).

III. FINITE GROUPS

NOTATION

$\mathscr{F}$, finite group considered

N, order of the group.

L, number of classes ($L \leqslant N$).

$\mathscr{C}$, class of the group.

l_a, number of elements in the class $\mathscr{C}_a$.

$\mathbf{F}$, representation of $\mathscr{F}$.

F, matrix representing f in $\mathbf{F}$.

$\mathbf{F}^r$, regular representation.

f, element of the group.

I, identity element.

f^a, element of class $\mathscr{C}_a$.

$k_a \equiv \sum_{i=1}^{l_a} f_i^a$, sum of the elements of $\mathscr{C}_a$.

χ, set of characters of $\mathbf{F}$.

$K_a \equiv \sum_{i=1}^{l_a} F_i^a$.

$\chi^{(a)} \equiv \mathrm{Tr}\, F^a = \frac{1}{l_a} \mathrm{Tr}\, K_a$.

$\mathbf{F}^{(j)}$, jth irreducible representation (to within an equivalence).

d_j, degree of $\mathbf{F}^{(j)}$: $d_j = \chi^{(j)}(I)$.

$F^{(j)}$, unitary matrix representing f in $\mathbf{F}^{(j)}$ for a standard choice fixed once and for all for the basis vectors of that representation.

$F^{(j)}_{\alpha\beta} \equiv (f|j\alpha\beta)$ element of the matrix $F^{(j)}$ relative to line α and column β.

$\chi^j(a) \equiv (a|j)$ character of the class $\mathscr{C}_a$ in the representation $\mathbf{F}^{(j)}$.

10. Generalities

REARRANGEMENT LEMMA

If $f_1, f_2, \ldots, f_N$ are the elements of the group written in a certain order, then each element of the group appears once and only once in the sequence $f_1 f, f_2 f, \ldots, f_N f$ obtained by multiplying each of them by the same element of the group f: it is simply the elements of the group written in a different order.

(All the properties of this paragraph result from this lemma.)

SUBGROUP OF $\mathscr{F}$

If $\mathscr{H}$ is a subgroup of $\mathscr{F}$ of order N_h, N is a multiple of N_h:

$$N = hN_h \qquad (h, \text{ integer} > 0).$$

The integer h is called the *index* of the subgroup.

If $\mathscr{H}$ is an invariant subgroup, its index h is the order of the factor group $\mathscr{G}/\mathscr{H}$.

CLASSES OF $\mathscr{F}$

N is a multiple of the number of elements l_a in any one of its classes

$$N = p_a l_a \qquad (p_a, \text{ integral} > 0)$$

(the elements of $\mathscr{F}$ which commute with a given element f^a of the class $\mathscr{C}_a$ form a subgroup whose index is l_a).

THE GROUP ALGEBRA. CLASS SUM k_a

The linear combinations of the elements of the group $\sum_{s=1}^{N} x^s f_s$, where $x^1, x^2, \ldots, x^s$ are any complex numbers, form an algebra (of hypercomplex numbers): the group algebra.

The L hypercomplex numbers obtained by summing the elements of each class are called class sums:

$$k_a \equiv \sum_{i=1}^{l_a} f_i{}^a \qquad (a = 1, 2, \ldots, L). \tag{D.24}$$

These L operators commute with all the elements of the group algebra, and any other number having this property is a linear combination of these L operators.

CLASS ALGEBRA

The group algebra is not commutative unless the group is Abelian. However, the linear combinations of the L operators k_a form a commutative algebra, the class algebra, or the centre of the group algebra (if the group is Abelian, the class algebra is identical with the group algebra). One finds:

$$k_a k_b = k_b k_a = \sum_{c=1}^{L} g_{ab}^c k_c. \tag{D.25}$$

The coefficients g_{ab}^c are *integers* $\geqslant 0$.

Although linearly independent, the L class sums are related by relations (D.25) and can therefore all be expressed as functions of a limited number of them.

Each class sum obeys an algebraic equation of degree inferior or equal to L.

11. Representations

REGULAR REPRESENTATION $\mathbf{F}^r$

This is the degree-N representation obtained by taking the N elements of the group itself as basis vectors. The vectors of the representation space are therefore the elements of the group algebra.

All but one of the elements of the $N \times N$ matrix representing f in $\mathbf{F}^r$ are null. The non-vanishing element is equal to 1.

The principal properties of the representations of the group nearly all follow from the rearrangement lemma, Schur's lemma and the study of the regular representation.

PROPERTIES OF REPRESENTATIONS IN GENERAL

Any representation of a finite group is equivalent to a *unitary* representation of that group.

If two representations have the same set of *characters*, they are equivalent [1]) (the converse is evident).

If $F_1, F_2, \ldots, F_N$ is a set of linear operators forming a finite group $\mathscr{F}$, and if $|u\rangle$ is a given vector in ket space, the representation $\mathbf{F}^u$

[1]) Consequence of the uniqueness of the decomposition into irreducible parts and of the orthogonality of the characters of the irreducible representations (see below). This property remains when the group is infinite provided that the two representations are completely reducible.

according to which $|u\rangle$ transforms in the operations of the group, is a *component of the regular representation* $\mathbf{F}^{\mathbf{r}}$ (if $\mathbf{F}^{\mathbf{u}}$ is of degree N, $\mathbf{F}^{\mathbf{u}} \simeq \mathbf{F}^{\mathbf{r}}$).

In a given representation $\mathbf{F}$, $k_1, k_2, \ldots, k_L$ are represented by matrices $K_1, K_2, \ldots, K_L$ that can be *simultaneously diagonalized*, and that commute with any matrix representing an element of the group:

$$[K_a, F] = 0. \tag{D.26}$$

IRREDUCIBLE REPRESENTATIONS

a) Number. The number of inequivalent, irreducible representations is equal to the number of classes, L.

b) Degree. If d_j is the degree of the jth irreducible representation $\mathbf{F}^{(j)}$ $(j = 1, 2, \ldots, L)$

$$N/d_j \text{ is integral} \tag{D.27}$$

$$\sum_{j=1}^{L} (d_j)^2 = N. \tag{D.28}$$

c) Orthogonality relations. If the unitary, irreducible representations $\mathbf{F}^{(j)}$, $\mathbf{F}^{(k)}$ are either inequivalent or equal [1]),

$$\boxed{\frac{d_j}{N} \sum_{f=1}^{N} (f|j\,\alpha\,\beta)\,(f|k\,\gamma\,\delta)^* = \delta_{jk}\,\delta_{\alpha\gamma}\,\delta_{\beta\delta}} \tag{D.29}$$

From (D.29) we obtain the orthogonality relation for characters

$$\boxed{\sum_{a=1}^{L} \frac{l_a}{N} (a|j)\,(a|k)^* = \delta_{jk}} \tag{D.30}$$

Each representation being unambiguously defined by a standard choice of its basis vectors, the N^2 quantities

$$\sqrt{\frac{d_j}{N}}\, F^{(j)}_{\alpha\beta} \equiv \sqrt{\frac{d_j}{N}}\, (f|j\,\alpha\,\beta)$$

$$(f = 1, 2, \ldots, N;\ j = 1, 2, \ldots, L;\ \alpha, \beta = 1, 2, \ldots, d_j)$$

[1]) Demonstration: if S is a $d_j \times d_j$ matrix, the matrix $T \equiv \sum_{f=1}^{N} F^{(j)} S F^{(k)-1}$ obeys the relation $F^{(j)} T = T F^{(k)}$ for all f (rearrangement lemma). Thus (Schur's lemma), T is either identically null or a multiple of unity according as $\mathbf{F}^{(j)}$ and $\mathbf{F}^{(k)}$ are inequivalent or equal. One obtains relations (D.29) by fixing S in an appropriate fashion.

are the elements of an $N \times N$ unitary matrix. Similarly, the L^2 quantities $\sqrt{l_a/N}\chi^j(a) \equiv \sqrt{l_a/N}(a|j)$ are the elements of a $L \times L$ unitary matrix. From the two unitarity relations (D.29), (D.30) we respectively deduce the two following:

$$\sum_{j=1}^{L} \sum_{\alpha,\beta=1}^{d_j} \frac{d_j}{N} (f|j\alpha\beta)\,(g|j\alpha\beta)^* = \delta_{fg} \tag{D.31}$$

$$\frac{l_a}{N} \sum_{j=1}^{L} (a|j)\,(b|j)^* = \delta_{ab}. \tag{D.32}$$

d) *Special cases.* If $\mathbf{F}^{(k)}$ is the identical representation, $(k=1)$ (D.29) and (D.30) give

$$\sum_{f=1}^{N} (f|j\alpha\beta) = N\,\delta_{j1} \qquad \sum_{a=1}^{L} l_a(a|j) = N\,\delta_{j1}. \tag{D.33a}$$

If g is the unit element $(g=I)$, (D.31) and (D.32) give

$$\sum_{j\alpha\beta} d_j(f|j\alpha\beta) = N\,\delta_{fI} \qquad \sum_{j} d_j(a|j) = N\,\delta_{aI}. \tag{D.33b}$$

e) *Relation between class sums and characters.* The matrices $K_a{}^{(j)}$ representing the k_a are *multiples of the unit matrix* (Schur's lemma):

$$K_a{}^{(j)} = k_a{}^{(j)} 1^{(j)}. \tag{D.34}$$

$$k_a{}^{(j)} = \frac{1}{d_j} \operatorname{Tr} K_a{}^{(j)} = \frac{l_a}{d_j} (a|j). \tag{D.35}$$

From (D.25) and (D.35),

$$l_a\, l_b (a|j)\,(b|j) = d_j \sum_{c=1}^{L} g_{ab}^c\, l_c (c|j) \tag{D.36}$$

from which, due to the orthogonality relation (D.32),

$$g_{ab}^c = \frac{l_a\, l_b}{N} \sum_{j=1}^{L} \frac{1}{d_j} (a|j)\,(b|j)\,(c|j)^*. \tag{D.37}$$

12. Irreducible Components of a Representation

GENERAL METHOD

In order to obtain the coefficients n_j of the decomposition

$$\mathbf{F} \simeq \sum_{j=1}^{L} n_j \mathbf{F}^{(j)} \tag{D.38}$$

of a representation $\mathbf{F}$ into irreducible components, we need only to know its set of characters $\chi(a)$ and the sets of characters of the L irreducible representations of the group. According to eq. (D.9) and relations (D.30),

$$n_j = \frac{1}{N} \sum_{a=1}^{L} l_a (a|j)^* \, \chi(a). \tag{D.39}$$

N.B. We have:

$$p \equiv \frac{1}{N} \sum_{a=1}^{L} l_a |\chi(a)|^2 = \sum_{j=1}^{L} n_j{}^2.$$

Thus $p=1$ is a criterion for the irreducibility of $\mathbf{F}$.

REGULAR REPRESENTATION ($\chi^r(a) = N\delta_{a1}$)

$$\mathbf{F}^r \simeq \sum_{j=1}^{L} d_j \, \mathbf{F}^{(j)}. \tag{D.40}$$

The regular representation contains each irreducible representation of the group a number of times equal to the degree of that representation [relation (D.28) follows].

TENSOR PRODUCT OF IRREDUCIBLE REPRESENTATIONS

$$\mathbf{F}^{(g)} \otimes \mathbf{F}^{(h)} \simeq \sum_{j=1}^{L} n_j{}^{gh} \, \mathbf{F}^{(j)}. \tag{D.41}$$

Equation (D.9) reads in this case

$$(a|g) \, (a|h) = \sum_{j=1}^{L} n_j{}^{gh} (a|j) \tag{D.42}$$

and (D.39) gives

$$n_j{}^{gh} = \frac{1}{N} \sum_{a=1}^{L} l_a (a|g) \, (a|h) \, (a|j)^*. \tag{D.43}$$

These two relations are to be compared with (D.36) and (D.37).

Lemma:

$\mathbf{F}^{(j)}$ is contained as many times in $\mathbf{F}^{(g)} \otimes \mathbf{F}^{(h)}$ as $\mathbf{F}^{(g)}$ in $\mathbf{F}^{(j)} \otimes \mathbf{F}^{(h)*}$.

(*N.B.* If the irreducible representations are all self-conjugate, $n_j{}^{gh}$ is symmetrical with respect to its three indices.)

COMPONENTS OF DEGREE 1

The tensor product $F^{(g)} \otimes F^{(h)}$ of two irreducible representations contains at most one component of degree 1; a necessary and sufficient condition for the first-degree representation $\mathbf{F}_1$ to figure in this product is that

$$\mathbf{F}^{(g)} \simeq \mathbf{F}_1 \otimes \mathbf{F}^{(h)*}. \tag{D.44}$$

In particular, the representation space $\mathscr{E}_g \otimes \mathscr{E}_h$ of this product contains *at most one* vector invariant under the transformations of the group; this vector exists if, and only if,

$$\mathbf{F}^{(g)} \simeq \mathbf{F}^{(h)*}. \tag{D.45}$$

13. Construction of the Irreducible Invariant Subspaces

We shall henceforth suppose that a standard choice of the basis vectors has been made for each irreducible representation $\mathbf{F}^{(j)}$. We then say that a vector is of the type $(j\mu)$ if it transforms like the μth vector of a standard basis of $\mathbf{F}^{(j)}$. The unitary matrices $F^{(j)}$ being thereby unambiguously defined, we now wish to construct the vectors of a standard basis suited to the group $\mathscr{F}$ in the space $\mathscr{E}$ of the representation $\mathbf{F}$ introduced at the beginning of § 12. We limit ourselves to the case where $\mathscr{E}$ is generated by application of the operators $F_1, F_2, \ldots, F_N$ of the group to a given vector $|\rangle$ in ket-space ($n_j \leqslant d_j$). In the applications of Group Theory to Quantum Mechanics we can always reduce the problems to this special case.

BASIS OPERATORS $B^{(j)}_{\mu\nu}$ OF THE REGULAR REPRESENTATION

Let us introduce the N operators

$$B^{(j)}_{\mu\nu} \equiv \frac{d_j}{N} \sum_{f=1}^{N} (f|j\,\mu\,\nu)^*\, F \tag{D.46}$$

$$(j = 1, 2, \ldots, L; \quad \mu, \nu = 1, 2, \ldots, d_j).$$

It follows from the orthogonality relations (D.31) that the N operators of the group are linear combinations of these operators:

$$F = \sum_{j=1}^{L} \sum_{\mu,\nu=1}^{d_j} (f|j\,\mu\,\nu)\, B^{(j)}_{\mu\nu}. \tag{D.47}$$

From the unitarity of the F, the rearrangement lemma and relations (D.29), we have the fundamental properties

$$B^{(j)\dagger}_{\mu\nu} = B^{(j)}_{\nu\mu} \tag{D.48}$$

$$F\, B^{(j)}_{\mu\nu} = \sum_{\varkappa} B^{(j)}_{\varkappa\nu}(f|j\varkappa\mu) \tag{D.49}$$

$$B^{(j)}_{\mu\nu}\, B^{(k)}_{\varrho\sigma} = \delta_{jk}\, \delta_{\nu\varrho}\, B^{(j)}_{\mu\sigma}. \tag{D.50}$$

According to (D.49), the N elements of the group algebra that are represented by the $B^{(j)}_{\mu\nu}$ form a standard basis in the regular representation.

CONSTRUCTION OF A STANDARD BASIS WITH THE AID OF THE $B^{(j)}_{\mu\nu}$

If one knows how to construct the operators $B^{(j)}_{\mu\nu}$ – in other words, if one knows the N matrix elements $(f|j\mu\nu)$ – the problem of constructing a standard basis in $\mathscr{E}$ is practically solved.

Indeed

a) If not null, $B^{(j)}_{\mu\nu}|\rangle$ is a vector of type $(j\mu)$ [eq. (D.49)].

b) The N vectors $B^{(j)}_{\mu\nu}|\rangle$ span the whole of the space [eq. (D.47)].

c) The $d_j{}^2$ vectors $B^{(j)}_{\mu\nu}|\rangle$ corresponding to the same value of j have the following properties:

(*i*) the d_j vectors corresponding to the same value of

$$\nu\ (\mu = 1, 2, \ldots, d_j)$$

have the same norm, and form, except for the normalization a standard basis for the representation $\mathbf{F}^{(j)}$;

(*ii*) the d_j vectors corresponding to the same value of

$$\mu\ (\nu = 1, 2, \ldots, d_j)$$

span the n_j-dimensional space $\mathscr{E}_{j\mu}$ of the vectors of type $(j\mu)$; they are related by $(n_j - d_j)$ linear relations whose coefficients are independent of μ (in particular, if $n_j = 1$, these d_j vectors are multiples of each other and the multiplication factors are independent of μ).

Thus to construct a standard basis in $\mathscr{E}$ suited to the group, we need only to choose for each value of j a definite value $\bar{\mu}$ of μ and to construct, by the Schmidt orthogonalization method for example, a set of n_j basis vectors in the space $\mathscr{E}_{j\bar{\mu}}$:

$$|\tau j\bar{\mu}\rangle = \sum_{\nu=1}^{d_j} c^{\tau j}_{\nu}\, B^{(j)}_{\bar{\mu}\nu}|\rangle \qquad (\tau = 1, 2, \ldots, n_j)$$

$$\langle \tau j\bar{\mu}|\tau' j\bar{\mu}\rangle = \delta_{\tau\tau'}.$$

Then the vectors

$$|\tau j \mu\rangle \equiv \sum_{\nu=1}^{d_j} c_\nu^{\tau j} B_{\mu\nu}^{(j)} |\rangle$$

$$(j = 1, 2, \ldots, L; \quad \mu = 1, 2, \ldots, d_j; \quad \tau = 1, 2, \ldots, n_j)$$

form the sought-for standard basis, the d_j vectors $|\tau j\mu\rangle$ (τ and j fixed, $\mu = 1, \ldots, d_j$) forming a standard basis for the representation $\mathbf{F}^{(j)}$.

