The Historical Development of Quantum Theory

Jagdish Mehra
Helmut Rechenberg

The Historical Development of Quantum Theory

VOLUME 3

The Formulation of Matrix Mechanics and Its Modifications 1925–1926

Springer-Verlag

New York Heidelberg Berlin

Library of Congress Cataloging in Publication Data
Mehra, Jagdish.
The formulation of matrix mechanics and its modifications, 1925–1926.
(The historical development of quantum theory; v. 3)
Bibliography: p.
Includes index.
1. Matrix mechanics. I. Rechenberg, Helmut, joint author
I. Title. III. Series: Mehra, Jagdish. The historical development of quantum theory; v. 3.
QC173.98.M44 vol. 3 [QC174.3] 530.1'2109s 82-3253
[530.1'22] AACR2

All rights reserved. No part of this book may be translated or reproduced in any form without written permission from Springer-Verlag, 175 Fifth Avenue, New York, New York 10010, USA.

© 1982 by Springer-Verlag New York Inc.

Printed in the United States of America.

9 8 7 6 5 4 3 2 1

ISBN 0-387-90675-4 Springer-Verlag Berlin Heidelberg New York
ISBN 3-540-90675-4 Springer-Verlag New York Heidelberg Berlin

Contents

Foreword vii

Introduction 3

Chapter I The Rediscovery of a Mathematical Tool 5

I.1 Max Born's Interpretation of Heisenberg's Quantum Condition 5
I.2 The Development of Matrix Calculus 13
I.3 Early Applications of Matrix Methods in Physics 34
I.4 Born's New Collaborator: Pascual Jordan 44

Chapter II Matching the Tools and the Task 61

II.1 The Programme of Matrix Mechanics 62
II.2 Operations with Matrices 65
II.3 Dynamical Laws and Energy Conservation 72
II.4 An Example of Discrete Mechanics: The Oscillator 80
II.5 Preliminary Remarks on Radiation 87

Chapter III Completion of the Matrix Scheme 91

III.1 The Three-Man Collaboration 92
III.2 Towards a New Perturbation Theory 103
III.3 Several Degrees of Freedom and Degeneracy 109
III.4 Born's *Idée Fixe* and a Letter to Niels Bohr 118
III.5 The Eigenvalue Problem and the Transformation to Principal Axes 121
III.6 Continuous Spectra and the Significance of the Transformation Matrix 129

Chapter IV The Success of Matrix Mechanics 139

IV.1 The Treatment of Dispersion Phenomena 140
IV.2 Fluctuations in Cavity Radiation 149

IV.3 The Conservation of Angular Momentum 157
IV.4 Wolfgang Pauli's Conversion 166
IV.5 The Solution of the Hydrogen Problem 174
IV.6 The Problems of Intensities and the Diatomic Molecule 185

Chapter V Modifications and Extensions of Matrix Mechanics 195
V.1 Nonmechanical Stress versus Spin 196
V.2 Field-Like Representation of Quantum Mechanics 204
V.3 The Operator Mechanics 220
V.4 Multiply Periodic Systems: Action-Angle Variables and the Method of Complex Integration 247
V.5 The Electron Spin, Fine Structure and Anomalous Zeeman Effects 266
V.6 Key to the Helium Problem 282

References 303

Author Index 331

Foreword

This volume deals with the formulation of matrix mechanics by Max Born, Werner Heisenberg and Pascual Jordan in fall 1925, and the immediate applications of this scheme to treat atomic problems (Wolfgang Pauli's triumphant treatment of the hydrogen atom, fall 1925; Heisenberg and Jordan's treatment of the anomalous Zeeman effects by including electron spin, early in 1926; Heisenberg's treatment of the many-body problem and resonance in quantum mechanics, spring 1926). Also treated are the modifications of the matrix scheme by Cornelius (Kornel) Lanczos ('field-like' representation of quantum mechanics) and Max Born and Norbert Wiener's formulation of operator mechanics, both in fall 1925. Matrix mechanics and its modifications, as well as Paul Dirac's work on the fundamental equations of quantum mechanics and the algebra of *q*-numbers preceded Erwin Schrödinger's formulation of wave mechanics.

I had the great privilege of discussing these developments with all the participants involved. Helmut Rechenberg and I have made use of these discussions to weave the human context in the narrative of the physical and mathematical framework of the discoveries in question. We have also drawn upon the rich source materials of the *Archives for the History of Quantum Physics*, especially the interviews with Max Born, Werner Heisenberg and Pascual Jordan (cited here as 'AHQP Interviews').* In addition, we have made use of the scientific correspondence of the principal physicists in question.

I am grateful to Aage Bohr for allowing me complete access to the archives held at the Niels Bohr Institute in Copenhagen. I am indebted to the architects of quantum mechanics and the heirs of their literary and scientific estates for permission to use the source materials pertaining to them.

JAGDISH MEHRA

*An inventory of these sources is contained in: T. S. Kuhn, J. L. Heilbron, P. Forman and L. Allen, *Sources for History of Quantum Physics*, The American Philosophical Society, Philadelphia, 1967.

The Formulation of Matrix Mechanics and Its Modifications 1925–1926

Introduction

On 30 August 1925 Niels Bohr delivered an address on 'Atomic Theory and Mechanics' to the sixth Scandinavian Mathematical Congress in Copenhagen. In his remarks on 'an attempt at a rational quantum mechanics,' which he appended to the original address in November 1925 before publication, Bohr said:

> It will interest mathematical circles that the mathematical instruments created by the higher algebra play an essential part in the rational formulation of the new quantum mechanics. Thus the general proofs of the conservation theorems in Heisenberg's theory carried out by Born and Jordan are based on the use of the theory of matrices, which go back to Cayley and were developed by Hermite. It is to be hoped that a new era of mutual stimulation of mechanics and mathematics has commenced. To the physicist it will seem first deplorable that in atomic problems we have apparently met with such a limitation of our usual means of visualisation. This regret will, however, have to give way to thankfulness that mathematics, in this field too, presents us with the tools to prepare the way for futher progress. (Bohr, 1925b, p. 852)[1]

With these remarks Bohr addressed himself to the latest status of atomic theory: the difficulties arising in the usual mechanical interpretation, as well as the progress achieved by Heisenberg's discovery of the quantum-mechanical scheme and its subsequent mathematical formulation by Born and Jordan. In the latter development Bohr saw a revival of the long-established connection between mathematics and physical theory. Since the days of Isaac Newton, mechanics had always been regarded as a part of mathematics. However, in the twentieth century, with the perfection attained by classical mechanics in Henri Poincaré's *Les Méthodes Nouvelles de la Mécanique Céleste* (1892, 1893, 1899), on the one hand, and Einstein's special and general relativity theories (in which, again, the mathematicians like Poincaré, Hermann Minkowski, Marcel Grossmann, David Hilbert, Felix Klein, Curbastro Gregorio Ricci and Tullio Levi-Civita played an important role), on the other, one began to believe that mathematics and physics—of which the Newtonian and relativity mechanics were a completed segment—had no further overlap. Bohr's reminder to the

[1] On 25 November 1925 Bohr wrote to Pauli that, induced by a previous letter of Pauli (Pauli to Bohr, 17 November 1925), he had made a series of changes in his Copenhagen address of August 1925, taking into account the recent progress in quantum mechanics initiated by Heisenberg.

mathematicians, that the unsolved problem of atomic theory would still require the use of appropriate mathematical tools, was therefore timely. As proof, he cited the fact that 'a self-contained theory sufficiently analogous to classical mechanics' had been formulated by Max Born and Pascual Jordan on the basis of 'Heisenberg's quantum mechanics,' and that this new atomic mechanics employed methods which had been developed in higher algebra (Bohr, 1925b, p. 852).

The quick mathematical formulation of the physical ideas, which Heisenberg had proposed in his paper on the quantum-theoretical reformulation of kinematic and mechanical relations, in a consistent theory came as a great surprise to most physicists who had laboured for years on the problems of atomic theory. A wider dissemination of matrix mechanics was limited, however, by the fact that its mathematical methods could not be handled easily for the purpose of calculating the properties of atomic systems. An important breakthrough was achieved by Wolfgang Pauli, who calculated the energy terms of the hydrogen atom in late October 1925. But it was necessary to extend Born and Jordan's original matrix scheme to deal with atoms containing more than one electron. Several authors independently provided such extensions: Cornelius Lanczos of Frankfurt University and Max Born and Norbert Wiener, then at the Massachusetts Institute of Technology. While Lanczos' proposal concerning a field-theoretical formulation of quantum mechanics became fruitful only later on in the wave mechanical theory, which Erwin Schrödinger discovered independently, the operator mechanics of Born and Wiener helped Heisenberg, Pauli and Gregor Wentzel to generalize matrix mechanics into a symbolic calculus, which was also applicable to the action and angle variables. The observed properties of complex atoms could still not be accounted for until two further features were introduced into quantum mechanics: first, the electron spin, discovered by George Uhlenbeck and Samuel Goudsmit in the analysis of multiplet spectra (Uhlenbeck and Goudsmit, 1925); second, the conception of exchange degeneracy (Heisenberg, 1926b). By incorporating these extensions of the matrix scheme, Heisenberg and Jordan successfully calculated the anomalous Zeeman splittings of complex spectra (Heisenberg and Jordan, 1926), and Heisenberg arrived at a qualitative understanding of the energy states of the helium atom (Heisenberg, 1926c). Thus, within a period of less than ten months after the submission of Heisenberg's first paper (1925c), quantum mechanics developed into a theory which provided a satisfactory explanation of many important problems of atomic physics that had concerned physicists for a long time.

Chapter I
The Rediscovery of a Mathematical Tool

Heisenberg's observation that the products of the quantum-theoretical analogues of the classical Fourier series, which were supposed to describe the dynamical variables of atomic systems, did not necessarily commute, led to an obvious conclusion: the calculus used in classical mechanics would not serve as an appropriate tool in atomic theory, and different mathematical methods had to be applied. Max Born discovered that infinite matrices could be used to formulate quantum-mechanical relations consistently. The operations with such matrices had been developed by the mathematicians during the second half of the nineteenth century and brought to some perfection in the theory of linear integral equations by Born's teacher David Hilbert in Göttingen. Born had not only witnessed Hilbert's work on that subject, but had himself applied the methods of infinite matrices to physical problems, as for example to the problem of specific heat of solids; hence he was well acquainted with the details of the matrix calculus. In Pascual Jordan he found an able collaborator for the task of formulating matrix mechanics.

I.1 Max Born's Interpretation of Heisenberg's Quantum Condition

During the summer semester of 1925, Max Born, Professor of Theoretical Physics at the University of Göttingen, had been especially busy. In a long letter to his friend Albert Einstein in Berlin, dated 15 July 1925, he gave an account of the events that had taken place. A number of physicists had visited Göttingen: Paul Ehrenfest from Leyden, who lectured on Bose–Einstein statistics and gas degeneracy, Hendrik Kramers from Copenhagen, Peter Kapitza from Cambridge, and Philipp Frank from Prague. 'For us,' Born wrote to Einstein, 'this is very stimulating, but it is often too much for our wives. So they simply run away; my wife and Mrs. Courant have already left, and Mrs. [Philipp] Frank is due to leave in two days' time.'[2] Then he continued: 'But do not conclude from that that your visit is going to be unwelcome! We are greatly looking forward to it! But it ought to be at a quieter time. In July, most of the foreigners are already on holiday,

[2] Hedwig Born had left Göttingen on 13 July 1925, accompanied by the children, to go to Silvaplana in the Engadine, Switzerland. Max Born joined her soon after the end of the semester, i.e., early in August 1925.

and they descend on us in droves. But you know all this business. There is going to be another rumpus tomorrow; it is the inauguration of [Ludwig] Prandtl's new hydrodynamic institute, with a guided tour, official dinner and gala concert. It will cost me almost an entire working day' (Born to Einstein, 15 July 1925).

Besides being occupied with receiving visitors and other social activities, Born was deeply interested in the problems of quantum theory. For example, he reported to Einstein about the exciting discussions on Louis de Broglie's matter waves that had been taking place in Göttingen at that time, which finally led Walter Elsasser to his explanation of the Ramsauer effect (Elsasser, 1925). He continued: 'But my principal interest is the rather mysterious difference calculus on which the quantum theory of atomic structure is based. Jordan and I are systematically (though with a minimum of mental effort) examining every imaginable correspondence relationship between classical, multiple-periodic systems and quantum atoms' (Born to Einstein, 15 July 1925). Indeed, Born hoped that the reformulation of the equations of atomic theory as difference equations would lead to a breakthrough in the description of atomic phenomena, for his previous investigations had pointed in that direction.

Upon concluding that all the advanced methods of celestial mechanics had failed in the helium problem (Born and Heisenberg, 1923b), Born had decided that new mathematical tools, completely different from those provided by the differential calculus and applied so successfully to the problem of physics from Newton to Poincaré, had to be developed for atomic mechanics. The calculus of differences seemed to present a natural method to account for the most important result connecting the frequencies and energies of atomic systems; in this case, the classical relation

$$\nu_\tau^{\mathrm{cl}} = \tau \frac{\partial W}{\partial J} , \tag{1}$$

giving ν_τ^{cl}, the τth harmonic frequency, as a derivative of W, the energy, with respect to the action integral J, could be replaced by the difference equation

$$\nu^{\mathrm{qu}}(n, n-\tau) = \frac{1}{h}\left[W(n) - W(n-\tau) \right]. \tag{2}$$

Born had then concluded that it was necessary to reformulate all the differential formulae of classical multiply periodic systems as difference equations, and had started to search for appropriate examples (Born, 1924b). Although the mathematical method, the calculus of differences, was available to a certain extent (see, e.g., Nörlund, 1924), Born made only slow progress in his programme of 'discretization' of classical mechanical relations. The main difficulty was that it was not altogether clear as to which of these relations should be reformulated at all as difference equations. However, encouraged by the successful procedure employed in dispersion theory (Kramers, 1924a, b; Born, 1924b), Born had arrived eventually at an important prescription concerning the reformulation of

mechanical equations in quantum theory. In a paper, submitted together with Pascual Jordan in June 1925, he had proposed to deal only with those classical equations in which all quantities correspond to observable properties of atomic systems (see Born and Jordan, 1925a, especially p. 493). Born referred to that paper in his letter to Einstein with the following words: 'A paper on this subject [i.e., on the application of the difference calculus in atomic theory] in which we examine the effect of non-periodic [electromagnetic] fields on atoms will appear soon. This is a preliminary study for an investigation of the processes occurring in atomic collisions (quenching of fluorescence, sensitized fluorescence *à la* Franck, etc.); one can understand, I think, the essential characteristics of what goes on. The different behaviour of atoms depends mainly on whether they have an (average) dipole moment, or just a quadrupole moment, or even a still higher electric symmetry' (Born to Einstein, 15 July 1925). Born hoped that the long march towards the unknown quantum laws would include as steps his discretization procedure and the use of observable quantities.

Born had barely sent off his letter to Einstein when his collaborator Werner Heisenberg came to see him. He had in the meanwhile 'pursued some work of his own, keeping its idea and purpose somewhat dark and mysterious,' as Born recalled later (Born, 1978, p. 216). During the summer semester of 1925 Born had seen less of Heisenberg than in previous times. The two of them had not worked on common projects; actually Born had been absorbed in his official duties and in his collaboration with Jordan on the application of dispersion-theoretic methods to aperiodic atomic systems, while Heisenberg had devoted his efforts to the general questions arising from the problem of calculating the intensities of hydrogen lines, a problem which was not in the centre of Born's interests. In May and early June 1925, before he went to Helgoland to recover from hay fever, Heisenberg had not discussed his work with Born. As he recalled later: 'At that time everything was so vague. We may, of course, have talked in a general manner about what one could possibly do, but the only time when I really had discussions about it was when he gave me this book on the Bessel functions' (Heisenberg, Conversations, p. 269). These discussions had taken place already earlier in May, when Heisenberg had begun by guessing the intensities of the Balmer lines and Born had expected that Bessel functions would play a role in solving that problem. After returning from Helgoland, Heisenberg had communicated exclusively with Wolfgang Pauli by letters; Pauli had thus been the only person who was familiar with the difficulties and successes of the quantum-mechanical scheme prior to the completion of the paper (Heisenberg, 1925c). Heisenberg had even sent the manuscript first to Hamburg, requesting Pauli's opinion before showing it to anybody in Göttingen. Upon receiving a favourable response from Pauli, he had then completed the paper and gone to see Born.

Born recalled the circumstances of receiving Heisenberg's paper as follows: 'He came to me with a manuscript and asked me to read it and to decide whether it was worth publishing. At the same time he asked me for leave of absence for the rest of the term (which ended about [1] August [1925]), as he had an

invitation to lecture at the Cavendish Laboratory in Cambridge. He added that though he had tried hard, he could not make any progress beyond the simple considerations contained in his paper, and he asked me to try myself, which I promised to do' (Born, 1978, p. 216). In spite of the fact that he 'felt that something real was in the paper,' Heisenberg was 'very uncertain about it' (Heisenberg, Conversations, p. 269). Hence he left the question of whether or not the paper should be submitted to *Zeitschrift für Physik* to Max Born, and went away from Göttingen without waiting for the final decision.[3]

Born received Heisenberg's paper on the quantum-theoretical re-interpretation of kinematic and mechanical relations around the middle of July, the time when he wrote the letter to Einstein.[4] Because of his preoccupation with other events—such as the inauguration of Prandtl's hydrodynamic institute—he delayed studying the new work. Many years later he said: 'I remember that I did not read this manuscript at once because I was tired after the term and afraid of hard thinking. But when, after a few days, I read it I was fascinated' (Born, 1978, p. 216).[5] He began to think especially about the meaning of the strange rule, which Heisenberg had given for the multiplication of two quantum-mechanical transi-

[3] According to the recollection of Friedrich Hund (private communication), Heisenberg gave an informal seminar on his new quantum-mechanical scheme to some friends in Göttingen after the completion of his paper (Heisenberg, 1925c). Besides Hund, these people included (most probably) Pascual Jordan, Hertha Sponer and Wilhelm Hanle.

[4] Heisenberg sent his manuscript to Pauli with a letter, dated 9 July 1925; this was on a Thursday. He could scarcely have received Pauli's reply before Monday, 13 July; Born, therefore, got the paper a couple of days later.

[5] In Born's letter to Einstein, dated 15 July 1925, there occurred the sentence: 'Heisenberg's latest paper, soon to be published, appears rather mystifying but is certainly true and profound.' In his commentary on *The Born–Einstein Letters*, Born interpreted this sentence as follows:

> Then comes the most important matter: a few lines about Heisenberg's new paper, which must have appeared "mystifying" but nevertheless true. This must have been the treatise in which he formulates the basic concepts of quantum mechanics and explains them by using simple examples. As my recollection of this time, which marked the beginning of a revolution in physical thinking, is a little hazy, I wrote to Professor van der Waerden, who confirmed my assumption. His book [van der Waerden, 1967] will enable the reader to look up the sequence of events in complete detail. (Einstein and Born, 1971, p. 87)

Van der Waerden, in his introduction to the collection of papers, *Sources of Quantum Mechanics*, reconstructed the events as follows: 'On 11 or 12 July Heisenberg gave Born the final version of his paper, asking him to decide whether it was worth publishing . . . . Born must have studied Heisenberg's manuscript before July 15, for on July 15 he writes in a letter to Einstein: ' . . . *Heisenberg's neue Arbeit die bald erscheint, sieht sehr mystisch aus, ist aber sicher richtig und tief* . . . ' (van der Waerden, 1967, p. 36).

It is easy to see that this conclusion is wrong because the sentence on Heisenberg's new paper in Born's letter to Einstein (15 July 1925) continues as follows: 'It enabled Hund to bring into order the whole of the periodic system [of elements] with all its complicated multiplets.' Born also indicated that Hund's paper on the interpretation of complex spectra (Hund, 1925c), which was built on Heisenberg's paper, would appear soon; this was indeed the case as Hund's paper was published in August. In that paper, however, Hund referred only to Heisenberg's paper on complex multiplets, which he had submitted from Copenhagen in April 1925 and which appeared in July 1925 (Heisenberg, 1925b). Hence it is clear that Born in his letter to Einstein did not refer to Heisenberg's paper on the new quantum-mechanical scheme; moreover, he had not yet even studied it, for otherwise he would have mentioned to Einstein that it was even more 'mystical' than the paper on multiplets.

tion amplitudes, namely, the equation

$$C(n, n-\beta) = \sum_{\alpha=-\infty}^{+\infty} A(n, n-\alpha) B(n-\alpha, n-\beta), \tag{3}$$

in which the product amplitude $C(n, n-\beta)$ was expressed as a sum of products of amplitudes $A(n, n-\alpha)$ and $B(n-\alpha, n-\beta)$. This multiplication rule, which was different from the usual multiplication of two Fourier series describing the variables of classical periodic systems, seemed to provide the key to Heisenberg's quantum-mechanical scheme; hence it was necessary to understand its significance. 'I began to ponder over his symbolic multiplication,' Born recalled, 'and was soon so involved in it that I thought about it the whole day and could hardly sleep at night. For I felt there was something fundamental behind it, the consummation of our endeavours of many years. And one morning, about the 10 July 1925, I suddenly saw light: Heisenberg's symbolic multiplication was nothing but the matrix calculus, well known to me since my student days from the lectures of [Jakob] Rosanes in Breslau' (Born, 1978, p. 217).[6]

According to Pascual Jordan's recollection Born had himself thought about the mathematical process of multiplication in quantum theory even before seeing Heisenberg's paper. Thus he reminded Born: 'By the way, I recall with certainty that, following our joint work on absorption [Born and Jordan, 1925a], you considered a symbolic multiplication of quantum-theoretical "transition amplitudes" yourself; at that time we used to discuss together almost every day, and you told me about it. Only, at that time, we did not see clearly what purpose this multiplication should serve; that one could use it to establish analogous equations of motion, we had not seen immediately' (Jordan to Born, 3 July 1948). However, exactly this recognition had provided the starting point for Heisenberg's general considerations in May 1925, which had also led to the noncommutative multiplication rule, Eq. (3). Although Heisenberg had thus proceeded faster and arrived at a definite conclusion, Born was familiar with the problem and prepared to grasp clearly the importance of the result.

It seems to be remarkable that Born, at the moment of recognizing the meaning of Heisenberg's multiplication rule, thought so far back as the lectures on linear algebra, which he had attended in the beginning of the century at the University of Breslau. However, the fact that two matrices do not commute like ordinary numbers—which is mentioned in all elementary lectures on matrix

[6] As we have discussed above, the idea of identifying the multiplication rule, Eq. (3), with the known matrix multiplication, must have occurred to Born after 15 July.

Born studied in Breslau from 1901 to 1904, except the summer semesters 1902 (Heidelberg) and 1903 (Zurich). He attended Rosanes' lectures on algebra probably in his first semester. Rosanes lectured on '*Elemente der Determinantentheorie*' ('Elements of the Theory of Determinants,' summer semester 1901), '*Analytische Geometrie des Raumes*' ('Analytical Geometry of [Three-Dimensional] Space,' winter semester 1901–1902), and *Einführung in die Theorie der Invarianten* ('Introduction to the Theory of Invariants,' winter semester 1902–1903). (See *Vorlesungsverzeichnis*, *Phys. Zs.* **2**, 1900–1901, p. 395, p. 728; **3**, 1901–1902, p. 590.)

calculus—leaves a definite impression when one learns it for the first time.[7] In order to connect Eq. (3) with the familiar notation used in matrix algebra, Born had merely to rewrite it as

$$C_{nm} = \sum_{n'=-\infty}^{+\infty} A_{nn'} B_{n'm}\,, \tag{4}$$

where the two numbers on which the transition amplitudes depend are written as subscripts with $n' = n - \alpha$ and $m = n - \beta$. Equation (4) thus denoted the product of two quadratic matrices, A and B, whose discrete indices take on infinitely many values.

Born discovered the first example of the occurrence of two noncommuting matrices in Heisenberg's quantum condition (Heisenberg, 1925c, p. 886, Eq. (16)),

$$h = 4\pi m \sum_{\alpha=0}^{\infty} \left\{ |a(n, n+\alpha)|^2 \omega(n+\alpha, n) - |a(n, n-\alpha)|^2 \omega(n, n-\alpha) \right\}. \tag{5}$$

By expressing Heisenberg's q, the position coordinate of the periodic system under investigation, as

$$q_{nn'} = a(n, n') \exp\left[i\omega(n, n') t \right], \tag{6}$$

and the conjugate momentum $p_{nn'}$ by

$$p_{nn'} = m \frac{dq_{nn'}}{dt}\,, \tag{7}$$

where m is the moving mass of the system, Born reformulated the quantum condition (5) as the equation[8]

$$\frac{h}{2\pi i} = \sum_{n'=-\infty}^{+\infty} (p_{nn'} q_{n'n} - q_{nn'} p_{n'n}). \tag{8}$$

'I recognized at once its formal significance,' he recalled many years later. 'It meant that the two matrix products pq and qp are not identical. I was familiar

[7] Later, Born had to deal with matrices often, as for example in his continuation of Hermann Minkowski's work on electrodynamics (Born, 1909b). The noncommutativity property of matrices is such an elementary property that it is often not even mentioned in the more sophisticated applications of matrix calculus (see, e.g., Hermann Weyl's discussion of tensor algebra in *Raum-Zeit-Materie*, Chapter I, §§5–8, Weyl, 1918c). However, it was referred to in Richard Courant's book *Methoden der mathematischen Physik* (Courant–Hilbert, 1924, Chapter One, Section 3).

[8] Evidently, one can express the right-hand side of Eq. (5) as 2π times the factor $m\sum_{\alpha=-\infty}^{+\infty}\{|a(n, n+\alpha)|^2\omega(n+\alpha, n) - |a(n, n-\alpha)|^2\omega(n, n-\alpha)\}$, or with the help of Eqs. (6) and (7), as 2π times the factor $(1/i)\sum_{n'} (q_{nn'}p_{n'n} - p_{nn'}q_{n'n})$.

with the fact that matrix multiplication is not commutative; therefore I was not too much puzzled by this result' (Born, 1978, p. 217).

It was at about this stage that Born, full of excitement with his discovery about what Heisenberg's quantum rule of multiplication signified mathematically, travelled to Hanover on Sunday, 19 July 1925, to attend the meeting of the *Gauverein Niedersachsen* (Lower Saxony Section) of the German Physical Society.[9] He recalled afterwards: 'A considerable number of physicists from Göttingen went there by train; it is about an hour's journey on the North Express. In the train we met physicists from other universities, among them my former assistant Pauli' (Born, 1978, p. 218). Born happily 'joined him [Pauli] in his compartment' and 'at once told him about matrices.'[10] Pauli, who had been Heisenberg's confidant while his new ideas had taken shape, who knew all about his work and had approved it only a week ago, did not respond favourably to Born's excitement about matrices. When Born asked Pauli whether he would care to collaborate with him on the further development of Heisenberg's ideas, he gave a 'cold and sarcastic refusal.' 'Yes, I know,' Pauli retorted, 'you are fond of tedious and complicated formalisms. You are only going to spoil Heisenberg's physical ideas by your futile mathematics' (Born, 1978, p. 218).

Born's approach to Pauli and the latter's response were typical of the attitudes of both men, who had been the first to learn about the new quantum-mechanical scheme. Born was happy when he discovered the fact that Heisenberg's multiplication rule, Eq. (3), was none other than matrix multiplication, an operation well known to mathematicians in the work of Arthur Cayley since the 1850s, for he could now exploit the entire formalism of matrix methods more effectively than he had been able to do, e.g., in his work on electron theory sixteen years previously (Born, 1909b). It occurred to Born that the well-established methods of matrix algebra would probably offer a far better tool in discovering the real quantum laws of atomic mechanics than the calculus of differences, with which he had been stuck for almost two years. Quickly abandoned now was Born's own correspondence principle, i.e., the formulation of quantum laws by transforming

[9]The record of this meeting (*Verh. d. Deutsch. Phys. Ges. (3)* **6**, 1925, pp. 36–39) states that Born discussed there his recent work on aperiodic motions (Born and Jordan, 1925a) and Hund his interpretation of complex spectra (Hund, 1925c).

[10]At this particular point Born's memory was not accurate. As he recalled: 'Pauli had meanwhile become famous from many excellent papers, among them that on his celebrated exclusion principle, on which Niels Bohr had built his theory of the periodic system of elements. He was coming from Zurich (where the summer vacation started earlier than in Germany) to take part in our meeting in Hanover' (Born, *Recollections*, 1978, p. 218).

Pauli actually went to Hanover from Hamburg; it was almost three years before he moved to his professorship in Zurich (spring 1928). He did not give a talk at Hanover himself; however, he was writing his article on quantum theory for the *Handbuch der Physik* and was interested in discussing the latest status of spectroscopic theory with such experts as Ralph Kronig, Wilhelm Hanle and Friedrich Hund, who reported about their work at Hanover. As far as the meeting of Pauli and Born is concerned, Pauli certainly did not join Born in Göttingen to go to Hanover; however, the two might have had occasion to talk privately at the railway station after the meeting, before leaving Hanover in opposite directions to Hamburg and Göttingen, respectively.

differential equations into difference equations, and there appeared to him the vision of a new territory to be explored by means of an obviously erudite, fully developed, and general mathematical apparatus. Moreover, Born believed that he could make sense of Heisenberg's new ideas, and thereby respond to the latter's appeal which had been expressed to Pauli: 'But perhaps people, who know more, may be able to do something reasonable with them' (Heisenberg to Pauli, 9 July 1925).

Wolfgang Pauli, however, did not share Born's enthusiasm. He had in fact been glad that 'Heisenberg has learned a little philosophy from Bohr in Copenhagen, and does indeed turn away noticeably from the purely formalistic approach' (Pauli to Kramers, 27 July 1925). And now Born talked about a well-defined mathematical scheme, whereas he, Pauli, only saw in Heisenberg's work the first tender green shoots in the barren terrain of the difficulties of atomic physics. No, Pauli was not the right collaborator to work with Born in applying matrix methods to Heisenberg's scheme. Moreover, he was convinced that the product of the physical quantities, p and q, had to commute, and Born, in his opinion, just could not be on the right track.[11] In any case, Pauli did not wish to interfere with Heisenberg's thoughts and the manner in which he would develop his theory. As Jordan reported later: 'To Miss Mensing . . . he had mentioned at that time that he wanted to leave the topic to Heisenberg himself for the moment' (Jordan to Born, 3 July 1948).

Since Pauli had given him a negative answer concerning a collaboration on quantum mechanics, Born turned to Pascual Jordan, his main helper in those days, because he was 'extremely tired and felt unable to make progress alone' (Born, 1978, p. 218).[12] Jordan was glad to assist his professor again and, within a short time, achieved important progress in the matrix formulation of Heisenberg's ideas. Soon Born was able to spread the news about the new theory outside Göttingen. On the way to spend his vacation in Switzerland, he delivered a lecture in Tübingen on 30 July 1925. Alfred Landé still remembered Born's visit four decades later: 'I first heard of matrix mechanics when Born was in Tübingen . . . . He told me something about it; that they [i.e., the Göttingen theoreticians] had a completely new approach to quantum mechanics, and everything [was] dominated by multiplication, [in which] A times B differs from BA. I did not understand a single word about it, and I don't think that Born and the whole group in Göttingen "understood" much more than the mere formulas' (Landé, AHQP Interview, 3 July 1962, p. 21). However, the mathematical methods of matrices enabled Born and Jordan to formulate the first consistent theory of quantum mechanics within less than two months.

[11] Pascual Jordan, who had not witnessed Born and Pauli's discussion in Hanover, but heard about it immediately afterwards from Born, recalled: 'Pauli had then told you that one had to assume $pq = qp$' (Jordan to Born, 3 July 1948).

[12] One may wonder why Born did not think of Jordan right away. He probably believed that the task of formulating Heisenberg's ideas was very difficult and that Jordan might not be experienced enough.

I.2 The Development of Matrix Calculus

The concept of matrices and the methods of operating with these objects arose in mathematics from three sources. Perhaps the earliest was the renewed interest in conic sections and the subsequent classification of the latter in the language of algebra, which was basically initiated by the work of René Descartes in the first half of the seventeenth century.[13] More than a century later Leonhard Euler systematically studied second degree equations in three variables x, y, z, i.e., equations of the form

$$ax^2 + by^2 + cz^2 + dxy + exz + fyz + gx + hy + kz = 1, \tag{9}$$

with constant coefficients, a, b, c, d, e, f, g, h, and k, which describe surfaces in the three-dimensional space (Euler, 1748).[14] He introduced the transformation from the coordinate system, given by the coordinates, x, y, z, to another system, (x', y', z'), through the rotation of the coordinate axes by the three (Euler) angles in such a way that the new coordinate axes coincide with the principal axes of the geometrical object under consideration (cone, cylinder, ellipsoid, hyperboloid, hyperbolic paraboloid, or parabolic cylinder). With this 'principal-axes' transformation, he then succeeded in classifying the geometrical objects in space described by the second-order equation (9).

The second development that led to matrices emerged from a purely algebraic consideration, i.e., the investigation of the simultaneous solution of systems of linear equations, which were later written as

$$x_i = \sum a_{ik} y_k \qquad (i, k = 1, \ldots, n), \tag{10}$$

[13]Conic sections, such as ellipses and hyperbolae, were already studied in the ancient times, especially by the Greeks, but it was only after the Middle Ages that scientists again became interested in them. For example, Kepler's recognition that the orbits of planets were ellipses demanded a rational explanation of the latter.

René Descartes was born on 31 March 1596 in La Haye, Touraine. He graduated as a lawyer twenty years later from the University of Poitiers and studied mathematics for a year in Paris in the circle of Father Marin Mersenne (the so-called '*Académie Mersenne*,' a precursor of the *Académie des Sciences*). From 1617 he served in various armies for nine years; then he returned to Paris and learned how to construct optical instruments. In 1629 he moved to Holland, where he found a quiet atmosphere to live and work, until he was finally called to Stockholm (in 1649) to instruct Queen Christina of Sweden. Descartes died there soon afterwards of pneumonia on 11 February 1650.

The foundation of analytical geometry was contained in the appendix on geometry ('*La Géométrie*') in his famous book '*Discours de la Méthode*' (1637). There he went beyond the results of the Greeks when he tried to construct, e.g., the curve $y(x)$ describing an ellipse. Although Descartes' approach was basically constructive, rather than analytic, he explicitly introduced algebraic equations to represent geometrical figures.

[14]Leonhard Euler, born on 15 April 1707 in Basle, studied under Johann Bernoulli at the University of Basle. He worked from 1727 to 1741 and from 1766 to his death on 18 September 1783 in St. Petersburg; in between, from 1741 to 1766, he worked at the Prussian Academy of Sciences in Berlin. He was member of most of the prestigious scientific academies of the day. Euler published a vast number of papers on mathematical and physical problems (about 800 memoirs) and wrote more than ten books.

where x_i are the known and y_k the unknown variables. Gottfried Wilhelm Leibniz concerned himself at the end of the seventeenth century with special cases of systems of three equations (Leibniz, 1693).[15] Being interested in the conditions, under which solutions of Eqs. (10) existed, Leibniz discovered that such conditions could indeed be established: they required the vanishing of certain expressions involving the coefficients a_{ik}. These expressions were used afterwards throughout the eighteenth century, and later named 'determinants' by Augustin Louis Cauchy (Cauchy, 1815).[16] Cauchy not only extended Leibniz' methods to systems of any finite number (n) of linear equations—Leibniz had originally studied systems of three equations with three unknowns—but also took an important formal step when he arranged the coefficients a_{ik} $(i,k = 1, \ldots, n)$ in space arrays,

$$a_{ik} = \begin{matrix} a_{11} & & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & & a_{nn} \end{matrix}, \tag{11}$$

thus inventing the modern notation of using two subscripts. He further proved the general multiplication law for two determinants, stating that the determinant of a product of two arrays, (a_{ik}) and (b_{ik}), whose coefficients c_{ik} are defined as

$$c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}\,, \tag{12}$$

is given by the product of the determinants of the original arrays. However, it

[15] Gottfried Wilhelm Leibniz was born on 1 July 1646 in Leipzig. He studied law, mathematics and philosophy at the Universities of Leipzig (1661–1663), Jena (1663–1666) and Altdorf, where he received his doctorate in law in 1666. He served the Archbishop and Elector of Mainz, Johann Philipp von Schönborn, and the Dukes of Braunschweig–Lüneburg. He died in Hanover on 14 November 1716.

Leibniz, who is sometimes referred to as the last universal scientist, worked in diplomacy, law, history, philosophy, mathematics and physics. Together with Newton, he invented the calculus and, besides trying to develop a universal symbolism for logic, he introduced much of the symbols and notation that were later used in mathematics. He wrote many books on philosophy, theology, mathematics, physics and history, as well as hundreds of articles and thousands of letters on scientific topics. Leibniz, a member of the important academies of sciences extant in his time, was instrumental in founding the Prussian Academy of Sciences in Berlin (1700); he also urged Peter the Great to found the St. Petersburg Academy of Sciences.

[16] The word 'determinant' had been used earlier by Carl Friedrich Gauss to denote the discriminant $\Delta\ (= ac - b^2)$ of a quadratic form $ax^2 + 2bx + cy^2$.

Augustin Louis Cauchy was born on 21 August 1789 in Paris. He received his education at the *École Polytechnique* and served as instructor at the *Collège de France* and at *École Polytechnique* until 1830. Then, because of his adherence to the French monarchy, he left Paris and became professor of mathematical physics at the University of Turin. In 1837 he returned to Paris and to the *École Polytechnique*. He died in Sceaux, France, on 23 May 1857.

Cauchy wrote many books on mathematical and physical topics and over 500 articles. He is especially known for his development of rigorous methods in analysis and for his contributions to the theory of complex functions. In physics he dealt primarily with questions of optics.

remained to James Joseph Sylvester to develop fully the theory of determinants; he devoted fifty years of his scientific career to this subject, which was connected most closely with the origin of matrix theory.[17]

Finally, the third root of matrix methods has to be traced to the efforts to generalize the concept of numbers. In an unpublished paper of 1831 Carl Friedrich Gauss was among the first to consider an extension of complex numbers, the most famous of which are the quaternions of William Rowan Hamilton (1844).[18] A quaternion, q, may be defined as an object of the form

$$q = a + b\hat{i} + c\hat{j} + d\hat{k}, \tag{13}$$

where a is the so-called 'scalar' and the rest is the 'vector' part, with b, c and d representing the Cartesian coordinates of a point in the three-dimensional space, and $\hat{i}$, $\hat{j}$ and $\hat{k}$ are unit vectors. The algebraic operations with quaternions, such as the multiplication, exhibit properties which deviate somewhat from the analogous operations with real or complex numbers. Especially, the product of two quaternions, q_1 and q_2—obtained by observing the rules for the products of imaginary vectors, $\hat{i}^2 = \hat{j}^2 = \hat{k}^2 = -1$, and $\hat{i}\hat{j} = -\hat{j}\hat{i} = \hat{k}$, $\hat{i}\hat{k} = -\hat{k}\hat{i} = -\hat{j}$ and $\hat{j}\hat{k} = -\hat{k}\hat{j} = \hat{i}$—is associative but not commutative.