OTHER PROPERTIES OF THE $B_{\mu\nu}^{(j)}$. PROJECTORS ONTO THE SPACES $\mathscr{E}_{j\mu}$ AND TRANSFER FROM ONE TO ANOTHER

Let us write

$$\Pi_\mu^{(j)} \equiv B_{\mu\mu}^{(j)} \equiv \frac{d_j}{N} \sum_{f=1}^{N} (f|j\mu\mu)^* F. \tag{D.51}$$

The operators $\Pi_\mu^{(j)}$ ($j = 1, 2, \ldots, L$; $\mu = 1, 2, \ldots, d_j$) form a set of orthogonal projectors whose sum is equal to 1 [1]:

$$\Pi_\mu^{(j)\dagger} = \Pi_\mu^{(j)} \tag{D.52}$$

$$\Pi_\mu^{(j)} \Pi_\varrho^{(k)} = \delta_{jk} \delta_{\mu\varrho} \Pi_\mu^{(j)} \tag{D.53}$$

$$\sum_{j=1}^{L} \sum_{\mu=1}^{d_j} \Pi_\mu^{(j)} = 1 \tag{D.54}$$

$\Pi_\mu^{(j)}$ is the projector onto the space $\mathscr{E}_{j\mu}$ of the vectors of type $(j\mu)$ [eq. (D.49)]. *The decomposition of unity* (D.54) permits one to express any vector of $\mathscr{E}$ in the form of a sum of vectors each belonging to a particular $\mathscr{E}_{j\mu}$

$$|u\rangle \equiv \sum_{j\mu} \Pi_\mu^{(j)} |u\rangle.$$

If $\mu \neq \nu$, the operator $B_{\mu\nu}^{(j)}$ is the *transfer operator* from the subspace $\mathscr{E}_{j\nu}$ to the subspace $\mathscr{E}_{j\mu}$. The reason for this name is easily seen; from (D.50) and (D.48)

$$\Pi_\gamma^{(g)} B_{\mu\nu}^{(j)} \Pi_\eta^{(h)} = \delta_{gj} \delta_{jh} \delta_{\gamma\mu} \delta_{\nu\eta} B_{\mu\nu}^{(j)} \tag{D.55}$$

$$B_{\mu\nu}^{(j)\dagger} B_{\mu\nu}^{(j)} = \Pi_\nu^{(j)}. \tag{D.56}$$

---

[1]) This is true whatever $\mathbf{F}$, even if this representation is not a component of $\mathbf{F}^r$.

Thus $B^{(j)}_{\mu\nu}$ acting on any vector orthogonal to $\mathscr{E}_{j\nu}$ gives zero, and acting on any vector of $\mathscr{E}_{j\nu}$ transforms it into a vector of $\mathscr{E}_{j\mu}$, the correspondence thereby established between these two spaces being one-to-one and conserving the scalar product.

If $|\sigma j\nu\rangle$ is a vector of type $(j\nu)$, the d_j vectors

$$B^{(j)}_{\mu\nu}|\sigma j\nu\rangle \qquad (\mu = 1, 2, \ldots, d_j)$$

form a standard basis for the $\mathbf{F}^{(j)}$ representation.

USE OF THE CLASS SUMS K_a. PROJECTOR $P^{(j)}$

If we merely wish to find the irreducible invariant subspaces of $\mathscr{E}$ we do not need to know the N^2 standard matrix elements $F^{(j)}_{\mu\nu} \equiv (f|j\mu\nu)$.

The class sums

$$K_a = \sum_{i=1}^{l_a} F_i{}^a \qquad (a = 1, 2, \ldots, L)$$

have at least one common set of basis vectors. To each vector of this set there corresponds a certain sequence $k \equiv (k_1, k_2, \ldots, k_L)$ of eigenvalues of these operators. There exists in all L series of possible eigenvalues $k^{(1)}, k^{(2)}, \ldots, k^{(L)}$, defined by equation (D.35) and corresponding each to a definite irreducible representation of the group. Consequently, if we simultaneously diagonalize $K_1, K_2, \ldots, K_L$, each series of eigenvalues $k^{(j)}$ is $n_j d_j$-fold degenerate and the corresponding subspace $\mathscr{E}_j$ defines the component $n_j\mathbf{F}^{(j)}$ of $\mathbf{F}$. If the n_j are all equal to 1, the decomposition of $\mathscr{E}$ into irreducible invariant subspaces is thereby automatically realized. If not, it remains to make this decomposition in each of the subspaces $\mathscr{E}_j$ for which $n_j > 1$.

Recall that the L operators K are all functions of a limited number of them; we therefore need only to diagonalize these to diagonalize them all.

The problem of diagonalizing the K operators is practically solved if we know the set of characters for all the irreducible representations of the group (tables of characters exist for most of the groups used in physics; cf. references of note, p. 1079). Indeed, the projector $P^{(j)}$ onto the space $\mathscr{E}_j$ is

$$\begin{aligned} P^{(j)} &\equiv \sum_{\mu=1}^{d_j} \Pi_\mu{}^{(j)} = \frac{d_j}{N} \sum_{f=1}^{N} \chi^{(j)*}(f)\, F \\ &= \frac{d_j}{N} \sum_{a=1}^{L} (a|j)^*\, K_a. \end{aligned} \tag{D.57}$$

IV. PERMUTATIONS [1])

(Group $\mathscr{S}_n$)

14. Generalities. Cycles. Classes

DEFINITION

Suppose we have n objects distributed among n "boxes" (ex.: n particles in n quantum states). A permutation of the n objects is a modification of their distribution among these n "boxes". We can label the objects with the integers from 1 to n, and define a particular permutation by the symbol

$$p = \begin{pmatrix} \alpha_1 \alpha_2 \dots \alpha_n \\ \beta_1 \beta_2 \dots \beta_n \end{pmatrix},$$

where $\alpha_1, \alpha_2, \dots, \alpha_n$ are the n integers written in an arbitrary order and $\beta_1, \beta_2, \dots, \beta_n$ the same integers written in an order such that object β_i occupies in the new distribution the place that α_i occupied in the old. Thus in the permutation of 5 objects:

$$p_a = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 3 & 2 & 1 & 4 \end{pmatrix}$$

object 5 takes the place of 1, 3 that of 2, etc. The significance of the symbol is clearly unaltered by a change in the order of its columns.

The successive application of two permutations p_a, p_b is equivalent to the single permutation $p_c \equiv p_b p_a$. The latter can easily be written down if the upper line of the symbol p_b is identical with the lower line of the symbol p_a. If, for example, p_a is the permutation defined above, and if

$$p_b = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 3 & 5 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 5 & 3 & 2 & 1 & 4 \\ 2 & 5 & 3 & 4 & 1 \end{pmatrix},$$

$$p_c \equiv p_b p_a = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 3 & 4 & 1 \end{pmatrix}.$$

---

[1]) A simple and complete discussion of the theory of Young tableaux and its application to the permutation group is given in the book by D. E. Rutherford, *Substitutional Analysis* (University Press, Edinburg, 1948). The essential of the theory of the group $\mathscr{S}_n$ is given by B. L. van de Waerden, *Modern Algebra* (Ungar, Ed., New York, 1950), vol. II, p. 190–193.

In particular, the inverse of p is the permutation whose symbol results from that of p by interchange of the two lines:

$$p^{-1} = \begin{pmatrix} \beta_1 \beta_2 \dots \beta_n \\ \alpha_1 \alpha_2 \dots \alpha_n \end{pmatrix}.$$

The permutations of n objects form a group of order $n!$.

Circular permutations. Cyclic notation

The permutation

$$\begin{pmatrix} \alpha_1 \alpha_2 \dots \alpha_{k-1} \alpha_k \alpha_{k+1} \dots \alpha_n \\ \alpha_2 \alpha_3 \dots \alpha_k \quad \alpha_1 \alpha_{k+1} \dots \alpha_n \end{pmatrix},$$

in which α_2 takes the place of α_1, α_3 that of α_2, ..., α_k that of α_{k-1}, α_1 that of α_k and the $(n-k)$ remaining objects α_{k+1}, ..., α_n keep their place, is by definition a *circular permutation, or cycle, of the k objects* $\alpha_1, \alpha_2, \dots, \alpha_k$; k is the *length of the cycle*. A permutation of this type can be represented by the symbol

$$p = (\alpha_1 \alpha_2 \dots \alpha_k). \tag{D.58}$$

By convention, the first object in the list, α_1, takes the place of the last, α_k, and each of the others takes the place immediately on its left. With this notation the order of the k elements is defined only to within a circular permutation.

Two cycles having no common element commute.

Any permutation of n objects is a product of commuting cycles (these cycles have no common element) *and this decomposition is unique.*

Thus the permutation p_a defined above is the product of (154) with (23) and is written *in cyclic notation*:

$$p_a = (154)\,(23) = (23)\,(154).$$

Similarly,

$$p_b = (14)\,(235) \qquad p_c = (125)\,(3)\,(4).$$

The cycles of length 1 are equivalent to the identity transformation and may be omitted; we write simply: $p_c = (125)$. If this omission is not made, the sum of the lengths of the cycles of a permutation is equal to n.

A permutation is completely defined by:

(*i*) *its cycle structure*, that is, the number of its cycles, $h\,(\leqslant n)$, and their respective lengths $\lambda_1, \lambda_2, \dots, \lambda_h$ $(\lambda_1 + \lambda_2 + \dots + \lambda_h = n)$;

(*ii*) the numbers that appear in each cycle and their order to within a circular permutation.

If we reverse the order of the numbers in each cycle we get the inverse permutation. Thus

$$p_a^{-1} = (451)\,(32).$$

CLASSES

Two permutations with the same cycle structure belong to the same class, and conversely.

The cyclic notation for the conjugate element to p, $p' = xpx^{-1}$, is derived by application of the permutation x to the sequence of n numbers figuring in the cyclic notation for p. Example:

$$p_a = \begin{pmatrix} 1 & 5 & 4 & 2 & 3 \\ 5 & 4 & 1 & 3 & 2 \end{pmatrix} = (154)\,(23)$$

$$x = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 1 & 3 & 5 \end{pmatrix} = \begin{pmatrix} 5 & 4 & 1 & 3 & 2 \\ 5 & 3 & 2 & 1 & 4 \end{pmatrix} \qquad x^{-1} = \begin{pmatrix} 2 & 5 & 3 & 4 & 1 \\ 1 & 5 & 4 & 2 & 3 \end{pmatrix}$$

$$p' \equiv x\,p_a\,x^{-1} = \begin{pmatrix} 2 & 5 & 3 & 4 & 1 \\ 5 & 3 & 2 & 1 & 4 \end{pmatrix} = (253)\,(41).$$

TRANSPOSITIONS

A transposition is a permutation of two objects (cycle of length 2). The transpositions form a class of $\mathscr{S}_n$.

Any cycle of a given length k is equal to the product of $k-1$ transpositions:

$$(\alpha_1\alpha_2 \ldots \alpha_k) = (\alpha_1\alpha_2)\,(\alpha_2\alpha_3) \ldots (\alpha_{k-1}\alpha_k).$$

More generally, any permutation p can be written as a product of transpositions. Such a decomposition is not unique but the number of transpositions involved will have a definite parity, either even or odd, which we shall denote by $(-)^p$. By definition, a permutation is *even* or *odd* according as $(-)^p = +1$ or -1.

SUBGROUPS OF $\mathscr{S}_n$. THE GROUP $\mathscr{A}_n$

$\mathscr{S}_n$ has one, and only one *invariant subgroup*, the group of even permutations $\mathscr{A}_n$. The index of $\mathscr{A}_n$ is equal to 2, its coset is the ensemble of odd permutations and the factor group $\mathscr{S}_n/\mathscr{A}_n$ is Abelian.

Among the other subgroups of $\mathscr{S}_n$ are the groups $\mathscr{S}_m$ of permuta-

tions of m objects where $m<n$, the groups $\mathscr{A}_m(m<n)$, etc. The index of $\mathscr{S}_m$ is $(n!/m!)$ and the index of $\mathscr{A}_m$ $2(n!/m!)$.

Symmetrizers and antisymmetrizers of $\mathscr{S}_n$

Two linear combinations of the operations of $\mathscr{S}_n$ play a special rôle, the symmetrizer s and the antisymmetrizer a:

$$s \equiv \frac{1}{n!}\sum_p p \qquad a \equiv \frac{1}{n!}\sum_p (-)^p p \tag{D.59}$$

(the sums extend over all $n!$ elements of the group $\mathscr{S}_n$). s and a are particular linear combinations of the class sums; they commute with all the operations of the group and also have the following special properties:

$$qs = sq = s \qquad qa = aq = (-)^q a \qquad (q \in \mathscr{S}_n) \tag{D.60}$$

$$s^2 = s \qquad a^2 = a. \tag{D.61}$$

15. Partitions

Definition

A *partition* $\lambda = [\lambda_1\lambda_2 \ldots \lambda_h]$ of the integer n is a sequence of positive integers, arranged in the order $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_h$, whose sum is equal to n:

$$\lambda_1 + \lambda_2 + \ldots + \lambda_h = n.$$

Since the cycle structure of a permutation is defined by a definite partition of n, *each partition of n defines a class of $\mathscr{S}_n$.*

Inequalities

Let $\lambda \equiv [\lambda_1 \ldots \lambda_h]$, $\mu \equiv [\mu_1 \ldots \mu_k]$ be two partitions of n. By definition:

$\lambda = \mu$ if $h = k$ and if $\lambda_1 = \mu_1, \lambda_2 = \mu_2, \ldots, \lambda_k = \mu_k$;

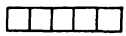
$\lambda > \mu$ if the first nonzero term in the sequence $\lambda_1 - \mu_1, \lambda_2 - \mu_2, \ldots,$ is positive;

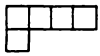
$\lambda < \mu$ if the first nonzero term in the sequence $\lambda_1 - \mu_1, \lambda_2 - \mu_2, \ldots,$ is negative.

Example: for $n = 5$ $[5] > [41] > [32] > [31^2]$, etc.

(By convention, $[31^2]$ is an abbreviated notation for $[311]$.)

YOUNG DIAGRAMS Y_λ


[5]


[4 1]


[3 2]


[3 1²]


[2² 1]


[2 1³]


[1⁵]

Fig. D.1. The partitions of 5 and their Young diagrams.

A given partition $[\lambda_1 \lambda_2 \dots \lambda_h]$ may be represented by a diagram Y_λ, called a Young diagram, made up of n squares arranged in h lines placed one above the other, the first line having λ_1 squares, the second λ_2 squares, ..., the hth λ_h squares (cf. Fig. D.1).

YOUNG TABLEAUX $\Theta_\lambda{}^p$

With the first n integers we can number the n squares of the diagram Y_λ in $n!$ ways, obtaining each time a certain *Young tableau*.

We shall denote by Θ_λ, and give the name of *normal tableau* to, the tableau in which the numbers 1, 2, ..., are arranged in their normal order: the sequence 1, 2, ..., λ_1 in the first row, λ_1+1, λ_1+2, ..., $\lambda_1+\lambda_2$ in the second, etc. By applying the permutation p



$\Theta_{[3\,2]}$



$\tilde{\Theta}_{[3\,2]}$



$\Theta_{[\widetilde{3\,2}]} \equiv \Theta_{[2^2\,1]}$



$\Theta^{(1\,4\,5)}_{[3\,2]}$



$\tilde{\Theta}^{(1\,4\,5)}_{[3\,2]}$

Fig. D.2. Several tableaux corresponding to the partitions $\lambda \equiv [3\,2]$ and $\tilde{\lambda} \equiv [2^2\,1]$ of the number 5.

to the n numbers of Θ_λ we obtain a new tableau $\Theta_\lambda{}^p \equiv p\Theta_\lambda$. Note that $q\Theta_\lambda{}^p \equiv \Theta_\lambda{}^{qp}$ and that there are $n!$ different tableaux for each Young diagram (cf. Fig. D.2).

ASSOCIATED PARTITIONS

Two partitions are associates one of the other if the Young diagram corresponding to one results from interchanging the rows and the

columns of that of the other (reflection in the principal diagonal). In what follows we denote this type of correspondence with the symbol $\sim$. Thus $\tilde{\lambda} \equiv [\tilde{\lambda}_1 \dots \tilde{\lambda}_k]$ denotes the partition associated with $\lambda \equiv [\lambda_1 \dots \lambda_h]$. Note that $k = \lambda_1$ and $h = \tilde{\lambda}_1$, and that $\tilde{\lambda}_s$ is the number of elements of the partition λ equal or greater than s and *vice versa*.

We similarly define the Young diagram, $\tilde{Y}_\lambda \equiv Y_{\tilde{\lambda}}$ associated with Y_λ and the tableau $\tilde{\Theta}_\lambda{}^p$ associated with $\Theta_\lambda{}^p$. Note that (cf. Fig. D.2)

$$\tilde{\Theta}_\lambda{}^p = p\tilde{\Theta}_\lambda$$

but that in general

$$\tilde{\Theta}_\lambda \neq \Theta_{\tilde{\lambda}}$$

16. Young "Symmetrizers". Construction of Irreducible Representations

"SYMMETRIZERS" s_λ, a_λ, $\tilde{s}_\lambda$, $\tilde{a}_\lambda$ OF THE LINES AND THE COLUMNS OF Θ_λ

We define the operators s_λ, the symmetrizer of the lines of Θ_λ, and a_λ, the antisymmetrizer of the lines of Θ_λ, as follows

$$s_\lambda \equiv \frac{1}{\lambda_1!\, \lambda_2! \dots \lambda_h!} \sum_h h_\lambda \tag{D.62}$$

$$a_\lambda \equiv \frac{1}{\lambda_1!\, \lambda_2! \dots \lambda_h!} \sum_h (-)^h h_\lambda \tag{D.63}$$

the summation extending over the $\lambda_1!\, \lambda_2! \dots \lambda_h!$ permutations h_λ that leave the *lines* of Θ_λ invariant, i.e. those for which the elements of each cycle are all on the same line of Θ_λ. The h_λ form a subgroup $\mathscr{S}_{[\lambda]}$ of $\mathscr{S}_n$; $\mathscr{S}_{\lambda_i}$ being the group of permutations of the λ_i objects on the ith line of Θ_λ, $\mathscr{S}_{[\lambda]}$ is the product of the subgroups

$$\mathscr{S}_{\lambda_1}, \mathscr{S}_{\lambda_2}, \dots, \mathscr{S}_{\lambda_h} \text{ of } \mathscr{S}_n;$$

s_λ and a_λ are respectively the products of the symmetrizers and the antisymmetrizers of these h subgroups.

The symmetrizer $\tilde{s}_\lambda$ and the antisymmetrizer $\tilde{a}_\lambda$ of the columns of Θ_λ (i.e. the lines of $\tilde{\Theta}_\lambda$) are similarly defined; they are the operators

$$\tilde{s}_\lambda \equiv \frac{1}{\tilde{\lambda}_1!\, \dots \tilde{\lambda}_k!} \sum_v v_\lambda \tag{D.64}$$

$$\tilde{a}_\lambda \equiv \frac{1}{\tilde{\lambda}_1!\, \dots \tilde{\lambda}_k!} \sum_v (-)^v v_\lambda \tag{D.65}$$

the summation extending over the $\tilde{\lambda}_1! \dots \tilde{\lambda}_k!$ permutations v_λ leaving invariant the *columns* of Θ_λ.

Principal properties

a) These are products of symmetrizers or of antisymmetrizers, and therefore

$$s_\lambda^2 = s_\lambda \qquad a_\lambda^2 = a_\lambda \qquad \tilde{s}_\lambda^2 = \dots \tag{D.66}$$

$$h_\lambda s_\lambda = s_\lambda h_\lambda = s_\lambda \qquad h_\lambda a_\lambda = a_\lambda h_\lambda = (-)^h a_\lambda \qquad v_\lambda \tilde{s}_\lambda = \dots \tag{D.67}$$

b) If $\varpi \equiv \sum_p x_p p$ is an arbitrary linear combination of the elements of $\mathscr{S}_n$ (an element of the group algebra) and if [1])

$$\lambda > \mu \qquad \text{or} \qquad \tilde{\lambda} < \tilde{\mu} \tag{D.68}$$

then

$$\begin{aligned} \tilde{a}_\mu \varpi s_\lambda &= s_\lambda \varpi \tilde{a}_\mu = 0 \\ \tilde{s}_\mu \varpi a_\lambda &= a_\lambda \varpi \tilde{s}_\mu = 0. \end{aligned} \tag{D.69}$$

In particular, it results from either of the conditions (D.68) that

$$\tilde{a}_\mu s_\lambda = s_\lambda \tilde{a}_\mu = \tilde{s}_\mu a_\lambda = a_\lambda \tilde{s}_\mu = 0. \tag{D.70}$$

"IRREDUCIBLE SYMMETRIZERS" i_λ AND j_λ

$$\begin{aligned} i_\lambda &\equiv s_\lambda \tilde{a}_\lambda \equiv \sum_{hv} (-)^v h_\lambda v_\lambda \qquad (i_\lambda \neq 0) \\ j_\lambda &\equiv \tilde{a}_\lambda s_\lambda \equiv \sum_{hv} (-)^v v_\lambda h_\lambda \qquad (j_\lambda \neq 0). \end{aligned} \tag{D.71}$$

We call i_λ the "irreducible symmetrizer" and j_λ the "irreducible antisymmetrizer" of Θ_λ. We similarly define the irreducible symmetrizer and antisymmetrizer of $\tilde{\Theta}_\lambda$:

$$\tilde{i}_\lambda \equiv \tilde{s}_\lambda a_\lambda \qquad \tilde{j}_\lambda = a_\lambda \tilde{s}_\lambda,$$

ϖ being defined as above, one can show that

$$s_\lambda \varpi \tilde{a}_\lambda = Cst.\ i_\lambda \qquad \tilde{a}_\lambda \varpi s_\lambda = Cst.\ j_\lambda. \tag{D.72}$$

From (D.69) and (D.72) we easily deduce

$$i_\lambda i_\mu = i_\mu i_\lambda = Cst.\ \delta_{\lambda\mu} i_\lambda \qquad j_\lambda j_\mu = j_\mu j_\lambda = Cst.\ \delta_{\lambda\mu} j_\lambda. \tag{D.73}$$

---

[1]) One of the conditions (D.68) does not necessarily imply the other, *Example*: if $\lambda = [41^2]$ and $\mu = [33]$, we have $\lambda > \mu$; however $\tilde{\lambda} = [31^3]$. $\tilde{\mu} = [2^3]$ and therefore $\tilde{\lambda} > \tilde{\mu}$.

FUNDAMENTAL THEOREM. (*i*) *In the space of the regular representation of the group* $\mathscr{S}_n$, *the vectors* ϖi_λ (ϖ, arbitrary element of the algebra of $\mathscr{S}_n$) *span an irreducible invariant subspace* $\mathscr{E}_\lambda$ *and are therefore associated with a certain irreducible representation* $\mathbf{P}^{(\lambda)}$ *of* $\mathscr{S}_n$.

(*ii*) *The vectors* ϖj_λ *span the same invariant subspace* $\mathscr{E}_\lambda$.

(*iii*) *If* $\lambda \neq \mu$, *the irreducible representations* $\mathbf{P}^{(\lambda)}$ *and* $\mathbf{P}^{(\mu)}$ *are inequivalent.*

Since the number of irreducible representations of $\mathscr{S}_n$ is equal to the number of its classes, and therefore to the number of partitions of n, this theorem permits the construction of them all. Consequently, each irreducible representation of $\mathscr{S}_n$ may be characterized by a well-defined Young diagram.

"SYMMETRIZERS" OF ASSOCIATED TABLEAUX AND ASSOCIATED REPRESENTATIONS

To each element $\varpi = \sum_p x_p p$ of the group algebra, we associate the element $\varpi' = \sum_p (-)^p x_p p$. This correspondence is linear, one-to-one, and has the following properties:

a) if $\eta = p\varpi$, $\eta' = (-)^p p \varpi'$;

b) if $\zeta = \varpi\eta$, $\zeta' = \varpi'\eta'$ (conservation of the product).

One notes that the "symmetrizers" of the associated tableaux Θ_λ, $\tilde{\Theta}_\lambda$ are in one-to-one correspondence in this way, i.e.

$$a_\lambda = s_\lambda' \qquad \tilde{a}_\lambda = \tilde{s}_\lambda'$$

whence

$$\tilde{i}_\lambda \equiv \tilde{s}_\lambda a_\lambda = j_\lambda' \qquad \tilde{j}_\lambda \equiv a_\lambda \tilde{s}_\lambda = i_\lambda'. \tag{D.74}$$

Let $\varpi_1 i_\lambda, \varpi_2 i_\lambda, \ldots$, be a set of basis vectors defining the representation $\mathbf{P}^{(\lambda)}$ in accordance with the Fundamental Theorem; from (D.74) the corresponding vectors are $\varpi_1' \tilde{j}_\lambda, \varpi_2' \tilde{j}_\lambda, \ldots$; they define the representation $\mathbf{P}^{(\tilde{\lambda})}$. With this choice of bases, it is clear that the matrices $P^{(\lambda)}$, $P^{(\tilde{\lambda})}$ representing in each of them a given permutation p are related by

$$P^{(\tilde{\lambda})} = (-)^p P^{(\lambda)}. \tag{D.75}$$

"SYMMETRIZERS" OF $\Theta_\lambda{}^p$

Just as we did for the tableau Θ_λ, we can define with the help of $\Theta_\lambda{}^p$

the permutations $h_\lambda{}^p$, $v_\lambda{}^p$ and the symmetrizers $s_\lambda{}^p, \ldots, i_\lambda{}^p, \ldots$. We note that

$$h_\lambda{}^p = p\, h_\lambda\, p^{-1} \qquad v_\lambda{}^p = p\, v_\lambda\, p^{-1}$$

whence

$$s_\lambda{}^p = p\, s_\lambda\, p^{-1}, \qquad a_\lambda{}^p = \ldots, \text{ etc.} \tag{D.76}$$

The "symmetrizers" of $\Theta_\lambda{}^p$ have properties analogous to the properties of the "symmetrizers" of Θ_λ; they can be deduced from the latter with the aid of relations (D.76).

17. Principal Properties of the Irreducible Representations of $\mathscr{S}_n$

Most of the properties of the irreducible representations of $\mathscr{S}_n$ follow from the properties of the "symmetrizers" given in the preceding paragraph.

Each irreducible representation $\mathbf{P}^{(\lambda)}$ of $\mathscr{S}_n$ is characterized by a certain Young diagram and may be constructed with the aid of the irreducible symmetrizer i_λ (or j_λ).

Each of them is *self-conjugate*

$$\mathbf{P}^{(\lambda)*} \simeq \mathbf{P}^{(\lambda)}. \tag{D.77}$$

Representations of degree 1

The *only two* representations of degree 1 are

(*i*) the identical (or symmetrical) representation **S**;

(*ii*) the antisymmetrical representation **A** in which each permutation p is represented by $(-)^p$.

They are respectively generated by s and a [eq. (D.59)].

The Young diagram for **S** has a single line and represents the partition $[n]$, that of **A** a single column and represents the partition $[1^n]$.

Associated irreducible representations

Two irreducible representations $\mathbf{P}^{(\lambda)}$, $\mathbf{P}^{(\tilde{\lambda})}$ are said to be associates one of the other if their Young diagrams are associates one of the other. It results from eq. (D.75) that:

$$\mathbf{P}^{(\tilde{\lambda})} \simeq \mathbf{A} \otimes \mathbf{P}^{(\lambda)}. \tag{D.78}$$

Components of degree 1 of the tensor product $\mathbf{P}^{(\lambda)} \otimes \mathbf{P}^{(\mu)}$

$\mathbf{P}^{(\lambda)} \otimes \mathbf{P}^{(\mu)}$ has one (and only one) component of degree 1 if, and only if, we have one of the following:

(*i*) $\lambda = \mu$, in which case that component is **S**;

(*ii*) $\lambda = \tilde{\mu}$, in which case that component is **A**.

IRREDUCIBLE REPRESENTATIONS OF $\mathscr{S}_{n-1}$ CONTAINED IN $\mathbf{P}^{(\lambda)}$

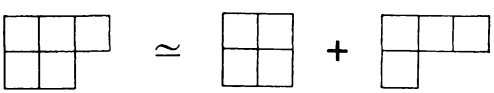
Any irreducible representation of $\mathscr{S}_n$ is a representation (not necessarily irreducible) of its subgroup $\mathscr{S}_{n-1}$.

Let us denote by the symbol $\mathbf{P}_t^{(\lambda)}$ the irreducible representation of the group $\mathscr{S}_t$ corresponding to the partition λ of the integer t. It can be shown that the decomposition of $\mathbf{P}_n^{(\lambda)}$ into parts irreducible *with respect to the group* $\mathscr{S}_{n-1}$ is given by:

$$\mathbf{P}_n^{(\lambda)} \simeq \sum_{\mu} \mathbf{P}_{n-1}^{(\mu)}, \tag{D.79}$$

the summation being extended over all partitions μ of the integer $(n-1)$ corresponding to Young diagrams that can be obtained by removing a square from the diagram corresponding to the partition λ of the integer n.

Example:



Eq. (D.79) can be used to relate certain characters of $\mathscr{S}_n$ to the characters of $\mathscr{S}_{n-1}$. In particular it can be used to deduce the degrees of the representations of $\mathscr{S}_n$ from the degrees of the irreducible representations of $\mathscr{S}_{n-1}$.

CONSTRUCTION OF THE IRREDUCIBLE INVARIANT SUBSPACES OF A GIVEN REPRESENTATION **P**

In the representation **P**, the various "symmetrizers" previously introduced are represented by *linear, Hermitean* operators (to within a constant, projectors) which we shall denote by the corresponding capital letter.

The method of constructing the irreducible components of **P** follows from the fundamental theorem of § 16. If $|u\rangle$ is an arbitrary vector of the space $\mathscr{E}$ of the representation **P**, $I_\lambda|u\rangle$, if not null, transforms like $\mathbf{P}^{(\lambda)}$ and its representation space contains the non-null vector $J_\lambda|u\rangle$. The set of vectors $I_\lambda|\rangle$ formed by application of I_λ to the basis vectors of $\mathscr{E}$ spans a subspace $\mathscr{J}_\lambda$ having a number of dimensions equal to the number n_λ of components $\mathbf{P}^{(\lambda)}$ contained in

P. Let $|\sigma\rangle$ be one of the vectors of an orthonormal basis in $\mathscr{J}_\lambda$; the space $\mathscr{E}_{\sigma\lambda}$ formed by application of the operators of the group to $|\sigma\rangle$ is a representation space for $\mathbf{P}^{(\lambda)}$; one obtains n_λ orthogonal representation spaces by effecting this manipulation on the n_λ basis vectors of $\mathscr{J}_\lambda$.

If Q is an operator of $\mathscr{E}$ invariant with respect to $\mathscr{S}_n$, it transforms the vectors of each of the subspaces $\mathscr{J}_\lambda$ linearly one into another. The problem of diagonalizing Q in $\mathscr{E}$ is completely solved once we have diagonalized it in each of the subspaces $\mathscr{J}_\lambda$.

Symmetry properties of the vectors of $\mathbf{P}^{(\lambda)}$

Ket vectors do not in general have well-defined properties of symmetry or antisymmetry. We say that a vector $|\rangle$ is of the well-defined symmetry S_λ if it belongs to the subspace of the projector S_λ: such a vector is symmetrical in any permutation of the elements of a same line of Θ_λ. In the same way a vector is A_λ-antisymmetrical if it belongs to the subspace of the projector A_λ.

Starting from the tableaux $\tilde{\Theta}_\lambda$, $\Theta_\lambda{}^p$, one similarly defines the symmetries of type $\tilde{S}_\lambda$, $S_\lambda{}^p$ and the antisymmetries of type $\tilde{A}_\lambda$, $A_\lambda{}^p$ respectively [1]).

Of the two vectors of well-defined symmetry $S_\lambda{}^p$, $S_\mu{}^q$ the *more symmetrical* is by definition the one that corresponds to the larger of the partitions λ, μ. Of the two vectors of well-defined antisymmetry $A_\lambda{}^p$, $A_\mu{}^q$, the *more antisymmetrical* is the one that corresponds to the larger of the partitions λ, μ.

We can now deduce from (D.69) and the Fundamental Theorem that:

The space of an irreducible representation $\mathbf{P}^{(\lambda)}$ *contains one and only one* S_λ*-symmetrical vector* (and consequently one, and only one $S_\lambda{}^p$-symmetrical vector, p being any permutation) *and no vector of greater symmetry* [2]).

---

[1]) The permutation P transforms any S_λ-symmetrical vector into a $S_\lambda{}^p$-symmetrical vector and any A_λ-antisymmetrical vector into an $A_\lambda{}^p$-antisymmetrical vector [cf. relations (D.76)].

There is no such correspondence between the symmetries of type S_λ, A_λ and the symmetries of type $\tilde{S}_\lambda$, $\tilde{A}_\lambda$; the latter are equivalent to the symmetries of the type $S_{\tilde{\lambda}}{}^q$, $A_{\tilde{\lambda}}{}^q$ respectively, q being the particular permutation defined by

$$\tilde{\Theta}_\lambda = q\Theta_{\tilde{\lambda}}.$$

[2]) More generally, there exists no S_μ-symmetrical vector corresponding to any partition μ satisfying one or the other of the inequalities

$$\mu > \lambda \qquad \tilde{\mu} < \tilde{\lambda}.$$

It contains one, and only one $A_{\tilde{\lambda}}$-antisymmetrical vector, and no vector of greater antisymmetry.

CLASS SUM OPERATOR FOR TRANSPOSITIONS K_T

To each partition μ there corresponds a certain class and a class sum operator K_μ. The various possible eigenvalues $k_\mu^{(\lambda)}$ of this operator correspond to the different irreducible representations $\mathbf{P}^{(\lambda)}$ of $\mathcal{S}_n$; they are functions of the integers $\lambda_1, \lambda_2, \ldots, \lambda_h$ occurring in the partition λ of n. (*N.B.* If two partitions are different, the corresponding eigenvalues of a particular K_μ are not necessarily different: $\lambda \neq \lambda'$ does not necessarily mean that $k_\mu^{(\lambda)} \neq k_\mu^{(\lambda')}$.) Let us consider in particular the class sum operator K_T for the transpositions:

$$K_T \equiv K_{[2\,1^{n-2}]} \equiv \sum_{i<j} (i\,j).$$

If we note that K_T commutes with all the permutations and that therefore

$$k_T^{(\lambda)} I_\lambda = K_T I_\lambda = S_\lambda K_T \tilde{A}_\lambda,$$

and if we take eqs. (D.67) into account, it is easy to show that $k_T^{(\lambda)}$ is equal to the difference between the number of transpositions of type h_λ (symmetrical pairs) and the number of transpositions of type v_λ (antisymmetrical pairs):

$$k_T^{(\lambda)} = \sum_{i=1}^{h} \frac{\lambda_i(\lambda_i - 1)}{2} - \sum_{j=1}^{k} \frac{\tilde{\lambda}_j(\tilde{\lambda}_j - 1)}{2}. \tag{D.80}$$

18. System of n Fermions of Spin $\frac{1}{2}$

SYMMETRY OF THE STATES OF n IDENTICAL SPINS $\frac{1}{2}$

THEOREM. *The space spanned by the vectors of total spin (SM) formed by n identical spins $\frac{1}{2}$ is associated with an irreducible representation of $\mathcal{S}_n$, the representation whose Young diagram corresponds to the partition $[\frac{1}{2}n+S, \frac{1}{2}n-S]$ of the integer n* (a Young diagram with at most two lines).

COROLLARY I. If $n=2$, there is one antisymmetrical state, the state $S=0$, and 3 linearly independent symmetrical states, the states $S=1$.

COROLLARY II. If $n>2$, there are no antisymmetrical states; there are in all $n+1$ linearly independent, completely symmetrical states, namely the $2S+1 \equiv n+1$ states for which the total spin is a maximum $(S=\frac{1}{2}n)$.

Proof

The space $\mathscr{E}^{(s)}$ formed of the n spins $\frac{1}{2}$ is of 2^n dimensions. Denote by u_i, v_i the dynamical states of the ith individual spin corresponding respectively to the eigenvalues $+\frac{1}{2}$ and $-\frac{1}{2}$ of s_z. An orthonormal basis in $\mathscr{E}^{(s)}$ is obtained by taking all possible products of n of these u and v vectors. Thus the vector

$$\zeta_M \equiv u_1 u_2 \dots u_\nu v_{\nu+1} \dots v_n$$

is a vector of this basis; it is also an eigenvector of the component S_z of the total spin

$$\boldsymbol{S} = \sum_{i=1}^{n} \boldsymbol{s}_i,$$

and the corresponding eigenvalue is $M = \nu + \frac{1}{2}n$. By effecting all possible permutations on ζ_M, we form in all

$$\binom{n}{\nu} = [n!/(\tfrac{1}{2}n+M)!\,(\tfrac{1}{2}n-M)!]$$

distinct vectors which span the subspace of the eigenvalue M; each of these vectors contains $\frac{1}{2}n+M$ vectors of type u and $\frac{1}{2}n-M$ vectors of type v.

This subspace can subsequently be subdivided into orthogonal subspaces corresponding to the different possible eigenvalues of $\boldsymbol{S}^2$; the quantum number S can take the $\frac{1}{2}n-|M|+1$ possible values: $|M|$, $|M|+1$, ..., $\frac{1}{2}n$. We denote by $\mathscr{E}^{(s)}(SM)$ the subspace formed by the vectors of total spin (SM). Since $\boldsymbol{S}^2$ and S_z *commute with all of the permutations*, each of these subspaces defines a representation o $\mathscr{S}_n$; moreover, since S_+ and S_- commute with all the permutations, the representations defined by two subspaces corresponding to the same value of S are equivalent.

We shall now prove Corollary II (Corollary I is evident). For this it suffices to consider the projections of the vectors of an orthonormal basis in $\mathscr{E}^{(s)}$ onto the space of symmetrical states and onto the space of antisymmetrical states. This is particularly simple with the basis vectors defined above. Since $n>2$, any vector ζ of that basis contains at least two individual spins in the same state, u or v; suppose that ζ contains the factor $u_i u_j$; since $A = A\frac{1}{2}(1-(ij))$ and $\frac{1}{2}(1-(ij))\,u_i u_j = 0$, necessarily $A\zeta = 0$. On the other hand there is one, and only one totally symmetrical linear combination of basis vectors belonging to

the subspace of the eigenvalue M, namely the sum of these $\binom{n}{\nu}$ basis vectors; since this is true for each possible value of M, this totally symmetrical vector corresponds necessarily to the maximum value $S=\frac{1}{2}n$ of the total spin. This completes the proof of Corollary II.