[17] James Joseph Sylvester was born on 3 September 1814 in London. After attending schools for Jewish boys in London and the University of London, he went in 1829 to the Royal Institution in Liverpool. Being an excellent student and having won several prizes for solving mathematical problems, he continued his education at St. John's College, Cambridge, from October 1831 to the end of 1833. Illness prevented him from taking the Tripos examination before January 1837, and he stood second on the list. Being Jewish, however, he was not allowed to take a degree or to compete for the Smith's Prize. (He obtained his degrees, B.A. and M.A., only in 1841 from Trinity College, Dublin.) Sylvester then became professor of natural philosophy in the University of London in 1837 and professor of mathematics at the University of Virginia, U.S.A., in 1841. He returned to England after two years, left his academic career for some time, and became engaged in actuarial work. He was called to the Bar in 1850, where he met and made friends with Arthur Cayley. From 1855 to 1870 Sylvester served as professor of mathematics at the Royal Academy in Woolwich, England, and in 1876 he began to lecture at the new Johns Hopkins University in Baltimore until he was called to the Savilian Chair of Geometry at Oxford University in 1883. He retired in 1894 and died on 15 March 1897 in Oxford.

Sylvester was elected Fellow of the Royal Society already in 1839 because of his work on the theory of optical crystals and the motion of rigid bodies. He contributed many papers to the theory of numbers, theory of equations, pure and analytic geometry, and algebra. He founded the theory of algebraic invariants. Besides his mathematical research, his educational and administrative activity was enormous. In the Woolwich time he edited the *Quarterly Journal of Pure and Applied Mathematics*; in Baltimore he founded the *American Journal of Mathematics*, thus initiating research in pure mathematics in the United States. He was a member of many associations and academies, including the London Mathematical Society (President, 1866–1869) and the *Académie des Sciences*, Paris (elected in 1863). He received numerous honours and prizes, including the Copley medal of the Royal Society in 1880. Sylvester also wrote poetry; he translated Horace and several German poets into English. His style is also reflected by the rich mathematical vocabulary which he created.

[18] William Rowan Hamilton (1805–1865) invented the quaternions in late 1843 and announced them at the meeting of the Royal Irish Academy. He rated this discovery very highly and wrote two books on this subject. Quaternions became a fashionable issue of debate later in the nineteenth century, without leading to important applications in mathematics and physics as Hamilton had hoped. However, quaternions stimulated the invention of other hypercomplex numbers, and also contributed to the development of vector and tensor methods.

Around the middle of the nineteenth century the above-mentioned concepts and methods had begun to merge towards a more unified calculus. For example, connections had been found between the problem of solving systems of linear algebraic equations and the problem of obtaining the standard expression for quadratic forms in three and more dimensions. For the latter purpose, Euler, when he discussed the reduction of quadratic forms in three variables, i.e., expressions of the type $\sum a_{ik}x_i x_k$ $(i,k = 1,2,3)$, to their principal axes, already had used what was later called (by Euler himself) the 'characteristic equation.' This equation,

$$\begin{aligned}(a_{11}-\lambda)\left[(a_{22}-\lambda)(a_{33}-\lambda)-a_{23}a_{32}\right]-a_{21}\left[a_{12}(a_{33}-\lambda)-a_{13}a_{32}\right]\\ +a_{31}\left[a_{12}a_{23}-a_{13}(a_{22}-\lambda)\right]=0,\end{aligned} \tag{14}$$

may be written as a determinantal equation involving a quadratic array with coefficients $(a_{ii}-\lambda)$ and a_{ik} for $i \neq k$. The three roots of the third-order equation in the unknown λ represent the lengths of the principal axes of the geometrical object associated with the quadratic form under investigation. Cauchy, in extending Euler's results to more than three dimensions, had shown that the characteristic equations, such as Eq. (14), remain the same whatever linear transformation is carried out with the coordinates x_i (Cauchy, 1826). Again, Sylvester put this result into the framework of a general algebraic theory by developing the theory of 'invariants of quadratic forms,' i.e., of the expressions obtained from quadratic forms that are unaffected by linear transformations of the coordinates (Sylvester, 1851). He quickly realized, for instance, that $\det(a_{ik})$, the determinant of the coefficients, is such an invariant. The theory of invariants became a powerful tool in various branches of mathematics during the second half of the nineteenth century and, a little later, also in the theory of infinite matrices.

The name 'matrix' was coined by Sylvester, who denoted by it an array of numbers (not necessarily quadratic), from which the determinant could be formed (Sylvester, 1850). Although matrices had been implicitly present in mathematics ever since Leibniz had taken up the solution of systems of linear equations (Leibniz, 1693), and Cauchy had introduced the arrangement of their coefficients into arrays (Cauchy, 1815), it was Arthur Cayley who, beginning in 1855, developed matrix theory.[19]

[19] Arthur Cayley was born on 16 August 1821 in Richmond, Surrey, the second son of Henry Cayley, an English merchant who was engaged in Russian trade at St. Petersburg, and his wife Maria Antonia Doughty. Cayley traced his father's family back to the Normans in the days of the Conquest, and even further to a noble family in Normandy (see Bell, 1937, Chapter 12). In 1829, after spending most of their time abroad, the family returned to England; young Arthur attended a private school at Blackheath and later, from his fourteenth year, King's College School in London, where he showed great talent in mathematics. At the age of seventeen he went up to Trinity College, Cambridge. There he distinguished himself as a mathematician and developed a passion for reading novels. He became Senior Wrangler in the Mathematical Tripos in 1842 and was placed first in the competition for the Smith's Prize; he was elected Fellow of Trinity College and Assistant Tutor for three years

Arthur Cayley produced a vast mathematical output over a period of fifty-four years: 996 papers, which filled the thirteen volumes of *Collected Papers* (Cayley, 1889–1898). His early work already indicated the fields, which would interest him throughout his life: these included the geometry of plane curves, the geometry of n-dimensions, elliptic functions and the theory of invariants. His investigations on algebra had a particular impact. Algebra, an ancient field of mathematics, had been greatly revived in the eighteenth century. However, in 1833 George Peacock of Cambridge University put algebra on a new level by introducing the concept of a 'symbolic algebra,' which follows combination laws similar to the ones used in arithmetic (Peacock, 1834).[20] Peacock's extension of algebra to objects other than ordinary numbers simplified the path to new concepts, such as the quaternions of Hamilton or the matrices of Cayley. Also the theory of invariants emerged as a field of algebra, and Cayley's work on this subject attracted scholars in England (George Boole, George Salmon, Joseph Sylvester), Germany (Alfred Clebsch, Paul Gordan), France (Camille Jordan, Charles Hermite) and Italy (Francesco Brioschi). It then became connected with the ideas of Bernhard Riemann, whose work would eventually prove to be one of the cornerstones of a fundamental physical theory: the general relativity theory.

Cayley's investigations on invariants cannot be separated from his other contributions to algebra, especially from those to matrix theory. His first paper on matrix theory, written in French and published in a German periodical (Cayley, 1855), already reflected the internationalism which would characterize the later development of this field. In this paper, Cayley introduced a shorthand

---

(1842–1845), an appointment which might have been renewed had he cared to take the holy orders, but Cayley did not. (Unlike his later friend Joseph Sylvester, who was not a practicing Jew but, being Jewish, was not even allowed to compete for the Smith's Prize, Cayley was an orthodox Church of England Christian; yet he preferred not to become a parson to keep his job.) At the age of twenty-five Cayley left Cambridge and, being attracted by the study of law, entered Lincoln's Inn to prepare himself for the Bar. Three years later, in 1849, he was called to the Bar and spent the following fourteen years in the routine of legal business. Though it meant a drudgery for him, it provided him an ample living and some spare time to devote himself to work in mathematics. At the Bar he also became acquainted with Sylvester who shared his interest in mathematics and classical literature. Their mathematical discussions in these years laid the foundations of the theory of algebraic invariants which would become fruitful in many fields of mathematics. In 1863 a new mathematical chair, the Sadlerian professorship, was established in the University of Cambridge, and Cayley, then forty-two, accepted it though it meant a cutback in his earnings. At the same time he married Susan Moline; they had a happy marriage, and had two children, a son and a daughter. His life in Cambridge was devoted to mathematical research and university administration, and he had a most successful career. Only once did he leave his university to spend, at Sylvester's invitation, six months during 1881–1882 at Johns Hopkins University, where he lectured on Abelian functions. Cayley received numerous honours, including the Royal medal (1859) and the Copley medal (1882) of the Royal Society. He died on 26 January 1895 in Cambridge. (For further details, see the article on Arthur Cayley by J. D. North in the *Dictionary of Scientific Biography*, Vol. 3, 1971, pp. 162–170.)

[20] George Peacock, born in Denton (England) on 9 April 1791, obtained his B.A. (1812) and M.A. (1816) from Cambridge University. He served as Tutor at Trinity College from 1823, and became Lowndean Professor of Astronomy and Geometry in 1836. He revived mathematical analysis in England and introduced Leibniz' notation in calculus. He worked and wrote books on calculus, arithmetic and algebra. Peacock died in Ely on 8 November 1858.

for writing the system of linear equations,

$$\begin{aligned} \xi &= \alpha x + \beta y + \gamma z + \cdots, \\ \eta &= \alpha' x + \beta' y + \gamma' z + \cdots, \\ \zeta &= \alpha'' x + \beta'' y + \gamma'' z + \cdots, \\ &\vdots \end{aligned} \tag{15}$$

as

$$(\xi, \eta, \zeta \ldots) = \begin{vmatrix} \alpha & , \beta & , \gamma & , \ldots \\ \alpha' & , \beta' & , \gamma' & , \ldots \\ \alpha'' & , \beta'' & , \gamma'' & , \ldots \\ & & \vdots & \end{vmatrix} \widehat{\ } (x, y, z \ldots), \tag{15a}$$

thereby casting Leibniz' formulation (i.e., Eq. (10), with the variables $x_1, x_2, x_3, \ldots$ and $y_1, y_2, y_3, \ldots$ replaced by $\xi, \eta, \zeta, \ldots$ and $x, y, z, \ldots$, respectively) into a practical shape by using a matrix, i.e., the quadratic array between two vertical bars in Eq. (15a). However, Cayley did not stop at this formal point. He proceeded to introduce the notion of the '*matrice inverse*' (inverse matrix), which appears in the solution of the system of linear equations and which he wrote as the original matrix to the power -1 (Cayley, 1855, p. 283). Finally, he considered the process of matrix multiplication, which yields what he called the '*matrice composée*' ('composed matrix,' Cayley, 1855, p. 284). The composed matrix appears when one solves connected systems of linear equations, in which the unknown variables $(x, y, z, \ldots)$ are given as linear expressions with coefficients $a, b, c, \ldots, a', b', c', \ldots, a'', b'', c'', \ldots,$ of the variables $x', y', z', \ldots,$ and the latter as linear expressions of the type of Eqs. (15) in terms of the known variables $(\xi, \eta, \zeta, \ldots)$, i.e.,

$$(\xi, \eta, \zeta, \ldots) = \left\| \begin{vmatrix} \alpha & , \beta & , \gamma & , \ldots \\ \alpha' & , \beta' & , \gamma' & , \ldots \\ \alpha'' & , \beta'' & , \gamma'' & , \ldots \\ \vdots & \vdots & \vdots & \end{vmatrix} \quad \begin{vmatrix} a & , b & , c & , \ldots \\ a' & , b' & , c' & , \ldots \\ a'' & , b'' & , c'' & , \ldots \\ \vdots & \vdots & \vdots & \end{vmatrix} \right\| \widehat{\ } (x, y, z, \ldots). \tag{16}$$

The coefficients or values of the composed matrix were given by the matrix equation

$$\begin{vmatrix} \alpha\ ,\beta\ ,\gamma\ ,\ldots \\ \alpha'\ ,\beta'\ ,\gamma'\ ,\ldots \\ \alpha''\ ,\beta''\ ,\gamma''\ ,\ldots \\ \qquad\vdots \end{vmatrix} \quad \begin{vmatrix} a\ ,b\ ,c\ ,\ldots \\ a'\ ,b'\ ,c'\ ,\ldots \\ a''\ ,b''\ ,c''\ ,\ldots \\ \qquad\vdots \end{vmatrix}$$

$$= \begin{vmatrix} (\alpha\ ,\beta\ ,\gamma\ ,\ldots)\frown(a,a',a'',\ldots), & (\alpha\ ,\beta\ ,\gamma\ ,\ldots)\frown(b,b',b'',\ldots),\ldots \\ (\alpha'\ ,\beta'\ ,\gamma'\ ,\ldots)\frown(a,a',a'',\ldots), & (\alpha'\ ,\beta'\ ,\gamma'\ ,\ldots)\frown(b,b',b'',\ldots),\ldots \\ (\alpha''\ ,\beta''\ ,\gamma''\ ,\ldots)\frown(a,a',a'',\ldots), & (\alpha''\ ,\beta''\ ,\gamma''\ ,\ldots)\frown(b,b',b'',\ldots),\ldots \end{vmatrix}, \tag{16a}$$

where $(\alpha, \beta, \gamma, \ldots)\ (a, a', a'', \ldots)$ denotes the sum $\alpha a + \beta a' + \gamma a'' + \ldots$, etc. (Cayley, 1855, p. 284). Cayley noticed that the matrix notation could also be profitably applied in the theory of quadratic and higher-order functions.

Three years later Cayley presented a systematic treatment of matrices, in which he established the laws governing the operations with matrices in detail (Cayley, 1858). Working almost exclusively with quadratic matrices, i.e., matrices in which the number of rows equals the number of columns, he defined the 'zero' and 'unit' matrices, the former having only 0 as coefficients or elements, the latter having 1 as diagonal elements and 0 as nondiagonal elements. Cayley also showed that the operation of adding two matrices, which yields a matrix whose elements are the sums of the corresponding matrix elements (at the same position), satisfies the associate and commutative laws.[21] Cayley introduced two kinds of multiplication operations: a 'scalar' multiplication by a constant factor, which implies the multiplication of each matrix element by this factor; and the process of 'compounding' two matrices, which leads to the composed matrix on the right-hand side of Eq. (16a). Cayley found that the matrix product was associative but in general not '*convertible*,' i.e., it was noncommutative. He called two matrices A and B '*convertible*' if they had identical products AB and BA; he called them '*skew-convertible*' if they satisfied $\mathsf{AB} = -\mathsf{BA}$. Then he discussed the repeated compounding of a matrix A and observed that this operation obeyed the relation

$$\mathsf{A}^{n+p} = \mathsf{A}^n\mathsf{A}^p \qquad (n, p \text{ integers}). \tag{17}$$

He further defined the '*inverse*,' A^{-1}, of a matrix A whose determinant is

[21] Addition can be defined only for matrices having the same structure, i.e., the same number of rows and columns. These matrices need not be quadratic.

different from zero (evidently, the product $\mathbf{A}\mathbf{A}^{-1}$ gives the unit matrix), and proposed a practical method of how to construct it. He also introduced the '*transpose*,' $\mathbf{A}^T$, of a matrix, which is obtained from $\mathbf{A}$ by exchanging the rows and columns, i.e., by replacing each element a_{ik} by a_{ki}. The transpose of a product of matrices is obtained by multiplying the transpose of the individual matrix factors in the reverse order.

With the results, contained in the memoirs of 1855 and 1858, Cayley not only founded matrix methods but also developed them almost fully. During the following decades the mathematicians applied these methods to new fields and problems. Two major trends arose: on the one hand, abstract algebraic structures were studied by employing matrix methods (predominantly by Cayley and Sylvester) and, on the other, matrix methods were successfully used in problems of geometry dealing with multidimensional spaces. Cayley himself opened up the field of applications by stating an important theorem: The determinant of the difference of a matrix $\mathbf{A}$ minus a diagonal matrix, whose elements are equal to $\det(\mathbf{A})$, will be equal to zero (Cayley, 1858). He proved this theorem for 2×2 and 3×3 matrices. In the first case, for example, he easily found that

$$\begin{aligned} &\det \begin{vmatrix} a - \begin{vmatrix} a & b \\ c & d \end{vmatrix}, & b \\ c, & d - \begin{vmatrix} a & b \\ c & d \end{vmatrix} \end{vmatrix} \\ &= (ad - bc) \det \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} - (a + d) \det \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \det \begin{vmatrix} a & b \\ c & d \end{vmatrix}^2 = \det \begin{vmatrix} 0 & 0 \\ 0 & 0 \end{vmatrix}. \end{aligned} \tag{18}$$

Equation (18) was the first equation involving matrices to be written down, and it says simply that the matrix is a root of its own characteristic equation.

Even before matrix theory was established, William Rowan Hamilton had proved a similar theorem for quaternions (Hamilton, 1853, p. 566). The close connection between many results in the theory of quaternions and matrix theory brought up the question, whether Hamilton's hypernumbers invented in 1843 had inspired Cayley directly to introduce matrices. To an explicit inquiry of Peter Guthrie Tait, however, Cayley replied: 'I certainly did not get the notion of a matrix in any way through quaternions; it was either directly from that of a determinant, or as a convenient mode of expression of the equation $x' = ax + by$, $y' = cx + dy$' (Cayley to Tait, 1894; quoted in Knott, 1911, p. 164). The quaternions were indeed a special type of matrices, as Cayley had noticed rather early. Therefore, the operations with matrices and those with quaternions (and other kinds of hypernumbers) obeyed largely the same rules. But in founding matrix methods, which Sylvester called 'a second birth of algebra, its *avatar* [incarnation] in a new and glorified form' (Sylvester, 1884, p. 35), Cayley had in mind Leibniz' general problem of treating systems of linear equations rather than Hamilton's favourite mathematical objects, the quaternions.

After Cayley and Sylvester, Ferdinand Georg Frobenius became one of the major contributors to matrix theory.[22] In a long memoir he generalized many of the specific results which had been obtained earlier; for example, he gave a universal proof of the above-mentioned theorem of Cayley and Hamilton (Frobenius, 1878). He studied the process of dividing a matrix, A, by another matrix, B, which he defined as multiplying A with the inverse of B, i.e., B^{-1}, paying proper attention to the order of the matrix factors. He also considered specific types of matrices, such as the orthogonal ones, which had been around before without having a proper definition. Frobenius now stated that a matrix, whose inverse is equal to its transpose, should be called 'orthogonal.' He further showed that any orthogonal matrix M can be expressed as follows:

$$\mathsf{M} = \frac{\mathsf{S} - \mathsf{T}}{\mathsf{S} + \mathsf{T}} = \frac{1 - \mathsf{T}'}{1 + \mathsf{T}'}\,, \tag{19}$$

where S denotes a symmetric matrix (which is identical to its tranpose), T and T' denote skew-symmetric matrices, and 1 denotes the unit matrix. Besides this definition he introduced the concept of 'cogredient' (now called 'congruent') matrices for matrices A and B related by the equation

$$\mathsf{A} = \mathsf{P}^T\mathsf{B}\mathsf{P}, \tag{20}$$

where P is a matrix having a nonzero determinant and P^T is its transpose.

The notion of cogredient or congruent matrices is related to that of similar matrices, developed before by Karl Weierstrass in extending the earlier work on determinants (Weierstrass, 1868).[23] Two matrices, A and B, are said to be 'similar,' if A equals the product $\mathsf{P}^{-1}\mathsf{B}\mathsf{P}$, with P being a matrix having nonzero determinant. Similar matrices possess important properties; e.g., they have the same characteristic equations.[24] Since matrices can be interpreted geometrically as linear homogeneous transformations, similar matrices in particular may be taken as representing the same transformation referring to different systems of

[22] Ferdinand Georg Frobenius was born in Berlin on 26 October 1849. He obtained his doctorate from the University of Berlin in 1870 and became a *Privatdozent* in 1874. The following year he obtained a professorship in Zurich, from where he returned to Berlin in 1892. Frobenius made important contributions to the theory of solving systems of linear differential equations, to various aspects of algebra, to the theory of finite groups and to number theory. He died in Berlin on 3 August 1917.

[23] Karl Theodor Wilhelm Weierstrass was born in Ostenfelde, Westphalia, on 31 October 1815. He became a professor at the University of Berlin in 1864. A dominant mathematical figure, Weierstrass contributed to many fields of mathematics, including the theory of linear and bilinear forms. He died on 19 February 1897 in Berlin.

[24] Matrices can be similar only if they have the same 'rank,' a notion introduced by Frobenius in connection with determinants (Frobenius, 1879). The rank of a matrix is defined as the number of rows (or columns) of the largest submatrix having a nonzero determinant.

coordinates. Camille Jordan[25] showed further that any quadratic matrix A might be brought by a similarity transformation (involving a matrix P and its inverse) into a canonical form (Jordan, 1870), i.e.,

$$\mathsf{PAP}^{-1} = \begin{vmatrix} \mathsf{J}_1 & 0 & 0 & \dots \\ 0 & \mathsf{J}_2 & 0 & \dots \\ 0 & 0 & \mathsf{J}_3 & \dots \\ \cdot & \cdot & \cdot & \dots \end{vmatrix}. \tag{21}$$

The J_i denote quadratic matrices with l_i rows and columns, and have the form:

$$\mathsf{J}_i = \begin{vmatrix} \lambda_i & 1 & 0 & & 0 \\ 0 & \lambda_i & 1 & & 0 \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \lambda_i \end{vmatrix}, \tag{21a}$$

where λ_i are the l_i-fold roots of the characteristic equation of the original matrix A; this equation may be written as

$$f(\lambda) = (\lambda - \lambda_1)^{l_1}(\lambda - \lambda_2)^{l_2} \cdots (\lambda - \lambda_k)^{l_k}. \tag{21b}$$

Of course, the canonical form of a matrix, all roots of which are different, is a diagonal matrix. Thus, Jordan's procedure proved to be identical, in special cases, with the reduction of bilinear forms to a sum of squares or with the transformation to principal axes.

Only about twenty-five years after its inauguration by Cayley, matrix theory had reached a state of considerable perfection through the investigations of relatively few mathematicians. It was then treated as a fully developed subject in textbooks, such as Maxime Bôcher's *Introduction to Higher Algebra* (1907) which,

[25] Camille Jordan was born in Lyon, France, on 5 January 1838. A brilliant student, he followed the usual career of French mathematicians from Cauchy to Poincaré: at seventeen he entered the *École Polytechnique* and was an engineer until 1885. He found ample time for mathematical research and wrote most of his 120 papers before he retired as an engineer. From 1873 until his retirement in 1912 he taught at the *École Polytechnique* and (from 1883) simultaneously at the *Collège de France*. He was elected a member of the Academy of Sciences in 1881. Jordan died in Paris on 20 January 1922.

'Jordan's place in the tradition of French mathematics is exactly halfway between Hermite and Poincaré,' writes Dieudonné. 'Like them he was a "universal" mathematician who published papers in practically all branches of the mathematics of his time.' Jordan took an active part in the movement which started modern analysis in the last decades of the nineteenth century. His most famous contribution to topology was to realize that the decomposition of a plane into regions by a simple closed curve was susceptible of mathematical proof and to imagine such a proof for the first time. Jordan was the first to embark on a systematic development of the theory of finite groups and of its applications in the directions opened by Galois. The most profound results obtained by Jordan in algebra were his 'finiteness theorems.' His last finiteness theorem was a powerful generalization of the results obtained earlier by Hermite in the theory of quadratic forms with integral coefficients. (See the article on Camille Jordan by J. Dieudonné in the *Dictionary of Scientific Biography*, Vol. 7, 1973, pp. 167–169.)

in turn, served as an authoritative reference book on matrix theory for many years to come. For example, it was cited as such in the relevant chapters of books by Courant and Hilbert (1924) and Erwin Madelung (1925). In the meanwhile, however, another important mathematical development had begun that led to a generalization of the concept of matrices, i.e., to matrices having an infinite number of rows and columns.

The earliest steps in this development go far back into the past. Already since the times of Isaac Newton and Gottfried Wilhelm Leibniz, one had considered besides finite systems of linear equations also infinite ones, the latter involving infinitely many coefficients, to obtain the solution of certain differential equations. However, all the systems of equations which the mathematicians had treated in the eighteenth century exhibited a characteristic feature: they could be solved by the method of recursion, i.e., step by step, each step involving only a finite number of equations for a finite number of unknowns. Hence these problems did not really demand the use of infinite matrices and determinants. The situation changed drastically when Joseph Fourier, in his *Théorie analytique de la chaleur* (1822) discussed the equation of heat conduction.[26] As is well known, Fourier solved the parabolic differential equation

$$\frac{\partial^2 T}{\partial x^2} = \kappa \frac{\partial T}{\partial t}\,, \tag{22}$$

determining the space-time dependence of the temperature T $(= T(x,t))$ in a linear rod of length l (which, for convenience, may be normalized to π), having κ as the coefficient of heat conduction, with the help of the following assumption. He represented T by an infinite series, i.e.,

$$T(x,t) = \sum_{\nu=1}^{\infty} b_\nu \exp\left(-\frac{\nu^2}{\kappa^2}\, t\right) \sin(\nu x), \tag{23}$$

where b_ν are the harmonic coefficients. He substituted this series in Eq. (22), thus obtaining an infinite system of linear equations determining the b_ν $(\nu = 1, 2, \ldots)$. Fourier solved it in a bold manner, applying at times methods of questionable rigour, and finally found the result expressing each b_ν as

$$b_\nu = \frac{2}{\pi} \int_0^\pi f(x) \sin(\nu x)\, dx, \tag{23a}$$

where the function $f(x)$ is identical with $T(x,0)$, the initial values of the temperature at the space points x. He was, thus, the first to deal successfully with

[26] Jean Baptiste Joseph (Baron) Fourier was born on 21 March 1768 in Auxerre, France. He joined the *École Normale* from its founding in 1795 and soon became professor of analysis at the *École Polytechnique*. He accompanied Napoleon to Egypt in 1798 and was appointed governor of Lower Egypt; after his return to France he was made prefect of the Department of Isère. From 1816 onwards he again settled in Paris and became member (1817) and Permanent Secretary (1822) of the Academy of Sciences. Fourier died in Paris on 16 May 1830.

a genuinely infinite system of equations. However, Fourier's intuitive manner of handling mathematical methods was viewed with suspicion by his contemporaries, and the following generation of mathematicians tried different approaches to arrive at the same results. Only a century later was it possible to continue, what Fourier had started, with improved and extended methods. The so-called 'harmonic analysis' would then be developed, containing a clear formulation of Fourier's ideas, and its relation to the algebraic methods of infinite matrices would also be recognized.

The factual beginning of the theory of infinite matrices can be traced to the theory of infinite determinants. The early approach to handle these objects was first to study finite sections before one passed over to an adequate infinite limit.[27] Nearly fifty years after Fourier, Th. Kötteritzsch, an otherwise not well-known mathematician, dealt with infinite systems of linear equations ($n = \infty$) of the type of Eqs. (10) (Kötteritzsch, 1870). He reminded himself of the fact that the solution of the finite system (n finite) can be written as

$$y_k = \sum_{i=1}^{n} \frac{A_{ik}}{A} x_i, \qquad k = 1, 2, \ldots, n, \tag{24}$$

which is known as Cramer's rule.[28] In Eq. (24), A_{ik} denotes the minor of the full determinant, A, from which the ith row and the kth column have been omitted, i.e.,

$$A_{ik} = \det \begin{vmatrix} a_{11} \ldots & a_{1k} \ldots a_{1n} \\ \vdots & \\ \cancel{a_{i1} \ldots} & \cancel{a_{ik} \ldots a_{in}} \\ \vdots & \\ a_{n1} \ldots & a_{nk} \ldots a_{nn} \end{vmatrix}. \tag{24a}$$

The addition of another linear equation,

$$x_{n+1} = \sum_{k=1}^{n+1} a_{n+1,k}\, y_k, \tag{25}$$

containing new variables, x_{n+1} and y_{n+1}, to the original system of n equations, Eqs. (10), does not change the formal structure of the result, expressed by Eqs. (24). The latter are then simply replaced by the equations

$$y_k = \sum_{i=1}^{n+1} \frac{A_{ik}}{A} x_i, \qquad k = 1, 2, \ldots, n+1, \tag{24'}$$

[27] Fourier, in his treatise on heat conduction quoted above (Fourier, 1822), also made use of the idea to approach an infinite system of equations by a finite one. However, this approach did not yield directly his result, Eqs. (23) and (23a).

[28] Gabriel Cramer (1704–1752), mathematician and consultant in the fortification of his hometown of Geneva, worked on algebraic curves, on which he also wrote a book (Cramer, 1750). In the appendix to the third chapter of this book he gave the result, Eq. (24), for the cases of $n = 2$, 3, etc. It was therefore called 'Cramer's rule,' though some special cases of it had already been given by Leibniz and Cramer's contemporaries.

with the A_{ik} now denoting the minors of the determinant, which is enlarged by including the coefficients $a_{n+1,k}$ and $a_{i,n+1}$ $(i,k = 1,2,\ldots,n+1)$. Therefore, in principle, the finite system can be replaced by an infinite system. Kötteritzsch concluded, in particular, that the solution for the variables, y_k $(k = 1,2,\ldots,\infty)$, of such systems may be expressed by the infinite series

$$y_k = \frac{A_{1k}}{A}x_1 + \frac{A_{2k}}{A}x_2 + \cdots, \qquad k = 1,2,\ldots,\infty, \tag{26}$$

which have the same structure as Eqs. (24), except that one sums on the right-hand side up to infinity and employs infinite determinants, i.e., determinants of infinite matrices and their minors. Remarkably enough for the time—after Cauchy, Riemann and Weierstrass had already established new standards of mathematical rigour—Kötteritzsch did not enter into any discussion of the conditions under which his solution (26) exists. He just tried to approximate the infinite systems in steps, i.e., by first solving the finite nth section of it, then the $(n+1)$st section, etc. A similar approach was proposed independently later on by Paul Appell (1885) and, stimulated by him, by Henri Poincaré (1885).

The above method of solving infinite systems of linear equations, the so-called 'section method' ('*Abschnittsmethode*'), did not involve infinite determinants directly but only ratios of them. It remained to the American mathematician George William Hill to deal first directly with infinite determinants in his work on the influence of celestial bodies on the motion of the moon (Hill, 1877, 1886).[29] Hill investigated in particular the solution of the second-order linear

[29] George William Hill, born in New York City on 3 March 1838, received his education at Rutgers University (A.B., 1859, M.A., 1862). In 1861 he joined the staff of scientists working in Cambridge, Massachusetts, on the *American Ephemeris and Nautical Alamanac*. At the *Almanac* office Hill was assigned the task of calculating the American ephemeris. Carolyn Eisele writes:

> When Simon Newcomb became director of the *American Ephemeris* in 1877, he undertook the reconstruction of the theories and tables of lunar and planetary motion. Hill was induced to work on the theories of Jupiter and Saturn, known to be exceptionally difficult in the determination of their mutual perturbations. Because the *Nautical Almanac* office had meanwhile been transferred to Washington to be under the more immediate jurisdiction of the Navy Department, Hill resided there for a ten-year period beginning in 1882. His success with the Newcomb assignment represented one of the most important contributions to nineteenth-century mathematical astronomy. The calculation of the effects of the planets on the moon's motion was a particular case of the famous three-body problem, which dates back to Newton (1686).
>
> Hill's "Researches in the Lunar Theory," published in the first issue of *American Journal of Mathematics* (1878), had, through its introduction of the periodic orbit, initiated a new approach to the study of three mutually attracting bodies. F. R. Moulton wrote in 1914 that no earlier work had approached it in practical application and no subsequent work had then surpassed it. The article became fundamental in the development of celestial mechanics.
>
> The memoir of 1877 entitled *On the Part of the Motion of the Lunar Perigee Which Is a Function of the Mean Motions of the Sun and Moon* contains the incontrovertible evidence of Hill's mathematical genius. He was led to a differential equation, now called Hill's equation, that is equivalent to an infinite number of algebraic linear equations. Hill showed how to develop the infinite determinant corresponding to these equations. (*Dictionary of Scientific Biography*, Vol. VI, 1972, p. 399)

G. W. Hill was honoured in his lifetime by the bestowal of advanced degrees and medals and by honorary memberships in the most prestigious professional scientific societies throughout the world: Member, U.S. National Academy of Sciences (1874); President, American Mathematical Society (1894–1896); Gold Medal of the Royal Astronomical Society for his researches on lunar theory (1887); Foreign Member of the Royal Society, the Paris Academy, and the Belgian Academy. He died in West Nyack, New York, on 16 April 1914.

differential equation

$$-\frac{d^2w}{d\tau^2} = \theta w \tag{27}$$

for the position coordinate w (the perigee of the moon) as a function of the time coordinate τ. The factor θ on the right-hand side of Eq. (27) takes into account the action of all (i.e., infinitely many) celestial bodies on the moon, and it is therefore given by an infinite sum

$$\theta = \sum_{k=-\infty}^{+\infty} \theta_k \zeta^{2k}, \qquad k = 1, 2, \ldots, \infty, \tag{27a}$$

where $\zeta = \exp(\tau i)$ and $\theta_k = \theta_{-k}$. Hill now proposed that the solution for the position coordinate takes the form

$$w = \sum_{k=-\infty}^{\infty} b_k \zeta^{c+2k}, \tag{28}$$

with an unknown coefficient c. By substituting the right-hand side of Eq. (28) for w in Eq. (27) and also replacing θ by the infinite sum, due to Eq. (27a), he obtained an infinite system of equations for the infinitely many coefficients b_k, i.e.,

$$\begin{aligned}
\ldots + [-2]b_{-2} - \quad \theta_1 b_{-1} - \theta_2 b_0 - \theta_3 b_1 - \theta_4 b_2 - \ldots &= 0, \\
\ldots - \quad \theta_1 b_{-2} + [-1]b_{-1} - \theta_1 b_0 - \theta_2 b_1 - \theta_3 b_2 - \ldots &= 0, \\
\ldots - \quad \theta_2 b_{-2} - \quad \theta_1 b_{-1} + [0]b_0 - \theta_1 b_1 - \theta_2 b_2 - \ldots &= 0, \\
\ldots - \quad \theta_3 b_{-2} - \quad \theta_2 b_{-1} - \theta_1 b_0 + [1]b_1 - \theta_1 b_2 - \ldots &= 0, \\
\ldots - \quad \theta_4 b_{-2} - \quad \theta_3 b_{-1} - \theta_2 b_0 - \theta_1 b_1 + [2]b_2 - \ldots &= 0,
\end{aligned}$$

with

$$[k] = (c + 2k)^2 - \theta_0, \qquad k = 0, \pm 1, \pm 2, \ldots . \tag{29}$$

The matrix of the coefficients of the system, Eqs. (29), is symmetric. To solve this system, Hill eliminated all b's except b_0 and constructed an inhomogeneous system. For the latter to have a solution, its determinant must be zero; hence there follows an equation which determines the value of the constant c in Eq. (28). Finally Hill worked out the solution for the infinitely many ratios b_k/b_0 and demonstrated that the motion of the moon's perigee was indeed periodic.

Hill's procedure, like Fourier's, for solving infinite systems of equations, was again very bold and not overly rigorous according to the standards of his time, although he tried to justify it by checking the convergence of the infinite determinant. In particular, he argued: 'All those determinants must be regarded as having a central constituent; when in computing in succession the deter-

minants formed from the 3^2, 5^2, 7^2, etc., constituents symmetrically situated with respect to the central constituent, we approach, without limit a determinate magnitude, the determinant may be called convergent, and the determinate magnitude is its value' (Hill, 1886, p. 26). A similar statement had already been made earlier by Kötteritzsch, but Hill surpassed his predecessor by giving a mathematical condition. 'In the present case,' he said, 'there can be scarcely a doubt that as long as $\sum \theta_k \zeta^{2k}$ is a legitimate expansion of θ, the determinant . . . must be regarded as convergent' (Hill, 1886, p. 26).

Hill's argument for the convergence of the infinite determinant was received with suspicion by the mathematicians, except by Henri Poincaré, who became interested in infinite systems of linear equations a second time because of Hill's work.[30] Poincaré immediately sought to put Hill's reasoning on a safer mathematical ground (Poincaré, 1886). To that end he considered the special type of matrices, (a_{ik}), whose diagonal elements, a_{ii} $(i = 0, \pm 1, \pm 2, \ldots)$, are unity. Then he took the determinant, Δ_n, of finite matrices, consisting of $(n \times n)$ elements, symmetric with respect to the diagonal of the infinite matrix, and calculated the infinite determinant Δ from a limiting process, i.e., as

$$\Delta = \lim_{n \to \infty} \Delta_n. \tag{30}$$

Poincaré proved that the limit (30) existed, provided the sum of the absolute values of the matrix elements, $|a_{ik}|$, with $i \neq k$, was finite. Finally he concluded that the convergence of the right-hand side of Eq. (30) was not altered if one row of the infinite matrix was replaced by a bounded sequence x_j, $j = 0, \pm 1, \pm 2, \ldots$. These results served to establish an essential part of Hill's intuitive procedure on firm mathematical grounds.

Hill and Poincaré had, without doubt, carried forward the notion of infinite matrices, although they had only discussed the determinants of the latter. The same is true of some of the next steps in the development. Gösta Mittag-Leffler, the Swedish mathematician, who, by reprinting Hill's memoir of 1877 in his journal, the *Acta Mathematica* (Hill, 1886), had directed Poincaré's attention to it, now got his student Helge von Koch interested in infinite systems of linear equations and their applications to various mathematical problems (Koch, 1892).[31] Von Koch first treated the so-called 'normal' determinants, i.e., determi-

[30] In introducing a reprint of Hill's work, Poincaré noted: 'The solution adopted by Mr. Hill is as original as it is bold . . . Did one have the right to set the determinant of these equations equal to zero? Mr. Hill ventured to do so and it was a very daring thing to do; until then an infinite number of linear equations had never been considered; determinants of infinite order had never been studied; no one even knew how to define them, and it was not certain that it was possible to give a precise meaning to this notion' (Poincaré, in Hill, 1905, p. xiii).

It is surprising that Poincaré did not refer to the occurrence of infinite systems of linear equations in Fourier's work on heat conduction. He did, however, know about Kötteritzsch's paper and stated: 'But his paper was hardly known in the scientific world and in any case was not known to Mr. Hill' (Poincaré, in Hill, 1905, p. xiii).