The demonstration carried out with A may be repeated with the antisymmetrizer $\tilde{A}_\lambda$ with the result that $\tilde{A}_\lambda\zeta=0$ if the Young diagram Y_λ contains more than two lines. It follows that the Young diagrams for the irreducible components of the representation of $\mathscr{S}_n$ defined in $\mathscr{E}^{(s)}$ have at most two lines.

Let $\lambda=[\lambda_1\lambda_2]$ be a partition of n that meets this condition. The number of irreducible components $\mathbf{P}^{(\lambda)}$ is equal to the number of linearly independent vectors of the type $I_\lambda|\rangle \equiv S_\lambda\tilde{A}_\lambda|\rangle$ that can be formed. For the enumeration of these we need only to consider the tableau Θ_λ:

1	2	.....		λ_2	λ_2+1	.....			λ_1
λ_1+1	λ_1+2	.....		$\lambda_1+\lambda_2$					

The n spins $\frac{1}{2}$ may be put into two separate sets, the $(\lambda_1-\lambda_2)$ elements $\lambda_2+1, \lambda_2+2, \ldots, \lambda_1$ of the first line having no partner in the second, and the λ_2 pairs of elements situated in the same column. Let the total spins of these sets be denoted by $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$ respectively:

$$\boldsymbol{S} = \boldsymbol{S}_1 + \boldsymbol{S}_2.$$

Denote by $S^{(\lambda_1-\lambda_2)}$ the projector onto the states symmetrical with respect to the $(\lambda_1-\lambda_2)!$ permutations of the spins of the first set. Clearly

$$I_\lambda = S_\lambda S^{(\lambda_1-\lambda_2)}\tilde{A}_\lambda.$$

However, from its very definition, $\tilde{A}_\lambda$ is the projector onto the singlet state of each of the pairs of the second set; there is only one vector of this set having this property (Corollary I), it corresponds to the spin $S_2=0$. As for $S^{(\lambda_1-\lambda_2)}$, it projects the vectors of the first set onto the $(2S+1)$-dimensional space corresponding to the largest possible value of S_1 (Corollary II), namely $\frac{1}{2}(\lambda_1-\lambda_2)$. Thus $S^{(\lambda_1-\lambda_2)}\tilde{A}_\lambda$ projects onto a $(2S+1)$-dimensional space corresponding to the value $S=\frac{1}{2}(\lambda_1-\lambda_2)$ of the total spin. Since the projector S_λ commutes with $\boldsymbol{S}$, its action on the vectors of this space is either to annul them all

or to transform them into vectors of the same total spin. The first possibility is excluded, otherwise it would be impossible to find vectors of total spin $\frac{1}{2}(\lambda_1-\lambda_2)$ in $\mathscr{E}^{(s)}$. Our theorem follows from the second. Q.E.D.

FORMATION OF COMPLETELY ANTISYMMETRICAL VECTORS

The dynamical states of n fermions of spin $\frac{1}{2}$ span a space formed by the tensor product $\mathscr{E}^{(0)} \otimes \mathscr{E}^{(s)}$ of the space $\mathscr{E}^{(s)}$ defined above and the space $\mathscr{E}^{(0)}$ of the orbital variables. $\mathscr{E}^{(0)}$ can be subdivided into mutually orthogonal, irreducible invariant subspaces $\mathscr{E}^{(0)}(\sigma\mu)$ with respect to the group $\mathscr{S}_n$; $\mathscr{E}^{(0)}(\sigma\mu)$ is associated with a certain irreducible representation $\mathbf{P}^{(\mu)}$; the index σ is to distinguish between the spaces associated with the same irreducible representation of $\mathscr{S}_n$. To form a complete set of orthogonal, totally antisymmetrical vectors, we need only to do so in each subspace $\mathscr{E}^{(0)}(\sigma\mu) \otimes \mathscr{E}^{(s)}$. The number of linearly independent, antisymmetrical vectors in such a subspace is equal to the number of times that the component $\mathbf{A}$ figures in the representation defined in this subspace and therefore, (§ 17), to the number of irreducible representations associated with $\mathbf{P}^{(\mu)}$ figuring in the decomposition of the representation defined in $\mathscr{E}^{(s)}$. *If Y_μ has more than two columns there are none. If Y_μ contains at most two columns, and if we put*

$$\mu = [2^{\frac{1}{2}n-S}\, 1^{2S}], \tag{D.81}$$

in accordance with the theorem above, there are $(2S+1)$ of them, and the antisymmetrical vectors that can be formed are the eigenvectors of $\mathbf{S}^2$ corresponding to the total spin (SM) $(M=-S, -S+1, \ldots, S)$.

SCALAR OBSERVABLES INDEPENDENT OF THE SPINS. LS COUPLING

If Q is a spin-independent, scalar observable, it may be regarded as an observable of the space $\mathscr{E}^{(0)}$ invariant under rotation and permutation of the orbital variables alone:

$$[Q, L] = 0 \qquad [Q, P^{(0)}] = 0.$$

$\mathscr{E}^{(0)}$ is the direct sum of $(2L+1)d_\mu$-dimensional subspaces $\mathscr{E}^{(0)}(\tau L\mu)$ irreducible with respect to the group of rotations and permutations [1]:

[1] $\mathscr{E}^{(0)}(\tau L\mu)$ is reducible with respect to the rotation group alone; it is the sum of d_μ equivalent irreducible subspaces (d_μ = degree of the representation $\mathbf{P}^{(\mu)}$). Similarly, it is reducible with respect to the group $\mathscr{S}_n$; it is the sum of $(2L+1)$ equivalent irreducible subspaces.

L is the angular momentum quantum number, μ the partition associated with the representation $\mathbf{P}^{(\mu)}$ of the permutation group, τ an additional quantum number to distinguish between equivalent subspaces. In particular, we can choose $\mathscr{E}^{(0)}(\tau L\mu)$ so as to have Q equal to a certain constant $q_\tau^{(L\mu)}$ in each of them (cf. § D.9).

The eigenvectors of Q corresponding to the eigenvalue $q_\tau^{(L\mu)}$ are the antisymmetrical vectors of the subspace $\mathscr{E}^{(0)}(\tau L\mu) \otimes \mathscr{E}^{(s)}$. According to the above demonstration, these can only be formed if Y_μ has at most two columns; the partition μ is then unambiguously defined by the specification of the quantum number S [eq. (D.81)]. The subspace of the antisymmetrical vectors of $\mathscr{E}^{(0)}(\tau L\mu) \otimes \mathscr{E}^{(s)}$ corresponds to a well-defined value S of the total spin of the particles; it is a certain $(2L+1)(2S+1)$-dimensional subspace $\mathscr{E}(\tau LS)$ in which one can choose eigenvectors of total orbital angular momentum and total spin

$$|\tau L S M_L M_S\rangle \qquad (M_L = -L, \ldots, +L;\ M_S = -S, \ldots, +S)$$

which form a standard basis $\{L^2 S^2 L_z S_z\}$.

GENERAL INDEX

(Page numbers relating to Volume 1 are in *italics*; starred page numbers refer to footnotes)

A

Absorption, emission of light, *9*, *18*, *23*, 991, 1007, 1041–5
  of matter, *60*, 875, 955
  (*see* Annihilation)
Accidental degeneracy, *355*, 932
Action, *10*
  atomism of, *41–42*, *115*
  integral, *36*
  principle of least action of Maupertuis–Hamilton, *32*, *53*, 966
Addition of angular momenta, 555–569
Adiabatic, 739
  approximation, 750–5, 781, 786–92, 795
  passage, 740, 744
  theorem, 740, 744, 747–50
Adjoint of an operator, *255*
d'Alembertian (*see* Dalembertian)
Alder, 728*
Amplitude, scattering, *372*, *381*, *383*, 607, 802
  Coulomb, *424*
  transition, 807, 836, 1002
Annihilation, creation of matter, *60*
  of corpuscles, *439*, 875, 964
  of pairs of corpuscles, 888, 955
  (*see* Operator)
Anticommutation, 545*
Antilinear (correspondence), *247*, 637
  (*see* Operator)
Antisymmetrical (*see* State)
Antisymmetrizer, 589–91, 630, 774, 1112
Antiunitary (*see* Operator, Transformations)
Atom, Bohr, *23*, *24*
  complex, 612–19, 700–708, 773–81
  hydrogen, *22*, *33–34*, *37*, *55*, *71*, *412–20*, 767–9, 930–3, 956, 958, 1051, 1052
  hydrogen-like, *39*, *413*
  Rutherford, *8*, *21–3*
Average (*see* Mean value)
Avogadro's number, *5**, *8*

B

Balmer (formula), *22*
Barn, *371*
Barrier, potential, *78*, *86*, *96*, *237*
Basis, *203*, *273*
Bauer, E., *157**
Bethe, H. A., *409**
  formula, *408*
Black body radiation, *10–11*
Blatt, J., 576*, 627*, 832*, 1043*, 1055*
Bloch, C., 719
Bloch, F., *449*
Bohm, D., *157**
Bohr, N., *23*, *29*, *30*, *34*, *38**, *47*, *152*, 960*
Bohr,
  atom, *23*, *24*
  frequency, *23*, 725, 753, 1000
  magneton, 543, 708, 755, 936
  orbit, *37*
  –Peierls–Placzek relation, *409*, 866–7
  (*see* Atom, Quantization)
  radius, *414*
Boltzmann, L., 598
  constant, *205*
  factor, *205*
  (*see* Statistics)
Born, M., *4**, *10**, *21**, *45**, *46*, *47*, *330**, 614, 740*, 763, 783*, 792, 954*
Born approximation, 736–8, 806, 808–15, 822, 824, 825, 832, 839, 1003
  for phase shifts, *405–6*, *430*, 818, 861
Bose (*see* Statistics)

Boson, 595, 596–8, 603, 799
Bound (*see* State)
Bra, bra vector, *246*
Branching (*see* Ratio)
Brillouin, L., *69**, *216**
Brownian motion, *7*, *56**

C

c, *7*, *11*
Canonical equations, *32* (*see* Hamilton)
Causal, causality, *156–9*, *310*
Central field approximation, 612–14, 701
Central potential, *344*, *355* (*see* Scattering)
Centre of mass, *361–6*, *380–5*, 1046
  system, *380–5*, 833
Channel, 833
Charge independence, 626–9, 659, 693
Chemical properties, 614
Chew, G., *372**, *409**
Classical,
  doctrine, *3*, *21*
  limit, *49*, *53*, *133*, *140–1*, *214–31*, *395*, *446–8*, 616, 685, 730, 1039
  mechanics, *3–11*, *23*, 645*, 650*, 664, 740*
  Theory *vs* Quantum Theory, *15–18*, *46*, *154*, *166*, *294*, *317*, *361–3*, *387–9*, *420*, *424*, *444–8*, 507, 527, 540, 582–3, 633, 644, 665–7
  (*see* Mixture, statistical)
Clebsch–Gordon coefficients, 560–3, 573, 576, 678, 731, 1053–60
Closure relation, *187*, *268–71*
Cloud chamber, *8*, *142**
Collision, 736–8, Chapter XIX, 1001–1009, 1045–9
  complex, 832–56, 862
  of identical particles, 603–9
  rearrangement, 832–3
  (*see* Scattering)
Combination principle of Rydberg–Ritz, *22*, *23*
Commutator, *63*, *163*, *206–12*, *250* (*see* Observables, Complete)
Compatible variables, *153–5*, *199–202* (*see* Complete)
Complementarity, complementary, *149–159*, *401*, 1007
  (*see* Subspace, Measurement)
Complete
  set of commuting observables, *202–4*, *273*, *301*
  set of compatible variables, *155*, *203*, *294*
  set of functions, vectors, *174–5*, *187*, *259*
  space, *165*
  theory, *151*
Compton, A. H., *13*, *15*
  effect, *13–18*, *41*, *45*, *144*, *198*
  formula, *13*, *16*
  wavelength, *13**, 943, 948, 981
Condon, E. U., 615*, 773*, 1055*
Configuration, 614
  space, *31*, *68*, *164*
Conjugate momentum, *32* (*see* Lagrange)
Conjugation
  between vectors, operators, *254–6*
  charge, 915, 949–50, 957
  complex, and time reversal, 671–3
  complex conjugation operator, 641, 670, 915
Connection formula for WKB Solutions, *234–7*
Conservation
  laws, *21*, *42*, *60*, Chapter XV, 914, 1011–14
  of angular momentum, 540, 657–8, 1013
  of flux, *108*, *121*
  of momentum, energy, *14*, 657–8, 664–5, 1013, 1046
  of parity, *211*, 657–8, 662, 915
Constant of the motion, *210*, *211*, *318–9*, 657, 914–15
Contravariant (*see* Covariant)
Correspondence
  principle, *27–42*, *46*, *50*, *61*, *64*, *68–71*, *214*, *317*, *338*, *446*, 876, 885, 889, 919, 921, 959, 1029

relations, wave-corpuscle, *14*, *20*, *55*, *59*, *117–8*
Coulomb (field, potential, interaction), *22*, *33*, *71*, Chapter XI, 612, 692, 1026, 1038
excitation, 728–32, 844–6
scattering, *228–31*, *421–30*, 608, 831–2
(*see* Amplitude)
Coupling, *LS*(Russel–Saunders), 703–6, 708, 720, 743, 1123
jj, 703, 720
(*see* Spin-orbit)
Covariant
form, covariant formalism, 883, 892–3, 895, 959, 991
vectors, covariant and contravariant indices, 877
Creation (*see* Annihilation, Operator) of matter, *60*
Cross section, *74*, *229*, *369–71*, 835, 1002
geometrical, 814
Current, *121*, *122*, *223*, *372*, 887, 893–5, 921, 937, 1038
Cut-off, 948, 981–2, 990

D

Dalembertian, *65*
Darwin (term), 816
Davisson and Germer, *57*
de Broglie, L., *46*, *47**, *48**, *49*, *61*, *65*, *66*
relation, *14*, *52*, *55*, *117*
wavelength, *50–7*
Debye, P., *13*
Debye–Scherrer rings, *57*
Decomposition into partial waves, *385*, *389*, *425–6*
Decomposition of unity, *190–1*, *267*
Degenerate, degeneracy, *78*, *172*, *258*, *354–5*, *417–18*, *453*, 709–11
accidental, *355*, 932
exchange, 583, 594
G-degeneracy, 657, 676*, 1097
Kramers, 675
removal of, 709–10
rotational, 540
δ-function of Dirac, *179–184*, *468–71*
Density of states, levels, 734, 736, 737, 806, 864, 865
(*see* Operator)
Determinant, of a matrix, *276*, *285*
Slater, 599, 613, 704, 773
Deuteron, 565
Diagonal (*see* Matrix)
Diagonalization, *287*
Diamagnetism, 541*, 553
Diffraction, *115*, *118*, *155*
by a hard sphere, *395* (*see* Interference)
Dipole, dipolar(*see* Moment, Transition)
Dirac, P. A. M., *45**, *47*, *245*, *246*, *255*, *434*
equation, Chapter XX, 844
identity, 629
matrices, 896–900
(*see* Representation, Self-consistent, Statistics)
Direct sum, *264*, 1087–88
Dispersion,
formula, 1001–1005
law, *50*
Distribution (Schwartz), *182*, Appendix A
Doppler effect, *15*
Dual space, *246–7*
Duality, wave-corpuscle, *21*, *42*, *47*, *49*, *59*, *115*, *155*, 964

E

Eckart (*see* Wigner)
Edmonds, A. E., 1053*
Effective range, *406–9*, *430*
Ehrenfest, 740*
theorem of, *216*
Eigendifferential, *112*, *170*, *185*, *249*
Eigenfunction, *72*
Eigenvalue
of a matrix, *287*
of an operator, *171*, *185*, *258*
problem, eigenvalue equation, *72*, *169*, *171–4*, *184–8*, *257–60*
(*see* Subspace)

Einstein, *7*, *8*, *11–12*, *13*, *21**, *48**
relation ($E = \hbar\omega$), *72*, *310*
(*see* Statistics)
Electrodynamics
classical (Maxwell–Lorentz), *4–7*, *7–11*, *12*, *15*, *22*, 1009–1011
quantum, *432*, *438–9*, 933
Electromagnetic field, 879–80
particle in, *43*, *47*, *49–50*, *67–8*, *70*, 551–2, 885, 890, 944–7
Electromagnetic radiation
classical theory, 1009–1029
quantum theory, 1029–1049
energy, momentum, angular momentum of the radiation, 1022
(*see* Electrodynamics)
Electron, *7*, *45*, *46*, *48*, *57–8*, *59–60*, *71*
(*see* Electrodynamics, Electromagnetic radiation)
Eliott, 1065
Emission of light (*see* Absorption)
Energy
conservation of unperturbed energy, 732–6, 738
ionization, 780
(*see* Conservation, Uncertainty, Measurement, Negative, Electromagnetic radiation)
Equation
radial, *349–53*, 928
radial for the hydrogen atom, *412–16*, 930
secular, *278*, *287*, 718
(*see* Dirac, Fock–Dirac, Hamilton, Hartree, Heisenberg, Klein–Gordon, Lagrange, Laplace, Poisson, Schrödinger, Sturm–Liouville, Eigenvalue)
Equilibrium (thermodynamical), *205*, *337*, 597–8
Ergodic hypothesis, *5**, 597
Euler angles, 523–4, 533, 1068
Exchange effect, 780 (*see* Degeneracy)
Exclusion (*see* Pauli)
Expansion (*see* Series expansion)

F

Fano, U., *331**, 569*, 1053*
Fermat's principle, *54*
Fermi (*see* Thomas, Statistics)
Fermion, 595, 598–600, 601, 603, 799
Feshbach, H., *231**, 886*, 888*
Filtering of a wave packet, *196–99*, *298*, *333**
Fine structure (hydrogen atom), *38*, *419*, 932–3
Flux, *369*
relative, *381* (*see* Conservation)
Fock, V., 740*
Fock–Dirac equations, 776–8
(*see* Self-consistent)
Foldy, L.L., 941* (*see* Transformations)
Foley, H. M., 936*
Form factor, 816, 842
Fourier transform, *118*, *127*, *130*, *471–8*
Franck–Hertz experiment, *24*, *40*, *46**
Frequency (*see* Bohr)
Friedrichs, K. O., 740*
Function
characteristic, *177*, *189*, *296*
eigen, *72*
partition, *338*, *448*
spherical Bessel, *355–6*, *488–90*
spherical Coulomb, *426*, *485–8*
(*see* Green, Hamilton, Lagrange, Wave)
Furry, W. H., *231**

G

Galilean transformation, *383**, 647, 680, 875
Gauge,
invariance, 918, 1015–6
Lorentz, 1016
radiation, 1017, 1020, 1025*, 1050
transformation, 1015–16
Gaussian, *131*, *449–51*
Geiger counter, *8*
Gell-Mann, M., 801*
Gerlach (*see* Meyer, Stern)
Germer (*see* Davisson)
Glauber, R., *442*
Goldberger, M., *409**, 801*

Good quantum number (*see* Quantum number)
Goudsmit, S., 544
Green (function), *497*, 712*, 810, 818, 819–22, 846–8, 857, 872
theorem, *120*, 804, 837
(*see* Series expansion)
Ground state, *23*, *24*
Group velocity, *50*, *55*
Groups
of transformation operators, 647–9
of transformations, 646–7
theory of, 528*, 632, Appendix D
(*see* Lorentz)

H

Half integral, 514, 519, 536, 543, 648, 652, 654, 663, 671, 799
Hamilton
equations, *33*, 968
function, *33*, *66–7*
–Jacobi formalism, *223**
(*see* Action)
Hamiltonian, *33*, *68*, *71*, *120*, *310*
in polar coordinates, *344–8*
of electromagnetic radiation, *439*, 1027–8, 1031
of the Dirac electron, 889–90
of the free scalar field, 963, 965, 967–8
of the harmonic oscillator, *432*
Hard sphere (*see* Scattering)
Harmonic (*see* Spherical)
approximation, 798
Harmonic oscillator, Chapter XII, 798, 963–6, 970
Hartree, D. R., 773*
equation of, 780–9
(*see* Self-consistent)
Heisenberg, W., *45**, *46*, *47*, *330**
equation, *316*, *433*
(*see* "Representation", Uncertainty relations)
Heitler, W., 960*, 1046*
Helium, 690–92, 771–3
Hermiticity, hermitean (*see* Matrix, Operator)
Hermite polynomials, *441*, *491–2*
Hertz, H. R., *6*
Hidden variables, *152*
Hilbert space, *164*, *248*
Holes (theory of), 953–6
Homomorphism, isomorphism, 648, 1084, 1089–90
Horowitz, J., 1053*
Hund's rule, 702*
Hydrogen (*see* Atom)
Hylleraas, 773

I

Identical particles, Chapter XIV, 773
Impact parameter, *228*, *373*, *378*, *387–9*
Indistinguishability, 582, 593, 598, 600
Inequality (*see* Schwartz)
Infinitesimal (*see* Operator, Rotations, Transformations)
Interference, *94*, 852–3
and diffraction, *4*, *18–20*, *41*, *47*, *56–9*, *115*, *118*
Intermediate (*see* "Representation")
Intersection of two subspaces, *264*
Invariance, of the Dirac equation, 896–918
of matrix equations, *285*
of the Hamiltonian, *343*, *365*, Chapter XV, 709–10
of the trace, *285*
(*see* Gauge)
Invariant (*see* Subspace)
Inverse,
of a matrix, *277*, *281*
of an operator, *251*, 638
Ionization (*see* Energy)
Irreducibility, reducibility, *302**, 536, 645, 711, 891–2, 898, 1088–9, 1090, 1102–4
(*see* Tensor)
Isomorphism (*see* Homomorphism)

J

Jauch, J. M., 960*, 1046*
Jeffreys, H., *216**
Jordan, P., *45**, *46*, *330**
Jost, R., 812*

K

k, Boltzmann constant (*see* Boltzmann)
Kato, T., 712, 716, 719, 740*, 744*
Kepler's laws, *30*
Ket, ket vector, *245*
Klein, O., 1046*
Klein–Gordon equation, *65*, 884–8, 890, 892, 919, 956, 957, 958, 961
Kohn, W., 812*
Kramers, H. A., *28*, *216**
(*see* Degeneracy)
Kronecker
product, *212**
symbol, *173*
Kusch, P., 936*

L

Laboratory system, *380–5*
Lagrange
conjugate momentum, *32*, 967
equation, *32*, 966
function, *32*, *33*, *43*, 966
Laguerre polynomials, *416*, *419*, *483*
Lamb, W. E., 932*
shift, 933
Landau, L. D., 785*, 793*
Landé factor, 543, 707–8, 721
Langer, R. E., *231**, *235*, *368*, *410*
Langevin, *26*
Laplace's equation, *415*, *422*, *425*, *480*
Laplacian, *63*
Larmor precession (frequency), *25*, *43*
Lenard, *12*
Length (*see* Scattering, Wave)
Lifetime of metastable state, *401–3*, 999, 1007
Lifshitz, E. M., 785*, 793*
Linear (*see* Operator, Superposition)
Lippmann, B., 801*, 857*, 863*
Logarithmic derivative, *81*, *99*, *389*
London, F., *157**
Lorentz, electron theory (*see* Electrodynamics)
force, 883, 957
group, 647, 880–2, 900–908, 967
transformation, *383**, 882*, 908
Loschmidt, *5**
Low, F., *372**
LS (*see* Coupling)

M

Macroscopic scale, limit, *3**, *21*
Magnetic (*see* Electromagnetic field, Moment, Reversal)
Magneton (*see* Bohr)
Mass, reduced, *31**, *362*
(*see* Centre, Negative)
Matrix, *273–9*
diagonal, *276*, *280*
hermitean, *276*
inverse, *277–81*
mechanics, *45–48*, *330**
orthogonal, *277*
reduced, 571, 1075–8
rotation, 536, 1068–75
singular, *277*
S, 863–6
T, 807, 829, 835–7, 863–70
unitary, *277*, *281*
(*see* Pauli, Dirac)
Matter wave (*see* Wave)
Maupertuis (*see* Action)
Maxwell, J. C., *4*, *5*
equations, 1009
(*see* Electrodynamics, Statistics)
Mean life, *28*, *137*, *157*, *403*, 992
relation to width, *137–8*, 992, 1007
Mean value, definition, *122–6*, *166–7*, *243*, *296*, *331–3*
law of motion, *210*, *216–18*, *319*
Measurement, ideal, *196–99*, *298*, *333*
and uncertainty relations, *139–49*
mechanism and complementarity, *149–59*, *157**
of energy, *138*
of momentum, *145–9*
of position, *142–5*
Mechanics (*see* Classical, Matrix, Wave)
Meson, π, 593, 885
Metastable state (*see* State)
Meyer–Gerlach experiment, *12*
Michelson–Morley experiment, *6*
Microreversibility, 673–5, 727, 808, 831, 867, 1009

Millibarn (mb), *371*
Millikan, *8*
Mixture,
  classical, *215*, *337*
  statistical, *204–5*, *331–4*
Moe, M., 857*
Molecule, 773, 781–800
  diatomic, 611, 696, 785, 793–800
  hydrogen, 782*, 798*
  (*see* Rotations, Vibrations)
Moment,
  electric dipole, 1044
  magnetic, *25*, 576
  quadrupole, 576, 610*
  2^l-pole (electric and magnetic), 576, 731, 1037
Momentum, *14*, *52–3*, *54**, *64*
  angular, *25*, *37–39*, *195–6*, *208–9*, *345–9*, Chapter XIII, 913, 921, 977–8, 1022, 1034–7, 1054 (*see* Addition)
  conjugate (*see* Lagrange)
  radial, *346*, *366*, 927
  (*see* Conservation, Measurement, Uncertainty)
Morse, P. M., *231**, *394**
  potential, 797, 800
Mott formula, 608–9
Multipole (expansion), 1031, 1034–7
  (*see* Moment, 2^l-pole)

N

Negative (mass, energy), 885, 887–8, 892, 949–56
Neumann, J. v., *157**, *166**, *191**, *263**
Neutron diffraction, *56**, *57*
  (*see* Nucleon)
Newton, *3*, *4*, *6*, *42*
Nishina, 1046*
Nodes, number of, *88*, *91*, *109–110*
Normal (coordinates, vibrations), 960–3, 969–72
Normalizable function, *185*
Normalization condition, *117*, *270*
  (*see* Orthonormality relations)
Norm of functions, vectors, *165*, *248*
Nucleon, 617
  -nucleon interaction, 553–5, 624–5, 626–9
  system of two nucleons, 553–5, 563–6
Nucleus, atomic, *8*, *60*, *71*, 692–3
  statistics, 799
  (*see* Radius, Spin)

O

Observables, *188*, *194*, *260*, *266–73*, *297*
  commuting, *199–204*, *272–3*
    (*see* Complete)
  fundamental, *298*, 645
  physical, 536, 594, 646, 679
Occupation number, 965
Old Quantum Theory, *27–31*, *241*, *419*, *420*
Operator
  antilinear, antiunitary, 637–43
  bounded, 713*
  creation, destruction, *438–9*, *453*, 977, 1031
  density (or statistical) *295*, *331–8*, 528, 598, 600, 644, 660, 679, 1032*
  differential, *62–4*
  evolution, *310–12*, 722–4, 801
  hermitean, *120*, *125*, *167–8*, *256*
  infinitesimal unitary, *291–2*
  linear, *62–3*, *250–2*
  positive-definite hermitean, *257*, *293*
  rotation, 526–36, 1068–70
  unitary, *257*, *291*
  (*see* Adjoint, Groups, Scalar, Tensor, Vector)
Oppenheimer, J. R., 763, 783*, 792
Optical theorem, 866
Orbit (*see* Bohr, Spin)
Orthogonal
  functions, *111*, *165*, *172–5*
  matrix, *277*
  spaces, projectors, vectors, *248*, *258*, *264*
  (*see* Schmidt)
Orthonormality relations, *174*, *187*, *259*
Oscillator (*see* Harmonic)

P

Pais, A., 812*
Parabolic coordinates, *415*
Paramagnetic, *25*
Parity
  as an observable, *211*, 564, 653, 914, 921, 926, 1035
  intrinsic, 908*
  of spherical harmonics, *495*, 523
  of wave functions, *88*, *112–3*, 559
  (*see* Conservation)
Parseval relation, *175*, *186*, *268*, *270*, *476*
Partial waves method, *385–9*, *425–6*
Partition, 1112–4
  function, *338*
Paschen–Back effect, 708, 712, 742
Passage (rapid), 740, 759
  (*see* Adiabatic)
Pauli, W., 888
  exclusion principle, 541, 582, 598, 612, 693*, 702
  matrices, 544–5, 578, 620
  theory of the electron, 544, 935–8, 940
Pauling, L., 615*, 690*, 691
Peierls, R. E. (*see* Bohr)
Penetration factor, *391*, *430*
Periodic table, 614
Periodicity condition, *193*, 792
Permutation, 587, 655, 658, 1109–24
  operators, 587–93
Perrin, J., *8*
Perturbation,
  stationary, Chapter XVI, 983, 1003, 1040
  time-dependent, 724–39, 991
  periodic, 738–9
Phase,
  Coulomb, *426*
  extension in, *241*
  factor, *121*, *296*, 646, 648
  shift, *353*, *385–409*, *428–30*, 811
  space, *33*, *36*
  velocity, *50–55*
Phonon, *456*
Photoelectric effect, *11–13*, *16*, *41*, 760, 1007–9, 1052
Photon, *11–21*, *41*, *59*, *138–9*, *147–9*, *156–9*, *438–9*, 960, 1030
Placzek (*see* Bohr)
Planck, M., *10*, *11*
  constant, *10*
  law, *10*, 1030
Poisson, equation, 615, 815
  bracket, *285*, *318*, *362*, 645*, 650*
Polar coordinates (spherical), *344–8*
Polarization, 680, 1032–3, 1044
Positive definite (*see* Operator)
Positron, *60*, 953–6
Potential,
  barrier, *78*, *86*
  central, *344–5*, 817, 925–8
  one-dimensional square well, *78–98*
  (*see* Scattering, Coulomb)
Probability, *49*, *116–21*, *178*, *189–90*, *296–8*, *332–3*
  of transition, 725–8
  of transition per unit time, *28*, *74*, 734–6, 992, 1042
Product (*see* Scalar, Tensor)
Projection, *249*
Projection operators (projectors), *260*, *271*, *297–8*, *313*, *333*
  elementary, *262*
  (*see* Orthogonal)
Proton, *22*
Pure state, *205*, *335–6*

Q

Quadrupole (*see* Moment)
Quanta
  appearance of – in physics, *7–11*
  light, *11–21*
  (*see* Old Quantum Theory)
Quantization
  Bohr–Sommerfeld rules, *34–41*, *56*, *241*
  in space, *24–7*
  of electromagnetic radiation, 1029–37
  of the energy, *23–24*
  of the free scalar field, 960–79
  WKB rules, *240–1*

Quantum number
electronic, of a molecule, 790, 797
good, *204*, *211–2*, *319*
in the Old Quantum Theory, *38–40*, *241*
principal, *417*
radial, azimuthal, magnetic, *354*
spin, 613
Quasi degeneracy, 711–2

R

Racah, G., 562, 569*, 615*, 1053*, 1055*, 1058, 1065, 1075,* 1080
W coefficients, 567–8, 1061–8
Radial (*see* Equation, Momentum, Quantum number)
Radiative correction, 933, 936* (*see* Transition)
Radioactive, radioactivity, *8*, *28*, *403*
beta, 626, 662, 760
Radius,
of atoms, *9*, 618–9
of nuclei, *9*, 693, 728
of the classical electron, 1014–15, 1049*
(*see* Bohr)
Raleigh, Lord, *216**
Ramsauer–Townsend effect, 815
Range (*see* Effective)
Ratio (branching), 992
Reality (and time reversal) 671–3
"Reality", 677–9
Rearrangement lemma, 897, 1100
(*see* Collision)
Reduced mass (*see* Mass, Matrix)
Reducibility (*see* Irreducibility)
Reflection
space, *212*, 554, 604*, 653–4, 658, 907, 978, 1014 (*see* Parity)
wave (*see* Wave)
Relativistic, relativity, *7*, *13**, *31**, *38**, *47*, *53*, *65*, *135*, *151*, *383**, *419*, 552, Chapter XX, 959
approximation, non-relativistic limit, 886, 933–48, 980–3
(*see* Covariant)
Renormalization
of the charge, 1003*
of the mass, 990, 991, 1003, 1014
Representation, *119*, *125*, *244*, *281–90*, *323–4*
$\{N\}$, *437–8*
$\{p\}$, *326–9*
$\{Q\}$, *440–1*
Dirac, 890, 923
real, 672
standard, 516–8, 569, 572–3, 656, 1093
standard "real", 677
symmetrical, 586
"Representation"
FW (Foldy–Wouthuysen), 941–44
Heisenberg, *314–17*, *330**, *443*, 922
intermediate, *321–3*, 723
rotation axis, 745–7
Schrödinger, *310–14*, *323*, 920
Representation of a group, 1084–99
(*see* Irreducibility)
Resolvent, 712–15, 819, 994
Resonance, *88–96*, 738, 846
scattering, *392*, *396–403*, *430*, 1005–7, 1050
(*see* Width)
Retherford, R. C., 932*
Reversal, of a magnetic field, 742–3, 755–9
time (*see* Time)
Ritz (*see* Combination, Variational method)
Röhrlich, F., 960*, 1046*
Röntgen, *9*
Rosenfeld, L., *7**, 960*
Rotations, 523–40, 652, 658, 662
infinitesimal, 524, 531
of molecules, 782, 796–8
(*see* Degeneracy, Matrix, Operator)
Rotator, 695–7
Rupp, *57*
Russel–Saunders (*see* Coupling, *LS*)
Rutherford, D. E., 1109*
Rutherford, E., *8*, *21*
atom, *8*, *21–3*
formula, *230*, *423–4*, 609, 729, 730
Rydberg constant, *22*, *29–31*
(*see* Combination principle)

S
S (*see* Matrix)
Saxon, D. S., 857*
Scalar, *274*
  operator, 529, 531, 569, 1075
  product, *164*, *248–50*, *290*
Scattering, *9*, *18*, Chapter X, 736–8
  by a central potential, *385–409*, 857–60
  hard sphere, *393–5*
  inelastic, 832
  length, *342*, *408*, 861
  multiple 854–6
  of an electron by an atom, 815–17 841–3
  potential, *397*, 1005
  (*see* Amplitude, Collision, Coulomb, Resonance)
Scherrer, P. (*see* Debye)
Schiff, L. L., 814*
Schilpp, P. A., *48**
Schmidt orthogonalization process, *173*
Schrödinger, E., *46*, *47*
  equation, *59–71*, *310–12*
    time independent, *71–4*
  (*see* "Representation")
Schur's lemma, 646, 1090
Schwartz, L., *463**, *471**
  (*see* Distribution)
Schwartz inequality, *165*, *248*
Schwinger, J., 801*, 857, 860, 863*
Screen effect, 816
Secular equation, *278*, *287*, 718
Selection rules, *22*, 575, 576, 629, 731, 993, 1051, 1056, 1063, 1096
Self-consistent,
  Hartree and Fock–Dirac self-consistent fields, 612, 773–81, 785
Self energy, 1014–15, 1026
Separation of variables, Chapter IX
Series expansion
  of eigenfunctions, *174–6*, *184–8*
  of eigenvectors, *266–73*
Series expansion in spherical harmonics
  of Coulomb wave, *426–7*
  of Green's function, *497*, 829, 872
  of plane wave, *357–9*
Shell, complete, incomplete, 614
Shortley, G. H., 615*, 773*, 1055*
Similar particles, 586
Singlet, 558, 608
Singular (*see* Matrix)
Slater (*see* Determinant)
Smoluchovski, *8*
Sommerfeld, A., *28*, *36*, *38**
  (*see* Quantization)
Sommerfield, C. M., 933*, 936*
Space
  configuration, *31*, *68*, *164*
  dual, *246–7*
  Hilbert, *164*, *248*, 713*
  phase, *33*, *36*
  state, *295*, *301–9*
  quantization, *24–7*
Spectrum
  continuous, *179–94*, *259*, *268–71*
  discrete, *171–9*, *266–8*
  energy, *72–4*, *195*
Spherical
  harmonics, *348–9*, *494–7*, 520–3, 678–9, 1073–4
  (*see* Function, Symmetry)
Spin, 519, 540–55, 586, 888–9, 907, 921, 983
  coupling, spin-orbit coupling, 551–2, 705–8, 709, 793*
  isotopic, 619–29, 659
  of a nucleus, 609–10, 799
  of nucleons, 544
  of the electron, *419*, 540–44
Spinor, 546, 888, 907
Spreading (*see* Wave packet)
Square well
  one-dimensional, *86–96*
  spherical, *359–61*, *396–9*
Standard (*see* Representation)
Stark effect, 662–3, 695–6, 710, 794
State
  bound, *73*
  dynamical, *3*, *28*, *116*, *294*, *295*, *296*
  ground, *23*, *24*
  metastable (decaying), *399–403*, 994, 1005–7

stationary, *23*, *48*, *72*, *137*, 835, 1008
symmetrical, antisymmetrical, 595
virtual, 727
(*see* Space, Superposition, Vacuum state)
Statistical interpretation, *21*, *48**, *59*, Chapter IV, Chapter V
Statistics
Bose–Einstein, 597, 600, 965, 1030
classical *vs* quantum, *337–8*
Fermi–Dirac, 600, 954
Maxwell–Boltzmann, 597, 600
quantum, *204–5*, *331–8*
(*see* Mixture, Nucleus, Operator, density)
Stern, *57*
Stern–Gerlach experiment, *24–7*, 541, 543
Stone, H. M., *245**, 713*
Sturm–Liouville equation, *78*
Subspace, *246*
complementary, *248*
invariant, 514, 1088
of an eigenvalue, *258*
Sudden approximation, 740–4
Sunakawa, S., 801*
Superposition
of waves, *42*, *63*, *66*, *82*, *117*
principle of states, *244*, *295*, *310*
property of waves, *61*
Symmetrical (*see* State)
Symmetry,
and invariance, Chapter XV
point (*see* Reflection of space)
spherical, *343*
Symmetry with respect to permutations
operator, symmetrical observable, 589, 630
symmetrization postulate, 585, 593–6
symmetrizer, 589–93, 630, 1112
(*see* Young)