[31] Helge von Koch was born in Stockholm on 25 January 1870. He began publishing in 1890 and worked mainly on the theory of infinite systems of linear equations; he also worked on the theory of differential equations and the theory of numbers. He died in Stockholm on 11 March 1924.

nants having the form

$$A = \det \begin{vmatrix} 1 + a_{11}, & a_{12}, & a_{13}, & \cdots \\ a_{21}, & 1 + a_{22}, & a_{23}, & \cdots \\ a_{31}, & a_{32}, & 1 + a_{33}, & \cdots \\ & \vdots & & \end{vmatrix}, \tag{31}$$

with the a_{ik} satisfying the condition

$$\sum_{i,k=1}^{+\infty} |a_{ik}| < \infty. \tag{31a}$$

These determinants were slightly more general than Poincaré's (in whose case, $a_{ii} = 0$). He also went beyond Poincaré by studying, besides the determinants, certain subdeterminants obtained from the former by omitting a finite number of rows and columns. Since these subdeterminants or (higher) minors were again normal determinants, von Koch could treat both inhomogeneous systems of linear equations, e.g.,

$$y_i + \sum_{k=1}^{\infty} a_{ik}\, y_k = x_i \qquad \text{with } |x_i| < \infty, \tag{32}$$

as well as homogeneous systems (with $x_i = 0$), which, in general, have solutions only if the determinant, A, equals zero—in which case a (higher) minor of A will be different from zero.

Von Koch's early interest in infinite systems of linear equations was connected with the search for solutions of the Fuchs differential equation. He found that solutions of this equation existed under certain severely restrictive conditions.[32] In order to remove them he delved deeper into the theory of infinite determinants and proved the following theorem: For the convergence of A, an infinite determinant given by Eq. (31), it is sufficient that the coefficients satisfy the conditions

$$\prod_{i=1}^{\infty} (1 + a_{ii}) < \infty \qquad \text{and} \qquad \sum_{i=1}^{\infty} \sum_{k=1}^{\infty} |a_{ik}| < \infty, \tag{33}$$

[32] The Fuchs differential equation reads:

$$\frac{d^n y}{dx^n} + P_2(x) \frac{d^{n-1} y}{dx^{n-1}} + \cdots + P_n(x)\, y = 0,$$

where each coefficient, $P_r(x)$, $r = 2, \ldots, n$, is expressed as a Laurent series in x, i.e., as $\sum_{\lambda=-\infty}^{+\infty} a_{r\lambda} x^{\lambda}$ (valid in an anulus in the complex plane around $x = 0$). The solution, y, can again be expanded in a Laurent series, and the substitution of the latter in the differential equation gives rise to a system of linear equations for the coefficients of the Laurent series for y.

where $|a_{ik}|$ is the absolute value of the coefficient a_{ik} $(i \neq k)$. Moreover, a convergent infinite determinant of the type (31) may be expanded into subdeterminants (or minors) in exactly the same manner as a finite determinant—that is, the determinant is essentially the sum of all possible subdeterminants. From the latter result it became evident that in the case of an infinite matrix, whose determinant assumes a finite value, the origin (or centre) could be chosen arbitrarily. Finally, von Koch concerned himself with another special type of determinants, the so-called 'uniformly convergent' determinants, which were again of the form (31), with the coefficients a_{ik} satisfying the conditions

$$\sum_{i=1}^{\infty} |a_{ii}| < \infty \qquad \text{and} \qquad \sum_{i,k=1}^{\infty} |a_{ik}|^2 < \infty, \qquad i \neq k. \tag{34}$$

He showed that these determinants assumed finite values and drew certain conclusions for the systems of linear equations having such determinants.[33]

On the whole, von Koch's work remained rather formal, and he did not consider applications of infinite determinants to mathematical problems other than the solution of the Fuchs differential equation. This task was performed by Erik Ivar Fredholm, another student of Mittag-Leffler's; he employed infinite matrices and determinants in a new powerful approach to the theory of linear integral equations.[34] Stimulated by the earlier work of Poincaré (1894, 1895), Fredholm began by attempting to solve the integral equation for the unknown function $\phi(x)$,

$$\phi(x) + \int_a^b K(x, y)\phi(y)\, dy = f(x), \qquad a \leqslant x \leqslant b, \tag{35}$$

where $f(x)$ and the kernel $K(x, y)$ are given functions of the variables (Fredholm, 1900, 1903). He divided the interval of integration into n equal parts at the points

$$\begin{gathered} x_1 = a + \delta, \qquad x_2 = a + 2\delta, \ldots, \qquad x_{n-1} = a + (n-1)\delta, \\ x_n = a + n\delta = b \end{gathered} \tag{36}$$

and denoted the corresponding values of the functions $\phi(x)$, $f(x)$ and $\delta \cdot K(x, y)$ at these points by ϕ_i, f_i and K_{ij}. Then he showed that the integral equation (35) was equivalent to a system of n linear equations for the unknown variables,

[33] Actually, von Koch's original definition of uniformly convergent determinants was different from the above one, but it is equivalent.

[34] Erik Ivar Fredholm was born on 7 April 1866 in Stockholm. He studied at the Stockholm Polytechnic (1885–1886), the University of Uppsala (1886–1888), and the University of Stockholm (1888–1893). He received his doctorate from Uppsala in 1898 and became a professor of mathematics and physics at the university of Stockholm in 1906. He died on 17 August 1927 in Mörby, Sweden.

Fredholm is known for his work on functional equations, especially integral equations, whose modern theory he inaugurated. He was a member of the Royal Swedish Academy of Sciences and a corresponding member of the Academy of Sciences, Paris.

$\phi_1, \ldots, \phi_n$, in the limit of arbitrarily large n. The latter system could be written as

$$\begin{aligned}
(1 + K_{11})\phi_1 + \quad & K_{12}\phi_2 + \cdots + \quad K_{1n}\phi_n = f_1, \\
K_{21}\phi_1 + (1 + K_{22})\phi_2 + \cdots + \quad & K_{2n}\phi_n = f_2, \\
& \vdots \\
K_{n1}\phi_1 + \quad & K_{n2}\phi_2 + \cdots + (1 + K_{nn})\phi_n = f_n.
\end{aligned} \tag{37}$$

Thus, Fredholm recognized that the theory of infinite systems of linear equations could be used to solve the integral equation (35), provided A, the determinant of the system (37), and all its subdeterminants, assumed finite values in the limit of infinite n. In particular, he was able to employ the generalization of Cramer's rule, Eq. (24), to infinite systems and to obtain the formal solution $\phi(x)$ as

$$\phi(x) = f(x) - \int_a^b \frac{A\binom{x}{y}}{A} f(y)\, dy, \tag{38}$$

where $A\binom{x}{y}$ corresponds to the minor of the determinant A (in which the ith row and the jth column have been dropped).

Fredholm's work on the theory of integral equations not only opened the field of the most important applications of infinite determinants, it initiated an even more important line of development in Göttingen, where Hilbert became interested in this subject.[35] As it happened to any other field which attracted Hilbert's attention, the theory of integral equations was generalized, fully developed and applied to new problems by him. During the first decade of the twentieth century this subject, especially the theory of infinite bilinear forms into which Hilbert turned the problem, moved into the centre of mathematical interest in Göttingen. Hilbert himself devoted six long memoirs to the theory of integral equations (Hilbert, 1904a, b; 1905; 1906a, b; 1910) and assembled them in a book (Hilbert, 1912c); he also supervised numerous doctoral theses of gifted students in this field. In his fourth communication (Hilbert, 1906a), he established the connection between linear integral equations and the theory of bilinear forms, giving Fredholm's earlier treatment a new direction.

Hilbert's main idea was to make use of the combined algebraic–geometric concepts, which had come into use since Euler's transformation of ellipsoids to principal axes. Hence he worked with quadratic forms, bilinear forms and linear forms of variables, extending the previous results right away to the case of infinitely many variables x_p, y_q $(p, q = 1, 2, \ldots, \infty)$. For these forms he introduced a convenient notation borrowed from the theory of n-dimensional spaces, like the scalar product of two linear objects, (x, y), which is defined as the sum

[35] Fredholm's first publication on integral equations (Fredholm, 1900) was brought to Hilbert's Seminar by a Swedish student in the winter semester of 1900–1901 (see Reid, 1970, p. 85). This student was E. Holmgren.

$x_1 y_1 + \cdots + x_n y_n$. Denoting a general infinite bilinear form by $A(x, y)$, where

$$A(x, y) = \sum_{p,q=1}^{\infty} a_{pq} x_p x_q, \tag{39}$$

Hilbert defined the 'folding' ('*Faltung*') of two such forms as (Hilbert, 1906a, pp. 160 and 179)

$$A(x, \cdot)B(\cdot, y) = \sum_{p,q,r=1}^{\infty} a_{pq} b_{qr} x_p y_r. \tag{40}$$

Evidently the right-hand side of this equation involves the product of two infinite matrices, (a_{pq}) and (b_{qr}). Such a matrix product occurs, for instance, if one wants to solve the equation

$$\overline{K}(x, y) - K(x, \cdot)\overline{K}(\cdot, y) = (x, y), \tag{41}$$

for the unknown bilinear form $\overline{K}(x, y)$. This equation is equivalent to a linear integral equation of the type of Eq. (35).

In order to derive more specific results, Hilbert considered two special cases of bilinear forms, which he called 'bounded' ('*beschränkte*') forms and 'completely continuous' ('*vollstetige*') forms—the latter are now called 'compact' forms. A bounded bilinear form satisfies the following condition:

$$|A_n(x, y)| = \left| \sum_{p,q=1}^{n} a_{pq} x_p x_q \right| \leqslant M \left(\sum_{p=1}^{n} x_p^2 \sum_{q=1}^{n} y_q^2 \right)^{1/2}, \tag{42}$$

i.e., any nth section ($n = 1, 2, 3, \ldots$) of the bilinear form possesses an upper bound determined by the finite number M (Hilbert, 1906a, p. 176).[36] To define a completely continuous form Hilbert started from the concept of a completely continuous function $F(x_1, x_2, \ldots)$ in the variables $x_1, x_2, \ldots$. F is said to be completely continuous if the values of $F(x_1 + \epsilon_1, x_2 + \epsilon_2, \ldots)$ converge towards the values of $F(x_1, x_2, \ldots)$ whenever the $\epsilon_1, \epsilon_2, \ldots$ go to zero. While any bounded linear form (i.e., a form $\sum l_p x_p$ with $|\sum l_p x_p| \leqslant N(\sum x_p^2)^{1/2}$, N being finite) is automatically completely continuous, this is not so in the case of bounded bilinear forms. However, a bounded bilinear form (bound M) which satisfies the condition

$$\begin{aligned} &|A(x, y) - A(x^{(\nu)}, y^{(\nu)})| \\ &\quad \leqslant M \left\{ \left[\sum x_p^2 \sum (y_q - y_q^{(\nu)})^2 \right]^{1/2} + \left[\sum (x_p - x_p^{(\nu)})^2 \sum y_q^2 \right]^{1/2} \right\}, \end{aligned} \tag{43}$$

[36] Hilbert always assumed that the sums $\sum x_p^2$ and $\sum y_q^2$ have unity as the upper bound. An immediate consequence of the definition of a bounded bilinear form is that the transpose, $A(y, x)$, of the bounded form, $A(x, y)$, is also bounded by the same number M.

for any sequences, $x_1^{(\nu)}, x_2^{(\nu)}, \ldots$ and $y_1^{(\nu)}, y_2^{(\nu)}, \ldots$ $(\nu = 1, 2, \ldots)$—that converge quadratically towards the values $x_1, x_2, \ldots$ and $y_1, y_2, \ldots$—turns out to be completely continuous. For example, a bilinear form $A(x, y)$, whose coefficients (a_{pq}) are square summable—i.e., $\sum_{p,q} a_{pq}^2$ exists—is always completely continuous (Hilbert, 1906a, p. 203, Theorem VI).

In his theory of infinite bilinear forms Hilbert tried to generalize the most important results already known from the theory of finite matrices, thereby pioneering the methods of infinite matrices. He particularly concerned himself with the problem of transforming quadratic forms, i.e., the special bilinear forms, $A(x, x)$, to principal axes. He showed that such a form can always be reduced to a sum of terms, $\alpha_1^2 x_1^2 + \alpha_2^2 x_2^2 + \cdots$, plus an integral. Hilbert called the α's the 'reciprocal eigenvalues' (nowadays they are defined as the eigenvalues themselves) of the quadratic form; all of them constitute the 'discrete spectrum' of the latter. On the other hand, the path of the integral may be associated with the 'continuous spectrum' of $A(x, x)$ (Hilbert, 1906a, p. 198, Theorem III). He further proved that a completely continuous quadratic form possesses only a discrete spectrum (Hilbert, 1906a, p. 205, Theorem VII).

Hilbert confined himself to establishing the fundamental concepts of the theory of infinite matrices and left the development of details to his collaborators, especially to Ernst Hellinger and Otto Toeplitz.[37,38] In a series of papers they translated Cayley's familiar operations on finite matrices into the new theory (Hellinger and Toeplitz, 1906, 1910). They noticed, in particular, that sums and differences of infinite matrices could be carried out as in the finite case and that Hilbert's procedure of folding infinite bilinear forms was equivalent to matrix multiplication, the latter being associative but not commutative. Further,

[37] Ernst Hellinger, born on 30 September 1883 in Striegau, Silesia, grew up in Breslau and studied at the Universities of Heidelberg, Breslau and Göttingen. In 1907 he received his doctorate with a thesis entitled '*Über die Orthogonalinvarianten quadratischer Formen von unendlich vielen Variablen*' ('On the Orthogonal Invariants of Quadratic Forms of Infinitely-Many Variables'), which was supervised by Hilbert. From 1907 to 1909 Hellinger served as an assistant in Göttingen, and from 1909 to 1914 as *Privatdozent* at the University of Marburg. Then he became a professor of mathematics at the newly founded University of Frankfurt. Being Jewish, he was forced to retire from this position in 1936, and at the end of 1938 he was held for a short time in a concentration camp. Upon his release he managed to go to the United States. He taught at Northwestern University, Evanston, Illinois, first as lecturer and later as full professor. In 1949 he took a position at the Illinois Institute of Technology, but fell ill soon afterwards and died on 28 March 1950 in Chicago.

Hellinger worked mainly on the theory of quadratic forms and infinite matrices. He also wrote, together with Otto Toeplitz, the review article on integral equations and systems of linear equations for the *Encyklopädie der mathematischen Wissenschaften* (Hellinger and Toeplitz, 1927). In later years he also became interested in certain topics of the history of mathematics.

[38] Otto Toeplitz, born on 1 August 1881 in Breslau, studied mathematics in Breslau and Berlin and became *Privatdozent* at the University of Göttingen in 1907. Later he became a professor of mathematics at the Universities of Kiel (1920–1928) and Bonn (from 1928) until he was dismissed in 1935 because of the Nazi racial laws. In 1939 he went to Palestine as scientific advisor to the administration of Hebrew University, but died soon afterwards on 15 February 1940.

Most of Toeplitz' papers dealt with quadratic forms of infinitely many variables. Later, he went more and more into the history of mathematics and was chiefly concerned with the mathematical ideas in Plato's work.

they were able to define the inverse of an infinite matrix and to prove that many theorems that were valid for finite matrices were also valid in the case of infinite matrices.[39] Erhard Schmidt, another disciple of Hilbert's, proceeded in a different direction.[40] He started to reverse the trend towards algebrization of mathematics, which, since Descartes, had led to such spectacular success. From his work emerged the powerful concept of 'Hilbert space,' which is built up from normed sequences

$$z = \{z_p\} \qquad \text{with } \|z\|^2 = \sum_{p=1}^{\infty} z_p \bar{z}_p < \infty, \tag{44}$$

where the double bars, $\|z\|$, denote the norm of the vector z with complex components z_p, and $\bar{z}_p$ denotes the complex conjugate to z_p (Schmidt, 1908).[41] The geometric point of view was already present in Hilbert's fourth communication on integral equations (Hilbert, 1906a), but Schmidt extended it considerably. It is interesting to note, however, that Schmidt sought to avoid the geometrical language altogether; e.g., he spoke of 'functions with integer index' rather than of points or vectors. In spite of this, the translation of his concepts into the modern language of Hilbert spaces is obvious. Any infinite sequence, $\{z_p\}$, can be regarded as a vector in an infinite-dimensional space; it can be normalized to unity by dividing each component z_p by the norm, $\|z\|$. Two such vectors, z and w, may be linearly dependent or not. A particular case of independent vectors is that of orthogonal ones, i.e., vectors z and w, for which the expression $\sum z_p \bar{w}_p$ (in geometrical language, the scalar product) is zero. From any set of n independent vectors one may construct n orthogonal unit vectors with the help of the so-called 'Schmidt orthogonalization procedure' (Schmidt, 1908). The total Hilbert space is finally spanned by a complete set of infinitely many unit vectors.

The methods which Schmidt introduced, and their subsequent geometrization, proved to be very helpful in constructing the solutions of infinite systems of linear equations involving real or complex variables. They, thus, also contributed to the theory of infinite matrices. In 1910 Hellinger and Toeplitz presented what they called an 'axiomatic' treatment of the calculus of infinite matrices; that is, they expressed the theory of the latter in a general form, regardless of its specific application to integral equations or algebraic problems (Hellinger and Toeplitz,

[39] For example, a bounded matrix, whose left-inverse exists, also possesses a unique right-inverse, which is identical to the (likewise unique) left-inverse (Toeplitz, 1907a, p. 106).

[40] Erhard Schmidt was born on 4 January 1876 in Dorpat, Estonia. He studied mathematics in Berlin and Göttingen (doctorate under Hilbert, 1905). In 1906 he became *Privatdozent* at the University of Bonn; two years later he was appointed professor of mathematics at the University of Zurich. From there he was invited ('*berufen*') in 1910 to succeed Paul Gordan in Erlangen. A year later Schmidt moved to the University of Breslau and stayed there until he obtained a chair of mathematics at the University of Berlin in 1917. Schmidt died in Berlin on 6 December 1959. His greatest contributions to mathematics were contained in his early work on functional analysis.

[41] The generalization from real variables x to complex variables z is rather obvious. Instead of Hilbert's quadratic forms with symmetric coefficients one has then to take quadratic forms of complex variables, with the coefficients forming a Hermitian matrix.

1910). The mathematical theory of infinite matrices had thus been perfected by relatively few people in less than twenty-five years since Poincaré's discovery of Hill's work on the moon's perigee in 1886. Poincaré had stimulated Fredholm's investigations on integral equations, and these had brought Hilbert to the scene. Hilbert and his students had worked out the theory in a way that not much seemed to have remained to be done. When in April 1909 Henri Poincaré was invited to deliver the Wolfskehl lectures at the University of Göttingen and announced his intention to devote a part of them to integral equations, the local mathematicians were very surprised that somebody could still say anything new on this subject in Göttingen. Poincaré did, and by his lectures (Poincaré, 1910) he closed a development in which he and his host David Hilbert had played a leading role. In Poincaré's audience was Max Born, who would profit most from an application of the theory of integral equations, especially in its algebraic version, to problems of theoretical physics.