T

T (*see* Matrix)
Tensor,
force, 554, 564
irreducible tensor operator, 569–77, 580, 581, 681, 696*, 1075–8, 1092, 1098
metric, 877
operator, 570, 1075–8
operator S_{12}, 554
product, of matrices, *275*, *278–9*, *285*
product, of vector spaces, *252–4*, *285*, *307–9*, *1076*, *1088*
Ter-Martyrosian, K. A., 728*
Thermodynamic (*see* Equilibrium)
Thirring, W. E., 960*
Thomas–Fermi model, 612, 615–19, 630, 705*, 817
Thomson, G. P., *57*
Thomson, J. J., *7*
formula, 1048
Time, reversal, 521*, 632, 647, 649, 664–79, 785, 831, 915, 978–9, 1014
(*see* Uncertainty)
Trace
of a matrix, *276*, *285*
of an operator, *290*, *331–8*
partial, *279*
(*see* Invariance)
Transform (*see* Fourier)
Transformations
antiunitary, 639–40
FW (Foldy–Wouthuysen), 940–6
infinitesimal unitary, *292*, 649
linear, of matrices, *285–88*
of variables and dynamical states, 643–6
unitary, of matrices, *286–7*
unitary, of vectors and operators, *290–92*
(*see* Galilean, Group, Gauge, Lorentz)
Transition,
electric dipole, 989*, 993, 1043–45
magnetic dipole, 1043, 1052
radiative, 985–6, 991, 1041
(*see* Amplitude, Probability)
Translations (group of) 649–52, 658, 662, 851, 880, 975
Transmission coefficient, *85*, *95*, *97–8*, *109*, *237–9*
Transpose of a matrix, *274*

Triplet, 558, 608
"Tunnel" effect, *96–8*
Turning points, *234*

U

Uehlenbeck, G. E., 544
Uncertainty relations,
  Heisenberg, *116*, *129–49*, *214*, *299–301*, *319–20*
  position-momentum, *129–35*, *139–149*, *301*, *446*
  time-energy, *135–8*, *319–20*, *447*, 733, 742
  (*see* Measurement)
Unimodular, 526, 1089
Unitary (*see* Matrix, Operator, Transformation)

V

Vacuum state, *439*, 954–6, 983
Variational method, Chapter XVIII, 788, 797, 856–63
Vector,
  basis, *246*
  column, row, *274*
  operator, 509, 530, 532, 573, 1075
  wave, *14*, *50*, *56*
  (*see* Bra, Complete, Ket)
Vibrations of molecules, 782, 797–9
  (*see* Normal)
Villars, F., 886*, 888*
Virtual (*see* State)

W

Waerden, B. L. v. d., 1080*, 1109*
Wave,
  -corpuscle duality (*see* Duality)
  distorted, 822, 826
  electromagnetic, light, *6*, *9*
  equation (*see* Schrödinger equation)
  function, *60–1*, *324*
  incoming, outgoing, *356*, *372*, *386*, *390*, *423*, *426*, 802, 835
  length, *14*, *19*, *55* (*see* de Broglie)
  mechanics, *45–9*, *324–6*
  partial (*see* Decomposition)
  reflection, *80–6*, *91–6*, *106–9*
  stationary scattering, *371–2*, *421–2*, 802, 835
  vector (*see* Vector)
Wave packet, *50–55*, *66*, *81–5*, *92–3*, *97–8*, *106–8*, *130–2*, *136*, *138*, 604, 1002
  filtering of, *196–9*, *298*, *333**
  free, *50–3*, *221–2*, *328–9*
  Gaussian, *131*
  minimum, *160*, *215*, *446–8*
  scattering of, *372–80*, *401–3*
  spreading of, *58*, *218–22*, *329*
Weisskopf, V., 576*, 627*, 832*, 888, 1043*, 1055*
Wentzel, G., *216**
Wick, G. C., 536*
Width,
  of a line, 994–1001
  of a resonance, *397–8*, 1006
  (*see* Mean life)
Wightman, A., 536*
Wigner, E. P., *242*, 536*, 676, 1055*, 1080*
  6*j* coefficients, 567, 1061
  3*j* symbols, 1054
  -Eckart theorem, 573–5, 629, 706, 707, 731, 1044, 1076, 1097
Wilkinson, D. H., 627*
Wilson cloud chamber, *8*, 844, 856
Wilson, E. B., 614*, 690*, 691
Wilson, H. A., *36*
WKB method, *231–41*, *368*, *410*, 686
Wronskian, *98*
  "radial", 804
  theorem, *98–101*

Y

Young
  diagram, tableau, 1109*, 1113
  symmetrizers, 1114–7

Z

Zeeman effect, *27*, *39*, *40*, 706–7
Zitterbewegung, 951–2

A CATALOG OF SELECTED

DOVER BOOKS

IN SCIENCE AND MATHEMATICS



Astronomy

BURNHAM'S CELESTIAL HANDBOOK, Robert Burnham, Jr. Thorough guide to the stars beyond our solar system. Exhaustive treatment. Alphabetical by constellation: Andromeda to Cetus in Vol. 1; Chamaeleon to Orion in Vol. 2; and Pavo to Vulpecula in Vol. 3. Hundreds of illustrations. Index in Vol. 3. 2,000pp. 6⅛ x 9¼.
Vol. I: 0-486-23567-X
Vol. II: 0-486-23568-8
Vol. III: 0-486-23673-0

EXPLORING THE MOON THROUGH BINOCULARS AND SMALL TELESCOPES, Ernest H. Cherrington, Jr. Informative, profusely illustrated guide to locating and identifying craters, rills, seas, mountains, other lunar features. Newly revised and updated with special section of new photos. Over 100 photos and diagrams. 240pp. 8¼ x 11. 0-486-24491-1

THE EXTRATERRESTRIAL LIFE DEBATE, 1750–1900, Michael J. Crowe. First detailed, scholarly study in English of the many ideas that developed from 1750 to 1900 regarding the existence of intelligent extraterrestrial life. Examines ideas of Kant, Herschel, Voltaire, Percival Lowell, many other scientists and thinkers. 16 illustrations. 704pp. 5⅜ x 8½. 0-486-40675-X

THEORIES OF THE WORLD FROM ANTIQUITY TO THE COPERNICAN REVOLUTION, Michael J. Crowe. Newly revised edition of an accessible, enlightening book re-creates the change from an earth-centered to a sun-centered conception of the solar system. 242pp. 5⅜ x 8½. 0-486-41444-2

ARISTARCHUS OF SAMOS: The Ancient Copernicus, Sir Thomas Heath. Heath's history of astronomy ranges from Homer and Hesiod to Aristarchus and includes quotes from numerous thinkers, compilers, and scholasticists from Thales and Anaximander through Pythagoras, Plato, Aristotle, and Heraclides. 34 figures. 448pp. 5⅜ x 8½. 0-486-43886-4

A COMPLETE MANUAL OF AMATEUR ASTRONOMY: TOOLS AND TECHNIQUES FOR ASTRONOMICAL OBSERVATIONS, P. Clay Sherrod with Thomas L. Koed. Concise, highly readable book discusses: selecting, setting up and maintaining a telescope; amateur studies of the sun; lunar topography and occultations; observations of Mars, Jupiter, Saturn, the minor planets and the stars; an introduction to photoelectric photometry; more. 1981 ed. 124 figures. 25 halftones. 37 tables. 335pp. 6½ x 9¼. 0-486-42820-8

AMATEUR ASTRONOMER'S HANDBOOK, J. B. Sidgwick. Timeless, comprehensive coverage of telescopes, mirrors, lenses, mountings, telescope drives, micrometers, spectroscopes, more. 189 illustrations. 576pp. 5⅜ x 8¼. (Available in U.S. only.) 0-486-24034-7

STAR LORE: Myths, Legends, and Facts, William Tyler Olcott. Captivating retellings of the origins and histories of ancient star groups include Pegasus, Ursa Major, Pleiades, signs of the zodiac, and other constellations. "Classic."–Sky & Telescope. 58 illustrations. 544pp. 5⅜ x 8½. 0-486-43581-4

Chemistry

THE SCEPTICAL CHYMIST: THE CLASSIC 1661 TEXT, Robert Boyle. Boyle defines the term "element," asserting that all natural phenomena can be explained by the motion and organization of primary particles. 1911 ed. viii+232pp. 5⅜ x 8½.
0-486-42825-7

RADIOACTIVE SUBSTANCES, Marie Curie. Here is the celebrated scientist's doctoral thesis, the prelude to her receipt of the 1903 Nobel Prize. Curie discusses establishing atomic character of radioactivity found in compounds of uranium and thorium; extraction from pitchblende of polonium and radium; isolation of pure radium chloride; determination of atomic weight of radium; plus electric, photographic, luminous, heat, color effects of radioactivity. ii+94pp. 5⅜ x 8½.
0-486-42550-9

CHEMICAL MAGIC, Leonard A. Ford. Second Edition, Revised by E. Winston Grundmeier. Over 100 unusual stunts demonstrating cold fire, dust explosions, much more. Text explains scientific principles and stresses safety precautions. 128pp. 5⅜ x 8½. 0-486-67628-5

MOLECULAR THEORY OF CAPILLARITY, J. S. Rowlinson and B. Widom. History of surface phenomena offers critical and detailed examination and assessment of modern theories, focusing on statistical mechanics and application of results in mean-field approximation to model systems. 1989 edition. 352pp. 5⅜ x 8½.
0-486-42544-4

CHEMICAL AND CATALYTIC REACTION ENGINEERING, James J. Carberry. Designed to offer background for managing chemical reactions, this text examines behavior of chemical reactions and reactors; fluid-fluid and fluid-solid reaction systems; heterogeneous catalysis and catalytic kinetics; more. 1976 edition. 672pp. 6⅛ x 9¼. 0-486-41736-0 $31.95

ELEMENTS OF CHEMISTRY, Antoine Lavoisier. Monumental classic by founder of modern chemistry in remarkable reprint of rare 1790 Kerr translation. A must for every student of chemistry or the history of science. 539pp. 5⅜ x 8½. 0-486-64624-6

MOLECULES AND RADIATION: An Introduction to Modern Molecular Spectroscopy. Second Edition, Jeffrey I. Steinfeld. This unified treatment introduces upper-level undergraduates and graduate students to the concepts and the methods of molecular spectroscopy and applications to quantum electronics, lasers, and related optical phenomena. 1985 edition. 512pp. 5⅜ x 8½. 0-486-44152-0

A SHORT HISTORY OF CHEMISTRY, J. R. Partington. Classic exposition explores origins of chemistry, alchemy, early medical chemistry, nature of atmosphere, theory of valency, laws and structure of atomic theory, much more. 428pp. 5⅜ x 8½. (Available in U.S. only.) 0-486-65977-1

GENERAL CHEMISTRY, Linus Pauling. Revised 3rd edition of classic first-year text by Nobel laureate. Atomic and molecular structure, quantum mechanics, statistical mechanics, thermodynamics correlated with descriptive chemistry. Problems. 992pp. 5⅜ x 8½. 0-486-65622-5

ELECTRON CORRELATION IN MOLECULES, S. Wilson. This text addresses one of theoretical chemistry's central problems. Topics include molecular electronic structure, independent electron models, electron correlation, the linked diagram theorem, and related topics. 1984 edition. 304pp. 5⅜ x 8½. 0-486-45879-2

Engineering

DE RE METALLICA, Georgius Agricola. The famous Hoover translation of greatest treatise on technological chemistry, engineering, geology, mining of early modern times (1556). All 289 original woodcuts. 638pp. 6¾ x 11. 0-486-60006-8

FUNDAMENTALS OF ASTRODYNAMICS, Roger Bate et al. Modern approach developed by U.S. Air Force Academy. Designed as a first course. Problems, exercises. Numerous illustrations. 455pp. 5⅜ x 8½. 0-486-60061-0

DYNAMICS OF FLUIDS IN POROUS MEDIA, Jacob Bear. For advanced students of ground water hydrology, soil mechanics and physics, drainage and irrigation engineering and more. 335 illustrations. Exercises, with answers. 784pp. 6⅛ x 9¼. 0-486-65675-6

THEORY OF VISCOELASTICITY (SECOND EDITION), Richard M. Christensen. Complete consistent description of the linear theory of the viscoelastic behavior of materials. Problem-solving techniques discussed. 1982 edition. 29 figures. xiv+364pp. 6⅛ x 9¼. 0-486-42880-X

MECHANICS, J. P. Den Hartog. A classic introductory text or refresher. Hundreds of applications and design problems illuminate fundamentals of trusses, loaded beams and cables, etc. 334 answered problems. 462pp. 5⅜ x 8½. 0-486-60754-2

MECHANICAL VIBRATIONS, J. P. Den Hartog. Classic textbook offers lucid explanations and illustrative models, applying theories of vibrations to a variety of practical industrial engineering problems. Numerous figures. 233 problems, solutions. Appendix. Index. Preface. 436pp. 5⅜ x 8½. 0-486-64785-4

STRENGTH OF MATERIALS, J. P. Den Hartog. Full, clear treatment of basic material (tension, torsion, bending, etc.) plus advanced material on engineering methods, applications. 350 answered problems. 323pp. 5⅜ x 8½. 0-486-60755-0

A HISTORY OF MECHANICS, René Dugas. Monumental study of mechanical principles from antiquity to quantum mechanics. Contributions of ancient Greeks, Galileo, Leonardo, Kepler, Lagrange, many others. 671pp. 5⅜ x 8½. 0-486-65632-2

STABILITY THEORY AND ITS APPLICATIONS TO STRUCTURAL MECHANICS, Clive L. Dym. Self-contained text focuses on Koiter postbuckling analyses, with mathematical notions of stability of motion. Basing minimum energy principles for static stability upon dynamic concepts of stability of motion, it develops asymptotic buckling and postbuckling analyses from potential energy considerations, with applications to columns, plates, and arches. 1974 ed. 208pp. 5⅜ x 8½. 0-486-42541-X

BASIC ELECTRICITY, U.S. Bureau of Naval Personnel. Originally a training course; best nontechnical coverage. Topics include batteries, circuits, conductors, AC and DC, inductance and capacitance, generators, motors, transformers, amplifiers, etc. Many questions with answers. 349 illustrations. 1969 edition. 448pp. 6½ x 9¼. 0-486-20973-3

ROCKETS, Robert Goddard. Two of the most significant publications in the history of rocketry and jet propulsion: "A Method of Reaching Extreme Altitudes" (1919) and "Liquid Propellant Rocket Development" (1936). 128pp. 5⅛ x 8½. 0-486-42537-1

STATISTICAL MECHANICS: PRINCIPLES AND APPLICATIONS, Terrell L. Hill. Standard text covers fundamentals of statistical mechanics, applications to fluctuation theory, imperfect gases, distribution functions, more. 448pp. 5⅛ x 8½. 0-486-65390-0

ENGINEERING AND TECHNOLOGY 1650–1750: ILLUSTRATIONS AND TEXTS FROM ORIGINAL SOURCES, Martin Jensen. Highly readable text with more than 200 contemporary drawings and detailed engravings of engineering projects dealing with surveying, leveling, materials, hand tools, lifting equipment, transport and erection, piling, bailing, water supply, hydraulic engineering, and more. Among the specific projects outlined-transporting a 50-ton stone to the Louvre, erecting an obelisk, building timber locks, and dredging canals. 207pp. 8⅜ x 11¼. 0-486-42232-1

THE VARIATIONAL PRINCIPLES OF MECHANICS, Cornelius Lanczos. Graduate level coverage of calculus of variations, equations of motion, relativistic mechanics, more. First inexpensive paperbound edition of classic treatise. Index. Bibliography. 418pp. 5⅛ x 8½. 0-486-65067-7

PROTECTION OF ELECTRONIC CIRCUITS FROM OVERVOLTAGES, Ronald B. Standler. Five-part treatment presents practical rules and strategies for circuits designed to protect electronic systems from damage by transient overvoltages. 1989 ed. xxiv+434pp. 6⅛ x 9¼. 0-486-42552-5

ROTARY WING AERODYNAMICS, W. Z. Stepniewski. Clear, concise text covers aerodynamic phenomena of the rotor and offers guidelines for helicopter performance evaluation. Originally prepared for NASA. 537 figures. 640pp. 6⅛ x 9¼. 0-486-64647-5

INTRODUCTION TO SPACE DYNAMICS, William Tyrrell Thomson. Comprehensive, classic introduction to space-flight engineering for advanced undergraduate and graduate students. Includes vector algebra, kinematics, transformation of coordinates. Bibliography. Index. 352pp. 5⅛ x 8½. 0-486-65113-4

HISTORY OF STRENGTH OF MATERIALS, Stephen P. Timoshenko. Excellent historical survey of the strength of materials with many references to the theories of elasticity and structure. 245 figures. 452pp. 5⅛ x 8½. 0-486-61187-6

ANALYTICAL FRACTURE MECHANICS, David J. Unger. Self-contained text supplements standard fracture mechanics texts by focusing on analytical methods for determining crack-tip stress and strain fields. 336pp. 6⅛ x 9¼. 0-486-41737-9

STATISTICAL MECHANICS OF ELASTICITY, J. H. Weiner. Advanced, self-contained treatment illustrates general principles and elastic behavior of solids. Part 1, based on classical mechanics, studies thermoelastic behavior of crystalline and polymeric solids. Part 2, based on quantum mechanics, focuses on interatomic force laws, behavior of solids, and thermally activated processes. For students of physics and chemistry and for polymer physicists. 1983 ed. 96 figures. 496pp. 5⅛ x 8½. 0-486-42260-7