I.3 Early Applications of Matrix Methods in Physics

Matrix methods had proved to be very useful in a variety of mathematical investigations: systems of linear equations, linear differential and integral equations, etc. But unlike other mathematical tools invented in the nineteenth century, such as quaternions and vectors, they did not seem to arouse the interest of physicists—notwithstanding Peter Guthrie Tait's prophecy that 'Cayley is forging the weapons for future generations of physicists' (Tait, quoted in Knott, 1911, p. 166). One wonders why this was so; why, for instance, such special objects as the quaternions made their way into the physical literature, but not the more general matrices?[42] After all, since the times of Joseph Louis Lagrange, quadratic forms and their determinants had played some role in mechanics—as, for instance, in the so-called 'secular' equation, whose roots serve to describe the slow periodic motions of, say, a system of several celestial bodies. However, there occurred no physical problem in the nineteenth century, in which one sought to make use of the specific properties of matrices, although in retrospect one might think of the possibility of applying them in the description of symmetry properties of crystals. When, towards the end of the nineteenth century, physicists

[42] Hamilton applied his quaternions, for instance, to lunar theory and to light propagation in biaxial crystals. James Clerk Maxwell, in his *Treatise on Electricity and Magnetism* (1873), also referred to quaternions. And Peter Guthrie Tait (1831–1901), the influential British mathematician and physicist, tried to persuade his contemporaries to make use of them. Interestingly enough, however, he did not even succeed with his friend William Thomson (Lord Kelvin), who remarked later:

> We had a thirty-eight year war over quaternions. He [i.e., Tait] had been captured by the originality and extraordinary beauty of Hamilton's genius in this respect, and had accepted I believe definitely from Hamilton to take charge of quaternions after his death, which he had most loyally executed. Times without number I offered to let quaternions into Thomson and Tait [*Treatise on Natural Philosophy* (1867)] if he could only show that in any case our work could be helped by their use. You will see that from the beginning to the end they were never introduced. (Lord Kelvin to G. Chrystall, 1910; quoted in Knott, 1911, p. 185)

started to pay attention to the transformations of coordinate systems in space and time, Cayley's matrices seemed to offer themselves as *the* tool for a proper mathematical formulation. Still one had to wait much longer until Hermann Minkowski in Göttingen made the first application of finite matrices in physics, and this too only after his friend Hilbert had been interested for several years in developing the methods of infinite matrices.

In his great memoir on the electrodynamics of moving bodies, Minkowski introduced matrix notation and formulated the fundamental equations of electrodynamics in matrix language (Minkowski, 1908, Section 11). He restricted himself to 4×4 matrices, with the help of which he could express the linear transformations of the four space-time coordinates, x_1 $(= x)$, x_2 $(= y)$, x_3 $(= z)$ and x_4 $(= ict)$. A linear transformation, $x_h \to x'_k$, given by the equations

$$x_h = \sum_{k=1}^{4} a_{hk} x'_k, \qquad h = 1, 2, 3, 4, \tag{45}$$

which satisfies the condition

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = x_1'^2 + x_2'^2 + x_3'^2 + x_4'^2, \tag{46}$$

Minkowski identified with a Lorentz transformation. He discovered that in this case the matrix of the coefficients, $\mathsf{A} = (a_{hk})$, obeys the equation

$$\mathsf{A}^T \mathsf{A} = 1, \tag{47}$$

where A^T is the transpose of the matrix A and 1 denotes the unit 4×4 matrix. In the case of a Lorentz transformation, in particular, the determinant of the matrix A assumes the value 1. The coefficients a_{14}, a_{24}, a_{34}, a_{41}, a_{42}, and a_{43} are purely imaginary and a_{44} assumes a positive value.

Minkowski also formed a 4×4 matrix from the vector components of the electric and magnetic fields, i.e., $\mathbf{E} = (E_x, E_y, E_z)$ and $\mathbf{H} = (H_x, H_y, H_z)$, namely, the skew-symmetric ('*alterniernde*') matrix f,

$$\mathsf{f} = \begin{vmatrix} 0, & H_z, & -H_y, & -iE_x \\ -H_z, & 0, & H_x, & -iE_y \\ H_y, & -H_x, & 0, & -iE_z \\ iE_x, & iE_y, & iE_z, & 0 \end{vmatrix}, \tag{48}$$

and showed that under a Lorentz transformation, described by the matrix A, it passes over into f′, with

$$\mathsf{f}' = \mathsf{A}^T \mathsf{f} \mathsf{A} = \mathsf{A}^{-1} \mathsf{f} \mathsf{A}. \tag{49}$$

He further introduced the nonquadratic 1×4 matrix of differential operators,

$|\partial/\partial x_1, \partial/\partial x_2, \partial/\partial x_3, \partial/\partial x_4|$, calling it 'the differential operator "lor,"' and rewrote Maxwell's equations for the current, $\mathbf{j} = (j_x, j_y, j_z)$, and the charge density ρ as

$$\mathsf{lor\,f} = -\mathsf{s}, \tag{50}$$

with s being the 1×4 matrix,

$$\mathsf{s} = |j_x, j_y, j_z, i\rho|. \tag{50a}$$

(In modern notation, s is replaced by the four-vector j_μ, $\mu = 1, 2, 3, 4$.) With the help of matrix methods Minkowski was able to apply the formal tools of the algebraic theory of invariants, in which he was a master, to the problems of special relativity theory, thereby deriving the fundamental equations of relativistic mechanics and electrodynamics in a compact and elegant manner.

Minkowski's plea for the use of matrix methods in relativity theory was not accepted in general by people afterwards. Even Arnold Sommerfeld, who was very close in spirit to Minkowski, moved away from these methods in his two papers on relativity theory published in 1910 (Sommerfeld, 1910b, c).[43] Thus, he pointed out in the very beginning of the first paper: 'I have put special emphasis on formulating the definitions and transformations of vectors as much geometrically as possible. The conceptions and calculations necessary for this purpose, which also provide a complete substitute for the matrix calculus used by Minkowski, are (if one admits imaginary coordinates) immediate generalizations of the conventional three-dimensional vector methods' (Sommerfeld, 1910b, p. 749). In particular, Sommerfeld replaced the skew-symmetric matrix f, defined in Eq. (48), by the concept of a 'six-vector,' which he interpreted geometrically as a rotation in four-dimensional space. And in all cases he sought to establish a close connection between the three-dimensional vector methods and Minkowski's form of relativity theory, thereby eliminating the use of matrix methods completely. Sommerfeld's procedure was followed by later authors, for example, by Max von Laue in his book *Das Relativitätsprinzip* (von Laue, 1911). Among the few persons who occasionally made use of matrices in their papers was Max Born, Minkowski's last collaborator. After Minkowski's death, he tried to continue the latter's work on relativity theory. In particular, he employed matrices in his research on the theory of the rigid electron (Born, 1909b)[44]; so did Gustav Mie in

[43] In the introduction to his first article on relativity theory, Sommerfeld stated explicitly: 'In this and some following investigations I want to show how remarkably the electrodynamical concepts and calculations become simplified if one lets oneself be guided by the profound space-time view of Minkowski. What follows here would hardly have offered anything new to the suddenly deceased friend; it might, however, be welcomed by some as an interpretation of Minkowski's ideas' (Sommerfeld, 1910b, p. 749).

[44] It is of interest to note that Sommerfeld demonstrated, for instance, that his use of four-dimensional vector methods simplified the result of Born's most important calculation, the one concerning the hyperbolic motion of the rigid electron. (Compare Born, 1909b, p. 27 and following pages, and Sommerfeld, 1910c, p. 670.)

his memoirs on the foundation of a theory of matter, in which he sought to unify the gravitational and electromagnetic fields (Mie, 1912a, c). Both Born and Mie referred to Minkowski, but they did not use matrix methods beyond what Minkowski had already invoked in his pioneering work of 1908.[45]

To be sure, nobody in Hilbert's mathematical circle—to which Born belonged—was afraid of applying matrix methods whenever they were needed or seemed to be adequate in the problem under investigation. For example, Ludwig Föppl studied the stability of an atom consisting of n electrons for his doctoral thesis under Hilbert (Föppl, 1912).[46] In the course of this work he had to solve the following determinantal equation for the unknown quantity $\tilde{D}$:

$$\det \begin{vmatrix} \tilde{D} & D_1 & D_2 & D_3 & \cdots & D_{n-1} \\ D_{n-1} & \tilde{D} & D_1 & D_2 & \cdots & D_{n-2} \\ D_{n-2} & D_{n-1} & \tilde{D} & D_1 & \cdots & D_{n-3} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ D_1 & D_2 & D_3 & D_4 & \cdots & \tilde{D} \end{vmatrix} = 0, \tag{51}$$

with the D's being given coefficients. Föppl's problem may be considered as the finite version of Hill's problem. The latter had treated, as we have mentioned in the previous section, the effect of an *infinite* number of celestial bodies on the motion of the moon (Hill, 1877), while Föppl was interested in the influence of a *finite* number of electrons on each other. As a result, the same type of determinants occurred in both problems.[47] The special determinant ('*Zirkulanten*') in Eq. (51), in which each subsequent row follows the preceding one by shifting the coefficients to the right, may be written as a product of simple factors, $F(\alpha_1)$, $F(\alpha_2)$, . . . , $F(\alpha_n)$. Thus, Eq. (51) assumes the simple form

$$F(\alpha_1)F(\alpha_2) \ldots F(\alpha_n) = 0, \tag{52}$$

where the factors $F(\alpha_i)$ are defined as the polynomials in α_i,

$$F(\alpha_i) = \tilde{D} + D_1\alpha_i + D_2\alpha_i^2 + \cdots + D_{n-1}\alpha_i^{n-1}, \tag{52a}$$

[45] This also holds for a later work, in which Born discussed the energy-momentum tensor in Mie's theory of matter (Born, 1914a).

[46] Ludwig Föppl was born on 27 February 1887 in Leipzig. He became *Privatdozent* at the University of Würzburg in 1914 and professor at the Technical University of Dresden in 1920. In 1922 he moved to Munich to succeed his father, August Föppl, in the chair of technical mechanics at the Technical University. He retired from the latter position in 1954, and died in Munich on 23 May 1976.

[47] To see the similarity, one has to compare the determinant on the left-hand side of Eq. (51) and the determinant of the system of inhomogeneous equations for the unknowns b_1/b_0, b_2/b_0, etc., which is obtained by dividing each of Hill's Eqs. (29) by b_0. Evidently, this system possesses a nontrivial solution only if the determinant of its coefficients has the value zero. Hence an equation similar in structure to Eq. (51) has, in Hill's theory of the moon, infinitely many coefficients.

and the α_i $(i = 1, 2, \ldots, n)$ are the roots of the equation

$$x^n - 1 = 0. \tag{52b}$$

Ludwig Föppl, in investigating the stability of atoms, regarded the atom as a mechanical system consisting of n electrons; hence he had to consider the action of $n - 1$ electrons on one electron. George Hill, on the other hand, had taken into account the action of infinitely many bodies on the moon. The actual physical situation lies between these extremes. Within an atom, as on the moon, only a finite (though large) number of bodies exert sensible forces. The problems of Föppl and Hill are even related more closely if one reminds oneself of the fact that every atom within matter is always surrounded by many other atoms and, depending upon the physical problem under consideration, their interactions have to be taken into account. The mathematical problems involved in such calculations of systems consisting of many atoms can often be handled only after replacing the finite system by an infinite system. Again it was Max Born in Göttingen who made the first application of infinite matrices to atomic problems.

Born had been acquainted with matrix methods ever since he had learned about them from his teacher Jakob Rosanes in Breslau in the beginning of the century. Rosanes was an expert in algebraic geometry and invariant theory, and he inspired many students, among them Born's friends Ernst Hellinger and Otto Toeplitz and, later, Richard Courant.[48] These students, especially the first two, did much to propagate the knowledge of matrices. Born recalled many years later:

> As a student in Breslau, I was much under the influence of Toeplitz, who was my senior by one year, and though my interest in algebra was not great he insisted on my learning matrix calculus properly and occasionally refreshed my knowledge when we were together again as young teachers in Göttingen. This turned out to be an advantage to me, first in developing Minkowski's form of relativity, then for the study of vibrations in crystals, which are determined, indeed, by a quadratic form of a (practically) infinite number of variables, but decidedly at the birth of quantum mechanics. (Born, 1940, p. 617)

Toeplitz' help proved to be especially valuable in the theory of crystal lattices, which Born and Theodore von Kármán developed in 1912.

[48] Jakob Rosanes was born on 16 August 1842 in Brody, Austria-Hungary (now in the Ukrainian S.S.R.) and studied at the Universities of Berlin and Breslau (1860–1865), obtaining his doctorate at the latter. Rosanes remained in Breslau for the rest of his life: he became *Privatdozent* in 1870, *Extraordinarius* in 1873 and full professor in 1876. He served as rector of his university during the academic year 1903–1904, retired in 1911, and died on 6 January 1922.

Rosanes was a good teacher, but not an equally good lecturer. As Courant recalled later, it was difficult to follow his lectures, especially since 'he scribbled equations which his students never quite saw because as he wrote he hid them with his body and as he moved along he rubbed them out with his sponge' (Reid, 1976, p. 13). At times, however, he could deliver inspiring lectures, such as his rectorial address of 1903 on the characteristic features of nineteenth century mathematics (Rosanes, 1904).

In order to determine the proper vibrations of a crystal, Born and von Kármán proceeded in the following way. They first considered a linear chain of N atoms and took into account the harmonic interaction (whose strength is given by the force constant α) between two adjacent atoms only (Born and von Kármán, 1912). Hence u_n, the displacement of the nth atom (with mass m) around its equilibrium position, follows from the equation of motion,

$$m\frac{d^2u_n}{dt^2} = \alpha(u_{n+1} - u_n) - \alpha(u_n - u_{n-1}), \qquad n = 2, \ldots, N-1. \tag{53}$$

Born and von Kármán's task consisted in bringing Eq. (53) into the standard form

$$m\frac{d^2U_n}{dt^2} + \alpha p_n U_n = 0, \qquad n = 2, \ldots, N, \tag{54}$$

for U_n, the so-called 'normal' coordinates of the system, with the products $(\alpha p_n/m)$ denoting the squares of the frequencies of the harmonic motion. For this purpose they had to perform a canonical transformation. They recognized, however, that this transformation was equivalent to solving the system of linear homogeneous equations,

$$-u_{n+1} + u_n(2-p) - u_{n-1} = 0, \qquad n = 2, \ldots, N. \tag{55}$$

This system has a nontrivial solution only if its determinant is zero, that is,

$$\det\begin{vmatrix} \ldots & . & . & . & . & . & \ldots \\ \ldots & 0, & -1, & 2-p, & -1, & 0, & \ldots \\ \ldots & 0, & 0, & -1, & 2-p, & -1, & \ldots \\ \ldots & 0, & 0, & 0, & -1, & 2-p, & \ldots \\ \ldots & . & . & . & . & . & \ldots \end{vmatrix} = 0. \tag{56}$$

The left-hand side yields a polynomial (of order $N-2$) for the unknown p, whose roots may be identified with the p_n in Eqs. (54).

The atoms at the boundary of the chain do not satisfy an equation of motion of the type of Eq. (53). In order to simplify their treatment, Born and von Kármán considered the limit of an infinite chain and imposed the boundary condition $u_{n+N} = u_N$. They expected this limit to yield a reasonably good physical approximation to their problem. Moreover, it did simplify the mathematical treatment considerably, as it allowed one to make use of the results obtained in the theory of infinite quadratic forms. Let us note that an infinite matrix of the type occurring in Eq. (56) is a special case of the type of matrix which had already occurred in Hill's theory of the moon. More importantly, however, Toeplitz had just recently studied in some detail quadratic forms, whose

matrices possessed the same structure (Toeplitz, 1907b, 1911). He wrote these forms as

$$L = \sum_{p,q=-\infty}^{+\infty} c_{p-q} x_p x_q, \tag{57}$$

with real or complex variables x_p and the coefficients depending just on the differences of the indices.[49] In case that the c_{p-q} form a symmetric or Hermitian matrix, he had shown that the spectrum had to be real (Toeplitz, 1911, p. 361, Theorem 6). As a result the roots of the infinite determinantal equation (56) had to be real as well. Born and von Kármán calculated them by making use of the factorization property, Eq. (52), and the periodicity condition (which implies the equality of the roots p_n and p_{n+N}), obtaining the result

$$p_n = 4\sin^2\left(\frac{\pi n}{N}\right), \qquad n = 0, 1, 2, \ldots . \tag{58}$$

The frequency spectrum of the infinite linear chain then followed with the help of the equation

$$\nu = \left(\frac{\alpha}{m} p_n\right)^{1/2}. \tag{59}$$

Although Born and von Kármán could derive their result, Eq. (58), by simply substituting in Eq. (53) the periodicity assumption, $u_n = u\exp(in\phi)$, where u is a constant amplitude and ϕ a constant phase which eventually equals $2\pi n/N$, only Toeplitz' theory of L-forms justified the steps in the derivation. It further guaranteed that the spectrum of a finite chain of N atoms was accounted for, in an acceptable approximation, by Eq. (58), even if no periodicity condition (which turns the finite chain into an infinite one) was assumed.

The discussion of the linear atomic chain or one-dimensional lattice served only as a preparation for the treatment of the three-dimensional lattice. With the latter, Born and von Kármán wanted to describe a realistic crystal and its thermodynamic properties. They had discussed it already in their first paper (Born and von Kármán, 1912) and returned to it in detail later (Born and von Kármán, 1913b). Instead of starting from the equations of motion, as we have outlined above in connection with the linear chain, they now preferred another approach: namely, they considered right away the potential energy of a three-dimensional lattice of mass points or atoms, whose coordinates (expressing the displacements from the equilibrium positions) are labelled by the indices l, m,

[49] Originally, Toeplitz (1907b) had used the L-forms to study the properties of Laurent series for functions of complex variables, i.e.,

$$f(z) = \sum_{n=-\infty}^{+\infty} c_n z^n,$$

by using c_n, the coefficients of the latter, as the coefficients c_{p-q} in Eq. (57).

and n. This means that the atoms are arranged (in equilibrium) at regular distances determined by multiples of the lattice constants, and the values of these multiples become the subscripts of the position coordinates. The latter are therefore written as $u_{l,m,n}$ (the coordinates in the x-direction), $v_{l,m,n}$ (the coordinates in the y-direction) and $w_{l,m,n}$ (the coordinates in the z-direction), where l, m and n assume integral values. V, the potential energy of a lattice consisting of atoms of the same mass, may be written as a quadratic form of the position coordinates, i.e.,

$$\begin{aligned} V = \frac{1}{2} \sum_{l,m,n} \sum_{\lambda,\mu,\nu} \Big[& u_{l,m,n} \left(a^{(1)}_{l-\lambda,m-\mu,n-\nu} u_{\lambda,\mu,\nu} + b^{(3)}_{l-\lambda,m-\mu,n-\nu} v_{\lambda,\mu,\nu} + b^{(2)}_{l-\lambda,m-\mu,n-\nu} w_{\lambda,\mu,\nu} \right) \\ & + v_{l,m,n} \left(b^{(3)}_{l-\lambda,m-\mu,n-\nu} u_{\lambda,\mu,\nu} + a^{(2)}_{l-\lambda,m-\mu,n-\nu} v_{\lambda,\mu,\nu} + b^{(1)}_{l-\lambda,m-\mu,n-\nu} w_{\lambda,\mu,\nu} \right) \\ & + w_{l,m,n} \left(b^{(2)}_{l-\lambda,m-\mu,n-\nu} u_{\lambda,\mu,\nu} + b^{(1)}_{l-\lambda,m-\mu,n-\nu} v_{\lambda,\mu,\nu} + a^{(3)}_{l-\lambda,m-\mu,n-\nu} w_{\lambda,\mu,\nu} \right) \Big]. \end{aligned} \tag{60}$$

Born and von Kármán had assumed that all other atoms (with position coordinates $u_{\lambda,\mu,\nu}$, $u_{\lambda,\mu,\nu}$, and $w_{\lambda,\mu,\nu}$) exert forces on the atom at the lattice point (l, m, n), as against the assumption of nearest-neighbour interaction that had led them previously to the equation of motion (53) for the one-dimensional chain. The six series of coefficients, $a^{(i)}_{l-\lambda,m-\mu,n-\nu}$ and $b^{(i)}_{l-\lambda,m-\mu,n-\nu}$ $(i = 1, 2, 3)$, now replaced the single force constant α of the former system.