Mathematics

FUNCTIONAL ANALYSIS (Second Corrected Edition), George Bachman and Lawrence Narici. Excellent treatment of subject geared toward students with background in linear algebra, advanced calculus, physics and engineering. Text covers introduction to inner-product spaces, normed, metric spaces, and topological spaces; complete orthonormal sets, the Hahn-Banach Theorem and its consequences, and many other related subjects. 1966 ed. 544pp. 6⅛ x 9¼. 0-486-40251-7

DIFFERENTIAL MANIFOLDS, Antoni A. Kosinski. Introductory text for advanced undergraduates and graduate students presents systematic study of the topological structure of smooth manifolds, starting with elements of theory and concluding with method of surgery. 1993 edition. 288pp. 5⅜ x 8½. 0-486-46244-7

VECTOR AND TENSOR ANALYSIS WITH APPLICATIONS, A. I. Borisenko and I. E. Tarapov. Concise introduction. Worked-out problems, solutions, exercises. 257pp. 5⅜ x 8¼. 0-486-63833-2

AN INTRODUCTION TO ORDINARY DIFFERENTIAL EQUATIONS, Earl A. Coddington. A thorough and systematic first course in elementary differential equations for undergraduates in mathematics and science, with many exercises and problems (with answers). Index. 304pp. 5⅜ x 8½. 0-486-65942-9

FOURIER SERIES AND ORTHOGONAL FUNCTIONS, Harry F. Davis. An incisive text combining theory and practical example to introduce Fourier series, orthogonal functions and applications of the Fourier method to boundary-value problems. 570 exercises. Answers and notes. 416pp. 5⅜ x 8½. 0-486-65973-9

COMPUTABILITY AND UNSOLVABILITY, Martin Davis. Classic graduate-level introduction to theory of computability, usually referred to as theory of recurrent functions. New preface and appendix. 288pp. 5⅜ x 8½. 0-486-61471-9

AN INTRODUCTION TO MATHEMATICAL ANALYSIS, Robert A. Rankin. Dealing chiefly with functions of a single real variable, this text by a distinguished educator introduces limits, continuity, differentiability, integration, convergence of infinite series, double series, and infinite products. 1963 edition. 624pp. 5⅜ x 8½. 0-486-46251-X

METHODS OF NUMERICAL INTEGRATION (SECOND EDITION), Philip J. Davis and Philip Rabinowitz. Requiring only a background in calculus, this text covers approximate integration over finite and infinite intervals, error analysis, approximate integration in two or more dimensions, and automatic integration. 1984 edition. 624pp. 5⅜ x 8½. 0-486-45339-1

INTRODUCTION TO LINEAR ALGEBRA AND DIFFERENTIAL EQUATIONS, John W. Dettman. Excellent text covers complex numbers, determinants, orthonormal bases, Laplace transforms, much more. Exercises with solutions. Undergraduate level. 416pp. 5⅜ x 8½. 0-486-65191-6

RIEMANN'S ZETA FUNCTION, H. M. Edwards. Superb, high-level study of landmark 1859 publication entitled "On the Number of Primes Less Than a Given Magnitude" traces developments in mathematical theory that it inspired. xiv+315pp. 5⅜ x 8½. 0-486-41740-9

CALCULUS OF VARIATIONS WITH APPLICATIONS, George M. Ewing. Applications-oriented introduction to variational theory develops insight and promotes understanding of specialized books, research papers. Suitable for advanced undergraduate/graduate students as primary, supplementary text. 352pp. 5⅜ x 8½.
0-486-64856-7

MATHEMATICIAN'S DELIGHT, W. W. Sawyer. "Recommended with confidence" by *The Times Literary Supplement,* this lively survey was written by a renowned teacher. It starts with arithmetic and algebra, gradually proceeding to trigonometry and calculus. 1943 edition. 240pp. 5⅜ x 8½. 0-486-46240-4

ADVANCED EUCLIDEAN GEOMETRY, Roger A. Johnson. This classic text explores the geometry of the triangle and the circle, concentrating on extensions of Euclidean theory, and examining in detail many relatively recent theorems. 1929 edition. 336pp. 5⅜ x 8½. 0-486-46237-4

COUNTEREXAMPLES IN ANALYSIS, Bernard R. Gelbaum and John M. H. Olmsted. These counterexamples deal mostly with the part of analysis known as "real variables." The first half covers the real number system, and the second half encompasses higher dimensions. 1962 edition. xxiv+198pp. 5⅜ x 8½. 0-486-42875-3

CATASTROPHE THEORY FOR SCIENTISTS AND ENGINEERS, Robert Gilmore. Advanced-level treatment describes mathematics of theory grounded in the work of Poincaré, R. Thom, other mathematicians. Also important applications to problems in mathematics, physics, chemistry and engineering. 1981 edition. References. 28 tables. 397 black-and-white illustrations. xvii + 666pp. 6⅛ x 9¼.
0-486-67539-4

COMPLEX VARIABLES: Second Edition, Robert B. Ash and W. P. Novinger. Suitable for advanced undergraduates and graduate students, this newly revised treatment covers Cauchy theorem and its applications, analytic functions, and the prime number theorem. Numerous problems and solutions. 2004 edition. 224pp. 6½ x 9¼. 0-486-46250-1

NUMERICAL METHODS FOR SCIENTISTS AND ENGINEERS, Richard Hamming. Classic text stresses frequency approach in coverage of algorithms, polynomial approximation, Fourier approximation, exponential approximation, other topics. Revised and enlarged 2nd edition. 721pp. 5⅜ x 8½. 0-486-65241-6

INTRODUCTION TO NUMERICAL ANALYSIS (2nd Edition), F. B. Hildebrand. Classic, fundamental treatment covers computation, approximation, interpolation, numerical differentiation and integration, other topics. 150 new problems. 669pp. 5⅜ x 8½. 0-486-65363-3

MARKOV PROCESSES AND POTENTIAL THEORY, Robert M. Blumental and Ronald K. Getoor. This graduate-level text explores the relationship between Markov processes and potential theory in terms of excessive functions, multiplicative functionals and subprocesses, additive functionals and their potentials, and dual processes. 1968 edition. 320pp. 5⅜ x 8½. 0-486-46263-3

ABSTRACT SETS AND FINITE ORDINALS: An Introduction to the Study of Set Theory, G. B. Keene. This text unites logical and philosophical aspects of set theory in a manner intelligible to mathematicians without training in formal logic and to logicians without a mathematical background. 1961 edition. 112pp. 5⅜ x 8½.
0-486-46249-8

INTRODUCTORY REAL ANALYSIS, A.N. Kolmogorov, S. V. Fomin. Translated by Richard A. Silverman. Self-contained, evenly paced introduction to real and functional analysis. Some 350 problems. 403pp. 5⅜ x 8½. 0-486-61226-0

APPLIED ANALYSIS, Cornelius Lanczos. Classic work on analysis and design of finite processes for approximating solution of analytical problems. Algebraic equations, matrices, harmonic analysis, quadrature methods, much more. 559pp. 5⅜ x 8½. 0-486-65656-X

AN INTRODUCTION TO ALGEBRAIC STRUCTURES, Joseph Landin. Superb self-contained text covers "abstract algebra": sets and numbers, theory of groups, theory of rings, much more. Numerous well-chosen examples, exercises. 247pp. 5⅜ x 8½. 0-486-65940-2

QUALITATIVE THEORY OF DIFFERENTIAL EQUATIONS, V. V. Nemytskii and V.V. Stepanov. Classic graduate-level text by two prominent Soviet mathematicians covers classical differential equations as well as topological dynamics and ergodic theory. Bibliographies. 523pp. 5⅜ x 8½. 0-486-65954-2

THEORY OF MATRICES, Sam Perlis. Outstanding text covering rank, nonsingularity and inverses in connection with the development of canonical matrices under the relation of equivalence, and without the intervention of determinants. Includes exercises. 237pp. 5⅜ x 8½. 0-486-66810-X

INTRODUCTION TO ANALYSIS, Maxwell Rosenlicht. Unusually clear, accessible coverage of set theory, real number system, metric spaces, continuous functions, Riemann integration, multiple integrals, more. Wide range of problems. Undergraduate level. Bibliography. 254pp. 5⅜ x 8½. 0-486-65038-3

MODERN NONLINEAR EQUATIONS, Thomas L. Saaty. Emphasizes practical solution of problems; covers seven types of equations. ". . . a welcome contribution to the existing literature. . . ."–*Math Reviews*. 490pp. 5⅜ x 8½. 0-486-64232-1

MATRICES AND LINEAR ALGEBRA, Hans Schneider and George Phillip Barker. Basic textbook covers theory of matrices and its applications to systems of linear equations and related topics such as determinants, eigenvalues and differential equations. Numerous exercises. 432pp. 5⅜ x 8½. 0-486-66014-1

LINEAR ALGEBRA, Georgi E. Shilov. Determinants, linear spaces, matrix algebras, similar topics. For advanced undergraduates, graduates. Silverman translation. 387pp. 5⅜ x 8½. 0-486-63518-X

MATHEMATICAL METHODS OF GAME AND ECONOMIC THEORY: Revised Edition, Jean-Pierre Aubin. This text begins with optimization theory and convex analysis, followed by topics in game theory and mathematical economics, and concluding with an introduction to nonlinear analysis and control theory. 1982 edition. 656pp. 6⅛ x 9¼. 0-486-46265-X

SET THEORY AND LOGIC, Robert R. Stoll. Lucid introduction to unified theory of mathematical concepts. Set theory and logic seen as tools for conceptual understanding of real number system. 496pp. 5⅝ x 8¼. 0-486-63829-4

TENSOR CALCULUS, J.L. Synge and A. Schild. Widely used introductory text covers spaces and tensors, basic operations in Riemannian space, non-Riemannian spaces, etc. 324pp. 5⅜ x 8¼. 0-486-63612-7

ORDINARY DIFFERENTIAL EQUATIONS, Morris Tenenbaum and Harry Pollard. Exhaustive survey of ordinary differential equations for undergraduates in mathematics, engineering, science. Thorough analysis of theorems. Diagrams. Bibliography. Index. 818pp. 5⅜ x 8½. 0-486-64940-7

INTEGRAL EQUATIONS, F. G. Tricomi. Authoritative, well-written treatment of extremely useful mathematical tool with wide applications. Volterra Equations, Fredholm Equations, much more. Advanced undergraduate to graduate level. Exercises. Bibliography. 238pp. 5⅜ x 8½. 0-486-64828-1

FOURIER SERIES, Georgi P. Tolstov. Translated by Richard A. Silverman. A valuable addition to the literature on the subject, moving clearly from subject to subject and theorem to theorem. 107 problems, answers. 336pp. 5⅜ x 8½. 0-486-63317-9

INTRODUCTION TO MATHEMATICAL THINKING, Friedrich Waismann. Examinations of arithmetic, geometry, and theory of integers; rational and natural numbers; complete induction; limit and point of accumulation; remarkable curves; complex and hypercomplex numbers, more. 1959 ed. 27 figures. xii+260pp. 5⅜ x 8½. 0-486-42804-8

THE RADON TRANSFORM AND SOME OF ITS APPLICATIONS, Stanley R. Deans. Of value to mathematicians, physicists, and engineers, this excellent introduction covers both theory and applications, including a rich array of examples and literature. Revised and updated by the author. 1993 edition. 304pp. 6⅛ x 9¼. 0-486-46241-2

CALCULUS OF VARIATIONS, Robert Weinstock. Basic introduction covering isoperimetric problems, theory of elasticity, quantum mechanics, electrostatics, etc. Exercises throughout. 326pp. 5⅜ x 8½. 0-486-63069-2

THE CONTINUUM: A CRITICAL EXAMINATION OF THE FOUNDATION OF ANALYSIS, Hermann Weyl. Classic of 20th-century foundational research deals with the conceptual problem posed by the continuum. 156pp. 5⅜ x 8½. 0-486-67982-9

CHALLENGING MATHEMATICAL PROBLEMS WITH ELEMENTARY SOLUTIONS, A. M. Yaglom and I. M. Yaglom. Over 170 challenging problems on probability theory, combinatorial analysis, points and lines, topology, convex polygons, many other topics. Solutions. Total of 445pp. 5⅜ x 8½. Two-vol. set. Vol. I: 0-486-65536-9 Vol. II: 0-486-65537-7

INTRODUCTION TO PARTIAL DIFFERENTIAL EQUATIONS WITH APPLICATIONS, E. C. Zachmanoglou and Dale W. Thoe. Essentials of partial differential equations applied to common problems in engineering and the physical sciences. Problems and answers. 416pp. 5⅜ x 8½. 0-486-65251-3

STOCHASTIC PROCESSES AND FILTERING THEORY, Andrew H. Jazwinski. This unified treatment presents material previously available only in journals, and in terms accessible to engineering students. Although theory is emphasized, it discusses numerous practical applications as well. 1970 edition. 400pp. 5⅜ x 8½. 0-486-46274-9

Math–Decision Theory, Statistics, Probability

INTRODUCTION TO PROBABILITY, John E. Freund. Featured topics include permutations and factorials, probabilities and odds, frequency interpretation, mathematical expectation, decision-making, postulates of probability, rule of elimination, much more. Exercises with some solutions. Summary. 1973 edition. 247pp. 5⅜ x 8½. 0-486-67549-1

STATISTICAL AND INDUCTIVE PROBABILITIES, Hugues Leblanc. This treatment addresses a decades-old dispute among probability theorists, asserting that both statistical and inductive probabilities may be treated as sentence-theoretic measurements, and that the latter qualify as estimates of the former. 1962 edition. 160pp. 5⅜ x 8½. 0-486-44980-7

APPLIED MULTIVARIATE ANALYSIS: Using Bayesian and Frequentist Methods of Inference, Second Edition, S. James Press. This two-part treatment deals with foundations as well as models and applications. Topics include continuous multivariate distributions; regression and analysis of variance; factor analysis and latent structure analysis; and structuring multivariate populations. 1982 edition. 692pp. 5⅜ x 8½. 0-486-44236-5

LINEAR PROGRAMMING AND ECONOMIC ANALYSIS, Robert Dorfman, Paul A. Samuelson and Robert M. Solow. First comprehensive treatment of linear programming in standard economic analysis. Game theory, modern welfare economics, Leontief input-output, more. 525pp. 5⅜ x 8½. 0-486-65491-5

PROBABILITY: AN INTRODUCTION, Samuel Goldberg. Excellent basic text covers set theory, probability theory for finite sample spaces, binomial theorem, much more. 360 problems. Bibliographies. 322pp. 5⅜ x 8½. 0-486-65252-1

GAMES AND DECISIONS: INTRODUCTION AND CRITICAL SURVEY, R. Duncan Luce and Howard Raiffa. Superb nontechnical introduction to game theory, primarily applied to social sciences. Utility theory, zero-sum games, n-person games, decision-making, much more. Bibliography. 509pp. 5⅜ x 8½. 0-486-65943-7

INTRODUCTION TO THE THEORY OF GAMES, J. C. C. McKinsey. This comprehensive overview of the mathematical theory of games illustrates applications to situations involving conflicts of interest, including economic, social, political, and military contexts. Appropriate for advanced undergraduate and graduate courses; advanced calculus a prerequisite. 1952 ed. x+372pp. 5⅜ x 8½. 0-486-42811-7

FIFTY CHALLENGING PROBLEMS IN PROBABILITY WITH SOLUTIONS, Frederick Mosteller. Remarkable puzzlers, graded in difficulty, illustrate elementary and advanced aspects of probability. Detailed solutions. 88pp. 5⅜ x 8½. 0-486-65355-2

PROBABILITY THEORY: A CONCISE COURSE, Y. A. Rozanov. Highly readable, self-contained introduction covers combination of events, dependent events, Bernoulli trials, etc. 148pp. 5⅜ x 8¼. 0-486-63544-9

THE STATISTICAL ANALYSIS OF EXPERIMENTAL DATA, John Mandel. First half of book presents fundamental mathematical definitions, concepts and facts while remaining half deals with statistics primarily as an interpretive tool. Well-written text, numerous worked examples with step-by-step presentation. Includes 116 tables. 448pp. 5⅜ x 8½. 0-486-64666-1

Math–Geometry and Topology

ELEMENTARY CONCEPTS OF TOPOLOGY, Paul Alexandroff. Elegant, intuitive approach to topology from set-theoretic topology to Betti groups; how concepts of topology are useful in math and physics. 25 figures. 57pp. 5⅜ x 8½. 0-486-60747-X

A LONG WAY FROM EUCLID, Constance Reid. Lively guide by a prominent historian focuses on the role of Euclid's Elements in subsequent mathematical developments. Elementary algebra and plane geometry are sole prerequisites. 80 drawings. 1963 edition. 304pp. 5⅜ x 8½. 0-486-43613-6

EXPERIMENTS IN TOPOLOGY, Stephen Barr. Classic, lively explanation of one of the byways of mathematics. Klein bottles, Moebius strips, projective planes, map coloring, problem of the Koenigsberg bridges, much more, described with clarity and wit. 43 figures. 210pp. 5⅜ x 8½. 0-486-25933-1

THE GEOMETRY OF RENÉ DESCARTES, René Descartes. The great work founded analytical geometry. Original French text, Descartes's own diagrams, together with definitive Smith-Latham translation. 244pp. 5⅜ x 8½. 0-486-60068-8

EUCLIDEAN GEOMETRY AND TRANSFORMATIONS, Clayton W. Dodge. This introduction to Euclidean geometry emphasizes transformations, particularly isometries and similarities. Suitable for undergraduate courses, it includes numerous examples, many with detailed answers. 1972 ed. viii+296pp. 6⅛ x 9¼. 0-486-43476-1

EXCURSIONS IN GEOMETRY, C. Stanley Ogilvy. A straightedge, compass, and a little thought are all that's needed to discover the intellectual excitement of geometry. Harmonic division and Apollonian circles, inversive geometry, hexlet, Golden Section, more. 132 illustrations. 192pp. 5⅜ x 8½. 0-486-26530-7

THE THIRTEEN BOOKS OF EUCLID'S ELEMENTS, translated with introduction and commentary by Sir Thomas L. Heath. Definitive edition. Textual and linguistic notes, mathematical analysis. 2,500 years of critical commentary. Unabridged. 1,414pp. 5⅜ x 8½. Three-vol. set.
Vol. I: 0-486-60088-2 Vol. II: 0-486-60089-0 Vol. III: 0-486-60090-4

SPACE AND GEOMETRY: IN THE LIGHT OF PHYSIOLOGICAL, PSYCHOLOGICAL AND PHYSICAL INQUIRY, Ernst Mach. Three essays by an eminent philosopher and scientist explore the nature, origin, and development of our concepts of space, with a distinctness and precision suitable for undergraduate students and other readers. 1906 ed. vi+148pp. 5⅜ x 8½. 0-486-43909-7