Having set up the expression for the potential energy of the three-dimensional lattice, Born and von Kármán were confronted with the task of bringing it into the special, standard form, where it becomes reduced to a sum involving the squares of the transformed, i.e., so-called normal coordinates, $u'_{l,m,n}$, $v'_{l,m,n}$ and $w'_{l,m,n}$, i.e.,

$$V = -\frac{1}{2} \sum_{l,m,n} \left[p^{(1)}_{l,m,n} (u'_{l,m,n})^2 + p^{(2)}_{l,m,n} (v'_{l,m,n})^2 + p^{(3)}_{l,m,n} (w'_{l,m,n})^2 \right]. \tag{61}$$

The $p^{(i)}_{l,m,n}$, $i = 1, 2, 3$, denote the eigenvalues of the quadratic form (60)—similar to the $(\alpha\ p_n)$ in the linear chain—and they are related to the frequency spectrum of the crystal through the equation

$$\nu^{(i)} = \left(\frac{p^{(i)}}{m} \right)^{1/2}, \qquad i = 1, 2, 3, \tag{62}$$

in which the indices have been dropped for the convenience of writing. Equation (62) is the generalization of Eq. (59) for the case of the three-dimensional lattice. Born and von Kármán performed the principal-axes transformation of V, Eq. (60) to Eq. (61), by assuming that the coordinates, $u_{l,m,n}$, $v_{l,m,n}$, and $w_{l,m,n}$, are periodic functions, all those belonging to a given lattice point (l, m, n) depending on the same phase angle $\omega_{l,m,n}$. The problem of the principal-axes transformation

of the quadratic form, given by Eq. (60), was then extremely simplified: it was reduced to the separate transformations to principal axes of those quadratic forms which are connected with single atoms (each of which has three degrees of freedom associated with the translational motions along the crystal axes). As a consequence of this simplification Born and von Kármán had to solve—instead of the large system of equations that arises in the process of transforming to principal axes the quadratic form describing the potential energy of the total crystal—a system of three homogeneous linear equations in three variables for each atom in the lattice, that is,

$$\begin{aligned} u(A_1 - p) + vB_3 + wB_2 &= 0, \\ uB_2 + v(A_2 - p) + wB_1 &= 0, \\ uB_2 + vB_1 + w(A_3 - p) &= 0. \end{aligned} \tag{63}$$

The variables u, v and w are the amplitudes of the harmonic coordinates $u_{l,m,n}$, $v_{l,m,n}$ and $w_{l,m,n}$; for example, u is the amplitude of the function $u\cos\omega_{l,m,n}$, which is assumed to represent the coordinate $u_{l,m,n}$, etc. The coefficients A_k and B_k $(k = 1, 2, 3)$ correspond to the sums

$$A_k = \sum_{\lambda,\mu,\nu} a^{(k)}_{\lambda,\mu,\nu} \exp(i\omega_{\lambda,\mu,\nu}), \tag{63a}$$

$$B_k = \sum_{\lambda,\mu,\nu} b^{(k)}_{\lambda,\mu,\nu} \exp(i\omega_{\lambda,\mu,\nu}), \tag{63b}$$

involving the coefficients which determine the elastic forces between the atoms in the lattice. Finally, the summations over λ, μ, and ν extend to all lattice points in the crystal. The system of equations, (63), yields a nontrivial solution for u, v and w, provided the determinant of the coefficients is zero. This condition yields three real values for p, namely, $p^{(1)}$, $p^{(2)}$ and $p^{(3)}$, and these quantities also occur as the factors multiplying the normal coordinates on the right-hand side of Eq. (61). Evidently, in order to carry out the principal-axis transformation of V, one has to solve as many systems of three linear equations as there are atoms in the lattice. However, once all $p^{(i)}$ have been determined, one also arrives at the frequency spectrum of the crystal by applying Eq. (62). This frequency spectrum depends, as in the case of the one-dimensional lattice, on the force parameters, $a^{(k)}_{l,m,n}$ and $b^{(k)}_{l,m,n}$, and on the phase angles, $\omega_{l,m,n}$.

The question arises whether Born and von Kármán, by using a clever trick, avoided the problem of treating a very large or—in the approximation they used—even the infinite three-dimensional lattice. The answer is no; the reason being that in performing the principal-axis transformation of the quadratic form, given by Eq. (60), one obtains a characteristic equation involving a determinant, which is 'cyclic' for each direction of the crystal—i.e., it possesses the form of the determinant of Eq. (56) in three dimensions, except that in each row all coefficients are supposed to be different from zero. In general, such a determi-

nant allows one to make the assumption of periodicity. Since there are three directions in a crystal lattice, the phase angles, $\omega_{l,m,n}$, will be composed of ϕ, ψ and χ, the phase angles associated with the individual directions, i.e.,

$$\omega_{l,m,n} = l\phi + m\psi + n\chi. \tag{64}$$

Now, instead of one phase, i.e., $2\pi/N$, as in the case of the linear chain, in the spatial lattice three phases—$\phi = 2\pi/N_1$, $\psi = 2\pi/N_2$ and $\chi = 2\pi/N_3$, where N_1, N_2 and N_3 are the numbers of atoms (or lattice points) along the three lattice directions—play a role.

The use of infinite quadratic forms—especially the compact L-forms of Toeplitz—in the lattice theory of solids by Born and von Kármán remained perhaps the only significant application of infinite matrices in physics until 1925. In the years following 1912 Born had continued to develop his theory of the solid state, going beyond the problems he had treated first with von Kármán (i.e., the determination of the proper vibrations in a crystal in order to calculate, say, its specific heat). In the course of this work he occasionally referred again to the methods of infinite matrices. For example, he employed L-forms in an investigation of the electrostatic potential of a crystal lattice (Born, 1921). Hence Born was perhaps the only scientist who, ever since he had learned about matrices, was aware of their potential use in physics. It is therefore rather surprising why Born had never applied his favourite algebraic method to problems of quantum theory during the early 1920s. One reason for this neglect lay in Born's scientific temperament. Whenever he pursued a new subject, as was the case with the theory of atomic structure after 1922, he would first stick to conventional approaches and exploit them systematically before turning to a new method. Now in atomic theory there had been applied for many years the methods of Hamilton and Jacobi, which seemed to be particularly adequate to describe the dynamics of multiply periodic systems. Only after these methods were found to be wanting did Born turn to other techniques. However, the new mechanical description of atomic systems, which made use of the difference calculus, still remained very close to the previous one. It was with this scheme, which Born already called 'quantum mechanics' in spring 1924, that he had hoped to make progress. At any rate, in summer 1925 he had not exhausted its possible applications by far.

However, the main reason for the absence of matrix methods in the earlier quantum theory has to be found in two facts. First, it was not known until Heisenberg's discovery that in the description of atoms quantities, which do not commute, play a crucial role; second, the fact that matrices are noncommuting mathematical objects had not appeared openly in any of Born's previous applications. Thus, Heisenberg, when he hit upon the noncommutativity property of his quantum-theoretical quantities, did not know that there existed sophisticated algebraic methods, which in fact had been perfected by Hilbert and his collaborators in Göttingen, and which could be adapted to handle his problems. Born, in turn, had mostly used matrices with very special properties—such as the ones occurring in Toeplitz' L-forms—for physical application in his lattice

theory. Therefore, when he first read Heisenberg's paper on the re-interpretation of mechanical relations in quantum theory, he reminded himself of the matrices in Rosanes' lectures in Breslau rather than of the infinite quadratic forms of Hilbert's theory of linear integral equations. After this recognition, however, he quickly established the connection between the two; and in doing so he brought into play all his mathematical prowess, his thorough knowledge of infinite matrices, and his experience of the principal-axes transformations of infinite quadratic forms.

To an outsider the progress achieved in the formulation of matrix mechanics by Born and his collaborators from July to October 1925 appears like a marvel. For Born, however, all the steps which he took were an almost straightforward application of the mathematical subtleties with which he had become familiar during the previous decade and a half. That a theoretical physicist like Born, having just the appropriate mathematical formalism at his command, was around at the right place in July 1925 may not have been an entirely fortuitous accident of history. But the fact that Born possessed this knowledge explains how Heisenberg's intuitive quantum-mechanical scheme could be transformed into a full-fledged matrix mechanics within a short three months, when everybody—including Heisenberg—expected years of hard intellectual labour before a rational atomic mechanics would arise.

In spite of his mathematical preparation in matrix theory, Born did not enter into its application to quantum mechanics alone. He knew about the complexity of the calculations involved in such things as the principal-axis transformation of an infinite matrix from his own work in solid state theory, and he also knew that the application of matrix methods to quantum theory would require much more than just calculating the spectrum of eigenvalues. Therefore, he desperately wanted someone with experience both in quantum theory and mathematics to assist him in formulating matrix mechanics. Wolfgang Pauli, whom he met in Hanover on 19 July 1925, seemed to be just the right man for this enterprise, but Pauli refused to collaborate. Born then turned to his second choice, Pascual Jordan, who, though less experienced than Pauli, was interested in mathematical formalism and was willing to study in particular the theory of infinite matrices and its application to quantum mechanics.

I.4 Born's New Collaborator: Pascual Jordan

Pascual Jordan was born on 18 October 1902 in Hanover, Germany, as the son of the painter Ernst Pascual Jordan and his wife Eva, *née* Fischer.[50] The fact that both of his parents were interested in natural sciences influenced his intellectual development. Jordan recalled later: 'My mother knew a good deal about plants, flowers, birds and stars. I learnt from her that light takes eight minutes in coming

[50] Jordan's Christian name was, as he told Born, inherited from his great-grandfather, a Spaniard, who had come to Germany as a soldier in Napoleon's army and settled there for good (Born, 1978, p. 216).

from the sun to the earth. She was also interested in numbers and calculation; I learnt arithmetic from her' (Jordan, AHQP Interview, 17 June 1963, p. 1). His father used to paint landscapes, buildings and portraits and knew a lot about perspective; young Pascual learnt much from him about geometry even before he began to attend school. 'My father used to read popular books about natural science,' he recalled, 'especially the series of *Kosmos* books which was popular at the beginning of this century. Each year several books came out in this series, written by people such as [Wilhelm] Bölsche who was well known as a popular author. All these *Kosmos* books belonged to my first reading in natural science' (Jordan, AHQP Interview, 17 June 1963, p. 1).[51]

At the age of ten or twelve Jordan thought of becoming a painter or architect, and it was only several years later that he decided to study natural sciences. At first he was attracted by biology. As he recalled later: 'To be sure, already very early as a small boy I liked animals very much. I often visited the zoological garden, and then I also possessed a collection of pictures of extinct animals like saurians, etc.' (Jordan, AHQP Interview, 17 June 1963, p. 4). With the passage of time, however, his interests shifted to physics and mathematics. Jordan went to the *Reform Gymnasium* in Hanover, a school which laid emphasis on scientific studies. There, for instance, in mathematics he learned differentiation and binomial series; in physics he learned mechanics and other subjects in an elementary manner.

In science, the larger view of all things attracted him. 'Rather early,' he remembered later, 'that is, at the age of fourteen, I had conceived a plan of once writing a big book about all fields of science' (Jordan, AHQP Interview, 17 June 1963, p. 5). He liked fundamental philosophical questions, and on these he would try to find books to study. First he turned to biology: A. Pauly's book on the theories of Charles Darwin and Jean Baptiste Lamarck on the propagation of species (*Darwinismus und Lamarckismus*, 1905) had come into his hands. Jordan was particularly interested in the discussions, which were going on in those days, about neo-vitalism and neo-Lamarckism and the question whether organic life could be understood by purely mechanical or physical means. At the age of fourteen he also read Friedrich Albert Lange's *Geschichte des Materialismus* (*History of Materialism*) and was deeply impressed by it.[52] At about this time he

[51] *Kosmos*—more accurately, *Kosmos, Gesellschaft der Naturfreunde*—was the name of a German association for natural science. From 1904 this association published a journal called *Kosmos*. (The publication was interrupted in 1944, but renewed in 1946.) The journal was addressed mainly to young and lay readers. To supplement the articles on specific topics, a series of small books dealing with certain fields of biology, astronomy, physics and the development of civilization was added. W. Bölsche (1861–1939), a popular writer in Germany, who propagated the ideas of Darwin and Haeckel on evolution in biology—and supported Haeckel's philosophical scheme, the so-called monism ('*Monismus*')—wrote the first volume, entitled '*Abstammung des Menschen*' (1904) and contributed numerous others.

[52] F. A. Lange (1828–1875), philosopher and writer, edited newspapers in Germany and Switzerland before becoming a professor of philosophy in Zurich (1870) and later in Marburg (1872). As a philosopher he belonged to the founders of the so-called 'neo-Kantianism,' i.e., a renewed and altered version of Kantian philosophy; as a political writer he propagated early socialistic demands. He also treated topics of the history of philosophy. His book on the history of materialism, first published in 1866, went through many editions, and was reprinted as late as 1926.

obtained some insight into classical physics; in particular, he understood the principle of action and reaction in mechanics and how a rocket could be propelled in space. 'From this I concluded that it must be possible to reach the moon,' he recalled. 'And this thought greatly stimulated my scientific interests' (Jordan, AHQP Interview, 17 June 1963, p. 1). In the last years of high school he was intrigued by the question about how matter became converted into different forms of energy, and he thought of learning all about atomic physics.

During the years at the *Reform Gymnasium*, Jordan widened his knowledge considerably by reading. Thus, for instance, he studied Konrad Knopp's small books, *Funktionentheorie 1* and *2*, in the *Göschen* series of volumes that had come out in 1913, and a textbook on Maxwell's electrodynamics. He obtained a lot of information from a book of Walther Nernst and Arthur Schoenfliess, entitled '*Kurzgefaßtes Lehrbuch der Differential- und Integralrechnung*' (1895), which he found accidentally in an *Antiquariat* when he was about sixteen years old. He was attracted by the book's title and looked curiously into its contents without understanding much in the beginning. Jordan bought other books, too, but the Nernst–Schoenfliess treatise proved to be especially valuable. He remarked about it almost forty-five years later: 'It presents (differential) calculus and the beginnings of integral calculus, also some considerations about multiple integrals and such things, with specific application to scientific problems . . . I could not read it all at once; at first I read only the beginning chapters, then stopped, and later went on a bit further. Thus it occupied me really for several years until the early period of my [university] studies' (Jordan, AHQP Interview, 17 June 1963, p. 3). Because of his philosophical interests he soon began to study the works of Ernst Mach, especially his *Mechanik* (1883) and *Prinzipien der Wärmelehre* (1896); later on he also read *Prinzipien der Physikalischen Optik* (1921) and the *Beiträge zur Analyse der Empfindungen* (1886). Jordan became a lifelong devotee of Mach's philosophy, which he found 'particularly helpful for the understanding of relativity and quantum theory.' He learned about relativity theory in 1919 from Moritz Schlick's book, *Raum und Zeit in der gegenwärtigen Physik* (1917), which he found to be a 'short, popular, but very nice book.'[53] He had thus acquired a broad intellectual background when he graduated from high school in spring 1921.

Pascual Jordan entered the *Technische Hochschule*, Hanover, in the summer

[53] Moritz Schlick was born in Berlin on 14 April 1882. He studied physics at Heidelberg, Lausanne and Berlin, obtaining his doctorate under the supervision of Max Planck in 1904. He then changed to philosophy, becoming a *Privatdozent* at the University of Rostock in 1911 and professor in 1917. Via Kiel (professor in 1921), Schlick finally moved to Vienna (in 1922) as professor of the philosophy of inductive sciences, his predecessors being Ernst Mach and Ludwig Boltzmann. In Vienna he gathered a group of philosophers, such as Rudolf Carnap, Otto Neurath and Friedrich Waismann, and of mathematicians and scientists, such as Kurt Gödel, Philipp Frank, Karl Menger and Hans Hahn. Thus, he became the founder of the Vienna circle of neo-positivism, i.e., a philosophical school of empiricism, which exerted much influence on philosophers in other countries and helped to create the international movement of logical positivism. Schlick was shot to death by a demented student on 22 June 1936 as he entered the University of Vienna to give his lecture.

semester of 1921 with the intention of studying physics.[54] He soon found that the physics lectures were rather bad, given as they were by 'a very undistinguished physicist' (Jordan, AHQP Interview, 17 June 1963, p. 6).[55] Hence he did not attend them, and turned to mathematics, taking a course on ordinary differential equations by Conrad Müller and another on algebra by Georg Prange.[56] His education was rather unsystematic: the course on differential equations belonged to the third semester, and Jordan took a course on integration methods (which should have preceded the course on differential equations) in the winter semester of 1921–1922. Besides mathematics he attended Wilhelm Friedrich Kohlrausch's lectures on *Elektrotechnik* ('Electrical Engineering') and Max Bodenstein's lectures on physical chemistry.[57] The only physics course he attended in Hanover was a specialized one on radioactivity.

Since physics at Hanover's *Technische Hochschule* did not satisfy Jordan, he went to study at the University of Göttingen in summer 1922. As he recalled: 'I had carried with me a letter of recommendation from one of my former teachers [at the *Reform Gymnasium*]—not a physics teacher, but a teacher of Latin and German—addressed to Mr. Pohl' (Jordan, AHQP Interview, 17 June 1963, p. 6). Robert Wichard Pohl was the director of one of the two experimental institutes in Göttingen, and he gave his lectures from 7 to 8 or 8 to 9 in the morning, an hour which was just too early for Jordan. 'Thus I became a physicist,' he claimed, 'who never attended a [regular undergraduate] course of lectures on physics' (Jordan, AHQP Interview, 17 June 1963, p.6).[58] Again he turned to mathematics, enrolling at once in Richard Courant's course on partial differential equations. He showed a keen interest in mathematics, and even volunteered to work out the lectures in detail.[59] He recalled later how he got into it: 'When I presented myself to Courant, I showed him what I had worked out from his first lecture. He liked that, he found it nice, and things got started this way. It was a

[54] Jordan began to study in his hometown for two reasons. First, he could stay with his parents. Second, his father still hoped that he might give up sciences and mathematics and eventually turn to architecture, which he could also pursue at the *Technische Hochschule*. When Pascual finally decided to stay on with physics, his mother helped to persuade his father to allow him to continue in that field.

[55] The professor of physics in question was Carl Christian Julius Precht (1871–1942), who had taught in Hanover since 1906.

[56] Conrad Heinrich Müller (1878–1953) served as professor of mathematics at the *Technische Hochschule*, Hanover, from 1910 to 1948. Heinrich Friedrich Wilhelm Georg Prange (1885–1941) was his colleague from 1921 to 1941. They collaborated in writing a textbook on mechanics, entitled '*Allgemeine Mechanik*,' which appeared in 1923.

[57] W. F. Kohlrausch (1855–1936) was professor of *Elektrotechnik* at the *Technische Hochschule* in Hanover until 1923.

Ernst August Max Bodenstein (1871–1942) was professor of electrochemistry at the *Technische Hochschule*, Hanover, from 1908–1923. Then he was called to the chair of physical chemistry at the University of Berlin, from which he retired in 1936.