GEOMETRY OF COMPLEX NUMBERS, Hans Schwerdtfeger. Illuminating, widely praised book on analytic geometry of circles, the Moebius transformation, and two-dimensional non-Euclidean geometries. 200pp. 5⅜ x 8¼. 0-486-63830-8

DIFFERENTIAL GEOMETRY, Heinrich W. Guggenheimer. Local differential geometry as an application of advanced calculus and linear algebra. Curvature, transformation groups, surfaces, more. Exercises. 62 figures. 378pp. 5⅜ x 8½. 0-486-63433-7

History of Math

THE WORKS OF ARCHIMEDES, Archimedes (T. L. Heath, ed.). Topics include the famous problems of the ratio of the areas of a cylinder and an inscribed sphere; the measurement of a circle; the properties of conoids, spheroids, and spirals; and the quadrature of the parabola. Informative introduction. clxxxvi+326pp. 5⅜ x 8½.
0-486-42084-1

A SHORT ACCOUNT OF THE HISTORY OF MATHEMATICS, W. W. Rouse Ball. One of clearest, most authoritative surveys from the Egyptians and Phoenicians through 19th-century figures such as Grassman, Galois, Riemann. Fourth edition. 522pp. 5⅜ x 8½. 0-486-20630-0

THE HISTORY OF THE CALCULUS AND ITS CONCEPTUAL DEVELOPMENT, Carl B. Boyer. Origins in antiquity, medieval contributions, work of Newton, Leibniz, rigorous formulation. Treatment is verbal. 346pp. 5⅜ x 8½. 0-486-60509-4

THE HISTORICAL ROOTS OF ELEMENTARY MATHEMATICS, Lucas N. H. Bunt, Phillip S. Jones, and Jack D. Bedient. Fundamental underpinnings of modern arithmetic, algebra, geometry and number systems derived from ancient civilizations. 320pp. 5⅜ x 8½. 0-486-25563-8

THE HISTORY OF THE CALCULUS AND ITS CONCEPTUAL DEVELOPMENT, Carl B. Boyer. Fluent description of the development of both the integral and differential calculus–its early beginnings in antiquity, medieval contributions, and a consideration of Newton and Leibniz. 368pp. 5⅜ x 8½. 0-486-60509-4

GAMES, GODS & GAMBLING: A HISTORY OF PROBABILITY AND STATISTICAL IDEAS, F. N. David. Episodes from the lives of Galileo, Fermat, Pascal, and others illustrate this fascinating account of the roots of mathematics. Features thought-provoking references to classics, archaeology, biography, poetry. 1962 edition. 304pp. 5⅜ x 8½. (Available in U.S. only.) 0-486-40023-9

OF MEN AND NUMBERS: THE STORY OF THE GREAT MATHEMATICIANS, Jane Muir. Fascinating accounts of the lives and accomplishments of history's greatest mathematical minds–Pythagoras, Descartes, Euler, Pascal, Cantor, many more. Anecdotal, illuminating. 30 diagrams. Bibliography. 256pp. 5⅜ x 8½.
0-486-28973-7

HISTORY OF MATHEMATICS, David E. Smith. Nontechnical survey from ancient Greece and Orient to late 19th century; evolution of arithmetic, geometry, trigonometry, calculating devices, algebra, the calculus. 362 illustrations. 1,355pp. 5⅜ x 8½. Two-vol. set. Vol. I: 0-486-20429-4 Vol. II: 0-486-20430-8

A CONCISE HISTORY OF MATHEMATICS, Dirk J. Struik. The best brief history of mathematics. Stresses origins and covers every major figure from ancient Near East to 19th century. 41 illustrations. 195pp. 5⅜ x 8½. 0-486-60255-9

Physics

OPTICAL RESONANCE AND TWO-LEVEL ATOMS, L. Allen and J. H. Eberly. Clear, comprehensive introduction to basic principles behind all quantum optical resonance phenomena. 53 illustrations. Preface. Index. 256pp. 5⅜ x 8½.
0-486-65533-4

QUANTUM THEORY, David Bohm. This advanced undergraduate-level text presents the quantum theory in terms of qualitative and imaginative concepts, followed by specific applications worked out in mathematical detail. Preface. Index. 655pp. 5⅜ x 8½. 0-486-65969-0

ATOMIC PHYSICS (8th EDITION), Max Born. Nobel laureate's lucid treatment of kinetic theory of gases, elementary particles, nuclear atom, wave-corpuscles, atomic structure and spectral lines, much more. Over 40 appendices, bibliography. 495pp. 5⅜ x 8½. 0-486-65984-4

A SOPHISTICATE'S PRIMER OF RELATIVITY, P. W. Bridgman. Geared toward readers already acquainted with special relativity, this book transcends the view of theory as a working tool to answer natural questions: What is a frame of reference? What is a "law of nature"? What is the role of the "observer"? Extensive treatment, written in terms accessible to those without a scientific background. 1983 ed. xlviii+172pp. 5⅜ x 8½. 0-486-42549-5

AN INTRODUCTION TO HAMILTONIAN OPTICS, H. A. Buchdahl. Detailed account of the Hamiltonian treatment of aberration theory in geometrical optics. Many classes of optical systems defined in terms of the symmetries they possess. Problems with detailed solutions. 1970 edition. xv + 360pp. 5⅜ x 8½. 0-486-67597-1

PRIMER OF QUANTUM MECHANICS, Marvin Chester. Introductory text examines the classical quantum bead on a track: its state and representations; operator eigenvalues; harmonic oscillator and bound bead in a symmetric force field; and bead in a spherical shell. Other topics include spin, matrices, and the structure of quantum mechanics; the simplest atom; indistinguishable particles; and stationary-state perturbation theory. 1992 ed. xiv+314pp. 6⅛ x 9¼. 0-486-42878-8

LECTURES ON QUANTUM MECHANICS, Paul A. M. Dirac. Four concise, brilliant lectures on mathematical methods in quantum mechanics from Nobel Prize-winning quantum pioneer build on idea of visualizing quantum theory through the use of classical mechanics. 96pp. 5⅜ x 8½. 0-486-41713-1

THIRTY YEARS THAT SHOOK PHYSICS: THE STORY OF QUANTUM THEORY, George Gamow. Lucid, accessible introduction to influential theory of energy and matter. Careful explanations of Dirac's anti-particles, Bohr's model of the atom, much more. 12 plates. Numerous drawings. 240pp. 5⅜ x 8½. 0-486-24895-X

ELECTRONIC STRUCTURE AND THE PROPERTIES OF SOLIDS: THE PHYSICS OF THE CHEMICAL BOND, Walter A. Harrison. Innovative text offers basic understanding of the electronic structure of covalent and ionic solids, simple metals, transition metals and their compounds. Problems. 1980 edition. 582pp. 6⅛ x 9¼. 0-486-66021-4

HYDRODYNAMIC AND HYDROMAGNETIC STABILITY, S. Chandrasekhar. Lucid examination of the Rayleigh-Benard problem; clear coverage of the theory of instabilities causing convection. 704pp. 5⅜ x 8¼. 0-486-64071-X

INVESTIGATIONS ON THE THEORY OF THE BROWNIAN MOVEMENT, Albert Einstein. Five papers (1905–8) investigating dynamics of Brownian motion and evolving elementary theory. Notes by R. Fürth. 122pp. 5⅜ x 8½. 0-486-60304-0

THE PHYSICS OF WAVES, William C. Elmore and Mark A. Heald. Unique overview of classical wave theory. Acoustics, optics, electromagnetic radiation, more. Ideal as classroom text or for self-study. Problems. 477pp. 5⅜ x 8½. 0-486-64926-1

GRAVITY, George Gamow. Distinguished physicist and teacher takes reader-friendly look at three scientists whose work unlocked many of the mysteries behind the laws of physics: Galileo, Newton, and Einstein. Most of the book focuses on Newton's ideas, with a concluding chapter on post-Einsteinian speculations concerning the relationship between gravity and other physical phenomena. 160pp. 5⅜ x 8½. 0-486-42563-0

PHYSICAL PRINCIPLES OF THE QUANTUM THEORY, Werner Heisenberg. Nobel Laureate discusses quantum theory, uncertainty, wave mechanics, work of Dirac, Schroedinger, Compton, Wilson, Einstein, etc. 184pp. 5⅜ x 8½. 0-486-60113-7

ATOMIC SPECTRA AND ATOMIC STRUCTURE, Gerhard Herzberg. One of best introductions; especially for specialist in other fields. Treatment is physical rather than mathematical. 80 illustrations. 257pp. 5⅜ x 8½. 0-486-60115-3

AN INTRODUCTION TO STATISTICAL THERMODYNAMICS, Terrell L. Hill. Excellent basic text offers wide-ranging coverage of quantum statistical mechanics, systems of interacting molecules, quantum statistics, more. 523pp. 5⅜ x 8½. 0-486-65242-4

THEORETICAL PHYSICS, Georg Joos, with Ira M. Freeman. Classic overview covers essential math, mechanics, electromagnetic theory, thermodynamics, quantum mechanics, nuclear physics, other topics. First paperback edition. xxiii + 885pp. 5⅜ x 8½. 0-486-65227-0

PROBLEMS AND SOLUTIONS IN QUANTUM CHEMISTRY AND PHYSICS, Charles S. Johnson, Jr. and Lee G. Pedersen. Unusually varied problems, detailed solutions in coverage of quantum mechanics, wave mechanics, angular momentum, molecular spectroscopy, more. 280 problems plus 139 supplementary exercises. 430pp. 6½ x 9¼. 0-486-65236-X

THEORETICAL SOLID STATE PHYSICS, Vol. 1: Perfect Lattices in Equilibrium; Vol. II: Non-Equilibrium and Disorder, William Jones and Norman H. March. Monumental reference work covers fundamental theory of equilibrium properties of perfect crystalline solids, non-equilibrium properties, defects and disordered systems. Appendices. Problems. Preface. Diagrams. Index. Bibliography. Total of 1,301pp. 5⅜ x 8½. Two volumes. Vol. I: 0-486-65015-4 Vol. II: 0-486-65016-2

WHAT IS RELATIVITY? L. D. Landau and G. B. Rumer. Written by a Nobel Prize physicist and his distinguished colleague, this compelling book explains the special theory of relativity to readers with no scientific background, using such familiar objects as trains, rulers, and clocks. 1960 ed. vi+72pp. 5⅜ x 8½. 0-486-42806-0

A TREATISE ON ELECTRICITY AND MAGNETISM, James Clerk Maxwell. Important foundation work of modern physics. Brings to final form Maxwell's theory of electromagnetism and rigorously derives his general equations of field theory. 1,084pp. 5⅜ x 8½. Two-vol. set. Vol. I: 0-486-60636-8 Vol. II: 0-486-60637-6

MATHEMATICS FOR PHYSICISTS, Philippe Dennery and Andre Krzywicki. Superb text provides math needed to understand today's more advanced topics in physics and engineering. Theory of functions of a complex variable, linear vector spaces, much more. Problems. 1967 edition. 400pp. 6½ x 9¼. 0-486-69193-4

INTRODUCTION TO QUANTUM MECHANICS WITH APPLICATIONS TO CHEMISTRY, Linus Pauling & E. Bright Wilson, Jr. Classic undergraduate text by Nobel Prize winner applies quantum mechanics to chemical and physical problems. Numerous tables and figures enhance the text. Chapter bibliographies. Appendices. Index. 468pp. 5⅜ x 8½. 0-486-64871-0

METHODS OF THERMODYNAMICS, Howard Reiss. Outstanding text focuses on physical technique of thermodynamics, typical problem areas of understanding, and significance and use of thermodynamic potential. 1965 edition. 238pp. 5⅜ x 8½. 0-486-69445-3

THE ELECTROMAGNETIC FIELD, Albert Shadowitz. Comprehensive undergraduate text covers basics of electric and magnetic fields, builds up to electromagnetic theory. Also related topics, including relativity. Over 900 problems. 768pp. 5⅜ x 8¼. 0-486-65660-8

GREAT EXPERIMENTS IN PHYSICS: FIRSTHAND ACCOUNTS FROM GALILEO TO EINSTEIN, Morris H. Shamos (ed.). 25 crucial discoveries: Newton's laws of motion, Chadwick's study of the neutron, Hertz on electromagnetic waves, more. Original accounts clearly annotated. 370pp. 5⅜ x 8½. 0-486-25346-5

EINSTEIN'S LEGACY, Julian Schwinger. A Nobel Laureate relates fascinating story of Einstein and development of relativity theory in well-illustrated, nontechnical volume. Subjects include meaning of time, paradoxes of space travel, gravity and its effect on light, non-Euclidean geometry and curving of space-time, impact of radio astronomy and space-age discoveries, and more. 189 b/w illustrations. xiv+250pp. 8⅜ x 9¼. 0-486-41974-6

THE VARIATIONAL PRINCIPLES OF MECHANICS, Cornelius Lanczos. Philosophic, less formalistic approach to analytical mechanics offers model of clear, scholarly exposition at graduate level with coverage of basics, calculus of variations, principle of virtual work, equations of motion, more. 418pp. 5⅜ x 8½. 0-486-65067-7

Paperbound unless otherwise indicated. Available at your book dealer, online at **www.doverpublications.com**, or by writing to Dept. GI, Dover Publications, Inc., 31 East 2nd Street, Mineola, NY 11501. For current price information or for free catalogues (please indicate field of interest), write to Dover Publications or log on to **www.doverpublications.com** and see every Dover book in print. Dover publishes more than 400 books each year on science, elementary and advanced mathematics, biology, music, art, literary history, social sciences, and other areas.

(continued from front flap)

ELEMENTS OF GASDYNAMICS, H. W. Liepmann and A. Roshko. (0-486-41963-0)

THE KINETIC THEORY OF GASES, Leonard B. Loeb. (0-486-49572-8)

THE THEORY OF ELECTRONS, H. A. Lorentz. (0-486-49558-2)

A GUIDE TO FEYNMAN DIAGRAMS IN THE MANY-BODY PROBLEM, Richard D. Mattuck. (0-486-67047-3)

MATTER AND MOTION, James Clerk Maxwell. (0-486-66895-9)

FUNDAMENTAL FORMULAS OF PHYSICS, VOLS. I & II, Donald H. Menzel (ed.). (0-486-60595-7, 0-486-60596-5)

OPTICKS, Sir Isaac Newton. (0-486-60205-2)

THEORY OF RELATIVITY, W. Pauli. (0-486-64152-X)

INTRODUCTION TO QUANTUM MECHANICS WITH APPLICATIONS TO CHEMISTRY, Linus Pauling and E. Bright Wilson, Jr. (0-486-64871-0)

THE PHILOSOPHY OF SPACE AND TIME, Hans Reichenbach. (0-486-60443-8)

BOUNDARY AND EIGENVALUE PROBLEMS IN MATHEMATICAL PHYSICS, Hans Sagan. (0-486-66132-6)

STATISTICAL THERMODYNAMICS, Erwin Schrodinger. (0-486-66101-6)

PRINCIPLES OF ELECTRODYNAMICS, Melvin Schwartz. (0-486-65493-1)

AN INTRODUCTION TO RELATIVISTIC QUANTUM FIELD THEORY, Silvan S. Schweber. (0-486-44228-4)

ELECTROMAGNETIC FIELD, Albert Shadowitz. (0-486-65660-8)

PARTIAL DIFFERENTIAL EQUATIONS OF MATHEMATICAL PHYSICS, S.L. Sobolev. (0-486-65964-X)

PRINCIPLES OF MECHANICS AND DYNAMICS, Sir William Thomson and Peter Guthrie Tait. (Vol. I: 0-486-49562-0; Vol. II: 0-486-49563-9)

PRINCIPLES OF STATISTICAL MECHANICS, Richard C. Tolman. (0-486-63896-0)

LIGHT SCATTERING BY SMALL PARTICLES, H.C. van de Hulst. (0-486-64228-3)

UNDERSTANDING THERMODYNAMICS, H.C. Van Ness. (0-486-63277-6)

MATHEMATICAL ANALYSIS OF PHYSICAL PROBLEMS, Philip R. Wallace. (0-486-64676-9)

STATISTICAL PHYSICS, Gregory H. Wannier. (0-486-65401-X)

X-RAY DIFFRACTION, B.E. Warren. (0-486-66317-5)

SPACE—TIME—MATTER, Hermann Weyl. (0-486-60267-2)

THEORY OF X-RAY DIFFRACTION IN CRYSTALS, William H. Zachariasen. (0-486-49567-1)

Paperbound unless otherwise indicated. Available at your book dealer, online at **www.doverpublications.com**, or by writing to Dept. 23, Dover Publications, Inc., 31 East 2nd Street, Mineola, NY 11501. For current price information or for free catalogs (please indicate field of interest), write to Dover Publications or log on to **www.doverpublications.com** and see every Dover book in print. Each year Dover publishes over 500 books on fine art, music, crafts and needlework, antiques, languages, literature, children's books, chess, cookery, nature, anthropology, science, mathematics, and other areas.

Manufactured in the U.S.A.

ALBERT MESSIAH

QUANTUM MECHANICS

Two Volumes Bound as One

"...this great work ought to be mastered by every postgraduate research student in theoretical physics . . . there is no other book like it."—*Proceedings of the Physical Society* (England)

Simple enough for students yet sufficiently comprehensive to serve as a reference for working physicists, this classic text initially appeared as a two-volume French edition and is now available in this convenient, all-in-one-book English translation. Formalism and its interpretation receive a detailed treatment in the first volume, starting with the origins of quantum theory and examinations of matter waves and the Schrödinger equation, one-dimensional quantized systems, the uncertainty relations, and the mathematical framework and physical content of formalism. An analysis of simple systems includes a look at the separation of variables, scattering problems and phase shifts, the Coulomb interaction, and the harmonic oscillator.

Volume II begins with an exploration of symmetries and invariance, including a consideration of angular momentum, identical particles and the Pauli exclusion principle, invariance and conservation laws, and time reversal. Methods of approximation discussed include those involving stationary perturbations, the equation of motion, variational method, and collision theory. The final chapters review the elements of relativistic quantum mechanics, and each volume concludes with useful appendixes.

The book has been hailed for the clarity and coherence of its presentation, and its scrupulous attention to detail.

Unabridged Dover (1999) republication in one volume of the English translation originally published in two volumes by John Wiley & Sons, New York, c.1958. Preface. Appendixes. Index. 1152pp. 5⅜ x 8½. Paperbound.

See every Dover book
in print at
www.doverpublications.com

$29.95 USA

PRINTED IN THE U.S.A.

ISBN-13:978-0-486-40924-5
ISBN-10:0-486-40924-4

52995

9 780486 409245