[58] Jordan shared with Wolfgang Pauli the habit of going to bed late and not rising early in the morning.

[59] These worked out lecture notes were then kept in the Mathematical Reading Room for the use of students.

bit difficult, for I had actually not attended the first lecture. Since I had to enroll at the university, I was prevented from doing so; hence I had to learn from a fellow student what he had treated. According to this report I worked out [the details of] the lecture' (Jordan, AHQP Interview, 17 June 1963, p. 12). In any case, he soon became familiar with the Mathematical Institute of Göttingen. Emil Artin, who was then Courant's assistant, advised Jordan to take pure mathematics rather than physics as his principal subject, and for a while he actually thought of becoming a pure mathematician.

Through Pohl, Jordan met Max Born. By the time of Jordan's arrival in Göttingen, Born's principal scientific interest had turned to a systematic examination of the problems of atomic mechanics, though he was still engaged upon the completion of his article on the dynamics of crystal lattices for the *Encyklopädie* (Born, 1923b). Jordan even gave him some technical help, which Born acknowledged in the foreword.[60] Together with Born he soon began to concentrate on the study of atomic structure. The two of them would often work together and discuss papers of common interest with each other. For Jordan, who stuttered, this friendly personal contact with Born was very helpful. And this, together with his interest in atomic physics, determined his future scientific career.

Already after finishing the *Gymnasium* Jordan had undertaken the study of Sommerfeld's *Atombau und Spektrallinien*, which introduced him to the excitement of atomic physics. Now in Göttingen in June 1922, during his first semester, he obtained the opportunity of attending Bohr's Wolfskehl lectures on atomic theory and of listening to the ensuing discussions, which involved the most outstanding experts in the field, including Arnold Sommerfeld and his brilliant students. Thus, during the *Bohr Festival*, he encountered Werner Heisenberg and Wolfgang Pauli, with whom he became better acquainted later. In the following winter semester, 1922–1923, Jordan attended Born's Seminar, where perturbation theory was discussed in detail, and together with Heisenberg and Friedrich Hund he studied the methods of Henri Poincaré and Karl Schwarzschild. He found this work satisfying, and it confirmed his decision to pursue physics: 'This thing is so interesting, I must stay in physics' (Jordan, AHQP Interview, 18 June 1963, p. 4).[61]

Before going to Göttingen Jordan had acquired knowledge of classical mechanics by studying the relevant portions of Hermann von Helmholtz' *Vorlesun-*

[60] In the foreword to the encyclopaedia article Born thanked 'especially P. Jordan,' whom he owed 'many valuable hints' (Born, 1923b, p. VI). Jordan, on the other hand, remembered that his help consisted mainly of inserting the formulae in the manuscript.

[61] Jordan decided in favour of physics for several reasons. One was certainly his excitement for atomic physics, for as he recalled: 'I still remember exactly that a paper of Bohr then appeared in *Zeitschrift für Physik*. *Zeitschrift für Physik* was displayed in the Mathematical Reading Room ("*Lesezimmer*"), and there I saw this paper, and I felt so fascinated that I said [to myself]: "No, it is certainly not right that I should go on with mathematics"' (Jordan, AHQP Interview, 18 June 1963, p. 4). This happened before the Bohr lectures in Göttingen.

gen über theoretische Physik ('*Lectures on Theoretical Physics*,' 1897–1907).[62] He now attended Born's lectures on thermodynamics and electrodynamics, and took Hilbert's course on statistical methods. In fact, he attended Hilbert's lectures on physical topics whenever they were given.[63] In spite of his earlier interest in astronomy, he did not attend lectures in it. As he recalled:

> The astronomer was the old [Johannes] Hartmann.[64] He had once discovered that the nova stars explode . . . . Otherwise he was not very modern. How much the new astronomy was stimulated by atomic physics, that we realized only from the title of Eddington's book *Stars and Atoms* [Eddington, 1927]. But even that book I saw only superficially, I did not study it. And otherwise there was no close contact between physicists and astronomers, for Hartmann was not—as I said—very modern. At that time one could not yet notice in Göttingen much of the close collaboration between modern spectroscopy and astrophysics. (Jordan, AHQP Interview, 18 June 1963, pp. 5–6)

Though enthusiastic about theoretical physics, Jordan did not find the courses in experimental physics very inviting. Not only did he skip Pohl's undergraduate lectures, he was also not fond of James Franck's courses on such topics as the conduction of electricity through gases, the experimental aspects of atomic physics, etc. Moreover, the experimental work required in Franck's physics laboratory did not interest him; he failed and was required to repeat it. This was due to his dislike of any experimental work. On the other hand, he enjoyed a laboratory course in general zoology. In Göttingen he returned to his old favourite, zoology, derived from his love of animals as a child and his later reading of the *Kosmos* books. He now eagerly sat in Alfred Kühn's lectures on heredity.[65] He even chose zoology as the second minor (*Nebenfach*) for his

[62] Hermann von Helmholtz' *Vorlesungen über theoretische Physik* (*Lectures on Theoretical Physics*, 1897–1907) included: I. an introductory volume on the basic ideas and foundations of physics; II. the dynamics of mass points; III. mechanics of continua; IV. acoustics; V. electrodynamics and magnetism, and the electromagnetic theory of light; VI. the theory of heat. Besides Vol. II Jordan certainly read Vol. I on the foundations of physical theory because of his interest in philosophical questions.

As we have mentioned, Jordan bought many books on science for private study. Thus, he also possessed Max Planck's *Vorlesungen über die Theorie der Wärmestrahlung* (Planck, 1906).

[63] According to the official announcements, Hilbert lectured during the summer semester on '*Statistische Methoden insbesondere in der Physik*' ('Statistical Methods, Especially in Physics,' *Phys. Zs.* **23**, 1922, p. 165), and in the following winter semester 1922–1923 on '*Mathematische Grundlagen der Quantentheorie*' ('Mathematical Foundations of Quantum Theory,' *Phys. Zs.* **23**, 1922, p. 532).

[64] Johannes Franz Hartmann (1865–1936) studied at the Universities of Tübingen, Berlin and Leipzig. He received his doctorate from Leipzig in 1891; then he served as an assistant in Vienna, Leipzig and Potsdam and became a professor of astronomy and director of the Göttingen Observatory in 1909. Hartmann was a leading astrophysicist of his day and an authority on spectroscopy around 1900.

[65] Alfred Kühn was born on 22 April 1885 in Baden–Baden. Before becoming professor of zoology and director of the Zoological Institute of the University of Göttingen in 1920, he served as a lecturer at the Universities of Freiburg and Berlin. In 1937 he was appointed director of the *Kaiser-Wilhelm-Institut für Biologie* in Berlin, which was transferred to Tübingen and renamed *Max-Planck-Institut für Biologie* in 1945. Kühn retired from his directorship in 1964.

doctorate. However, besides theoretical physics, he preferred mathematics most. Jordan took several courses from Courant, and he gave technical assistance during the latter's writing of the *Methoden der mathematischen Physik* (Courant and Hilbert, 1924). Thus he used to meet Courant almost daily to discuss mathematical problems.[66] Out of curiosity he went to Edmund Landau's lectures on number theory. He also attended some of Niels Erik Nörlund's lectures on difference equations, which he later found useful in his work on atomic theory.

Besides lectures and laboratory courses Jordan completed his education by attending the weekly Physics Colloquium as well as participating in the Born–Franck–Hilbert Seminar on the structure of matter ('*Struktur der Materie*'), which was held regularly each semester and dealt with atomic theory. Now and then he went to the meetings of the *Mathematische Gesellschaft*, where he made the acquaintance of Emmy Noether; after all, because of his work for Courant, he belonged partly to the mathematical circle.

After helping Born with his *Encyklopädie* article on crystal dynamics, Jordan became his closest student. As he recalled: 'There were also other students, or junior co-workers who had already obtained their doctorate, who spent much time together with Born for joint work in his apartment. I believe I was closer to him [than anyone else], for I had an extensive collaboration with him, first on the book on lattice theory of crystals and later on quantum-theoretical questions' (Jordan, AHQP Interview, 18 June 1963, p. 9). Besides Born, Jordan also discussed theoretical problems with other students, but these discussions were much less intense. Jordan used to report to Born about the progress of his theoretical work, and Born would propose new problems and give the necessary advice. On completing the *Encyklopädie* article Born, for a while, became deeply involved in the theory of molecules, a subject tied both to the theory of the solid state and to atomic theory. When Jordan was looking for a problem to treat in his dissertation, Born tried to interest him, too, in molecular theory. But Jordan did not warm up to the idea, for the physics involved was not to his liking. His preference was for more fundamental problems. Now, in 1923, there existed such a problem: in connection with the Compton effect, the problem of the light-quantum and how to fit it into quantum theory was much discussed. The quantum physicists, in particular Niels Bohr, had raised many objections against Einstein's old conception of light-quanta (Einstein, 1905b). On the other hand, Wolfgang Pauli treated the interaction between electrons and radiation consistently by taking into account the light-quantum properties of radiation (Pauli, 1923b). Then Albert Einstein and Paul Ehrenfest generalized Pauli's treatment in a further article (Einstein and Ehrenfest, 1923). That Einstein concerned himself again with the light-quantum problem set Jordan moving. Later on he reconstructed his reaction as follows: 'I believe, I became interested in this work [of Einstein and Ehrenfest] because I had also read Pauli's paper earlier and had now found this paper of Einstein and Ehrenfest, which I found to be very

[66] Jordan worked at least as much to help Courant on his book as he did for Born's review article on lattice dynamics. Courant thanked him in his foreword as one among a long list of 'true helpers.'

interesting. Then, I think, I announced on my own that I would like to give a talk on it sometime' (Jordan, AHQP Interview, 20 June 1963, p. 21). He thus presented the content of that paper in a seminar, with which Born was satisfied. As Jordan recalled: 'After listening to my seminar on the light-quanta he said that I could take this as my topic if I liked it better. In general one was skeptical [about the light-quanta] and one thought that the problem had to be solved differently, but somehow Born's opinion in this regard was not fixed and he thought that since I had raised the question I might as well pursue it' (Jordan, AHQP Interview, 17 June 1963, p. 15).[67] Jordan was glad to have escaped molecular physics, and he turned to the problems arising from the light-quantum hypothesis in atomic physics.

In his thesis Jordan sought to propose a compromise between the extreme positions of the light-quantum theory, as assumed by Einstein, and of the wave theory of radiation, as preferred by Niels Bohr on the basis of correspondence arguments (Jordan, 1924). In particular, he made an attempt to modify Einstein's theory of the interaction between electrons or atoms and radiation (Einstein, 1909a, 1916e) by avoiding an important assumption of the light-quantum theory: namely, the assumption that a light-quantum of energy $h\nu$ necessarily imparts a recoil momentum of magnitude $h\nu/c$ to the electron or atom, where h is Planck's constant, ν the frequency of radiation, and c the velocity of light. To that end, Jordan assumed that $\sigma(\theta, \phi)$, the distribution of the intensity of radiation scattered from an atom or electron, is not concentrated in one direction in space (denoted by the scattering angles θ and ϕ) but varies continuously, taking nonzero values for all θ and ϕ. As a result, in Jordan's theory, the momentum of the scattered radiation obtained a value much smaller than Einstein's value $h\nu/c$. And yet, Jordan was able to explain the conclusions derived from the Compton effect concerning the change in the frequency of the scattered radiation and the motion of the recoil electron. His thesis, therefore, sought to bridge the open conflict between Einstein's light-quantum theory and the Bohr–Kramers–Slater radiation theory by offering them a kind of coexistence.

Jordan's work, however, did not persuade his more advanced colleagues. Pauli, for instance, after reading the thesis, simply declared: 'Now I am more

[67] Born himself had not worked on the light-quantum problem before, though he followed the development with interest. On the one hand, his friend Albert Einstein had pioneered the light-quantum hypothesis; he had advocated it over many years against serious objections, and the discovery of the Compton effect seemed to prove that his conception was right. On the other, Niels Bohr—whose guidance Born had followed very closely in atomic theory—refused to believe in the reality of the light-quantum and developed (together with Hendrik Kramers and John Slater in 1924) the counterproposal of a statistical radiation theory. Other physicists were also concerned with the nature of radiation in the early 1920s. Thus Walther Bothe discussed the possibility of deciding the question about the reality of the light-quantum by studying the details of the photoelectric effect (Bothe, 1923c, 1924b). Furthermore he attempted to explain Einstein's fluctuation formula (Einstein, 1909a) and the interaction between radiation and free electrons by assuming the existence of multiple light-quanta having energies $sh\nu$, with $s = 1, 2, 3, \ldots$ (Bothe, 1923e, 1924a). Bothe's papers were discussed in Göttingen, especially in the private seminar for students that was arranged from time to time in Hertha Sponer's home (see Hund, AHQP Interview, 25 June 1963, p. 13), and they constituted some of the background of Jordan's interest in light-quanta.

convinced than ever that Einstein is right' (Jordan, AHQP Interview, 18 June 1963, p. 10). Einstein also studied Jordan's generalization and relaxation of the light-quantum theory and commented on it in a short note to *Zeitschrift für Physik*, where the thesis had been published (Jordan, 1924). He began by saying: 'In the ingenious investigation quoted above, Jordan has tried to disprove the thesis proposed by me, that for each elementary process of emission and absorption a momentum of magnitude $h\nu/c$ is transferred to the emitting or absorbing molecule' (Einstein, 1925c, p. 784). Then he pointed out that Jordan's use of a distribution function $\sigma(\theta, \phi)$, which is nonzero in all space directions θ and ϕ, implied that the emitted radiation depended on the direction of the incident radiation in a strange way. Thus, the probability that radiation, which impinges on the atom or molecule from opposite directions with intensities ρ_+ and ρ_-, is absorbed, becomes proportional to the square root of the product $\rho_+ \cdot \rho_-$, provided the two incident beams of radiation are not coherent. As a consequence, in the case of radiation coming from one direction—say, with zero intensity ρ_-—Jordan obtained no absorption at all, evidently in contrast to observation. In addition, Einstein noted that in Jordan's theory an atom or molecule would behave very strangely if exposed to a thermal radiation field and concluded:

> All this is connected with the fact that the author uses a hypothesis, according to which the amounts of radiation taken from rays of different directions are not treated as independent of each other. One must assume this, however, in order not to contradict the most elementary experiences of absorption. If one does so, then one necessarily arrives at the result that in every elementary process of absorption or emission a momentum of absolute magnitude $h\nu/c$ is transferred to the molecule. (Einstein, 1925c, p. 785)

In spite of this criticism Jordan was glad that Einstein took an interest in his work on radiation theory. He continued to work in this field, and some months later he published another paper on the thermal equilibrium between electrons and blackbody radiation (Jordan, 1925d). In it he did not persist in the ideas of his thesis, but followed Pauli's and Einstein and Ehrenfest's treatments of the processes involved. In particular, he showed that thermal equilibrium was still preserved if Planck's law governed the energy distribution in radiation and the free electrons or atoms obeyed Bose–Einstein statistics rather than Boltzmann statistics, as had been assumed previously by Pauli and Einstein and Ehrenfest. In conclusion Jordan thanked 'Mr. Einstein for the kind interest which he has taken in this investigation' (Jordan, 1925d, p. 655). He felt encouraged to continue to make further contributions to the fundamental problem of radiation.

In his student years at Göttingen, Jordan showed no particular eagerness to produce many publications, which was quite unlike Pauli and Heisenberg who were only slightly senior to him. This may have been due to his reticent character and the fact that he began to study physics seriously only in summer 1922. Furthermore, he quite early became involved in the laborious projects of other

people, such as Born's article for the *Encyklopädie* and Courant's *Methoden der mathematischen Physik,* from which he did not earn great merit or rewards. Afterwards, when he was ready to attack problems of atomic theory, he did not treat such problems as molecular structure, on which it seemed easier to write papers, but selected the difficult question of light-quanta for his research. So it happened that Jordan's first publication dealt with the results of his thesis. In this respect Jordan's scientific career resembled that of an average student, and yet he was considered to be a great talent by his professor. Thus, Born wrote to Einstein in summer 1925 about Jordan and his thesis:

> As regards your objections against Jordan's paper, I still feel very unsure of myself; but as I am now coming to grips with these things from my own somewhat complicated point of view, I will understand them one of these days. On the whole, you are certainly right; though Jordan's opinion is based on somewhat different considerations, as he allows coherent bundles of rays, whereas you only mention incoherent ones. Even if Jordan is mistaken in this, as I now think is highly probable, he is still exceptionally intelligent and astute and can think far more swiftly and confidently than I. (Born to Einstein, 15 July 1925)

He thought that Jordan was an able collaborator, who would make his way as a theoretical physicist.

It is surprising, however, that Born did not draw Jordan immediately into his research projects. After completing his doctorate in fall 1924, Jordan studied spectroscopic problems under James Franck. 'Franck engaged me to help him with the preparation of the book *Anregung von Quantensprüngen durch Stöße* [Franck and Jordan, 1926], and I learnt a great deal from this activity. I had to study a whole series of papers and take notes of their contents in order to collect the material. Later on we wrote it together' (Jordan, AHQP Interview, 18 June 1963, p. 7). Again he became involved in a laborious project, which took a long time. This time, however, he acted as a junior collaborator on equal terms, not just as an anonymous helper.[68] This work also led to certain publications. Jordan thus submitted his first paper on a spectroscopic problem already on 23 February 1925 (Jordan, 1925a), which was based on an experimental observation reported by Karl Meissner of Zurich University in the January issue of *Annalen der Physik* (Meissner, 1925a). While observing the absorption of light by excited neon

[68] In spite of the long time he had to invest in writing the book with Franck, Jordan later found the work rather useful. As he recalled: 'Well, it was also very important for me to obtain a feeling once for what the empirical facts look like . . . The experimental work on electron collisions and collisions of the second kind, which Franck pursued, about that I really got a clear idea first from working on this book' (Jordan, AHQP Interview, 18 June 1963). Franck agreed that Jordan learned a good deal in writing the book, though he admitted the limitations of the joint enterprise. He said: 'This book is by no means a good source of exact literature citations, because actually it was always this way: I dictated a great part to Jordan, and then Jordan went over it. I didn't look up the literature, so I ascribed a good deal of things to Mr. Miller which were done by Mr. Mayer, and what not. And Jordan thought that I would know, and I thought that he would look at it' (Franck, AHQP Interview, Fourth Session, p. 4).

atoms, Meissner noticed that there occurred two types of transitions, connected with strong and weak absorption, respectively; from this he concluded that the so-called $1s_3$ and $1s_5$ states of neon had a long lifetime, and the $1s_2$ and $1s_4$ states had a short lifetime. Born was interested in Meissner's investigation, which seemed to imply that the ground state of neon was a *p*-state, and he asked Jordan to examine the question on the basis of known theoretical principles.[69] In his short paper, Jordan tried to organize the terms according to the 'branching rule' of Landé and Heisenberg (1924), but he found that the regularity of the periodic system of elements required certain restrictions upon this rule. Wolfgang Pauli, his former critic, agreed with his result. As he recalled later: 'Pauli, who saw this manuscript [Jordan, 1925a], was very interested in my observation, as he had just then written his paper on the exclusion principle. His paper was ready in manuscript, but not yet published, and he noticed that my remark touched upon his subject' (Jordan, AHQP Interview, 17 June 1963, p. 15). Jordan, like Pauli, had used Stoner's classification of stationary states in atoms (Stoner, 1924) and concluded that the ground state of the neon atom must be a *p*-state (with the principal quantum number 2). He thereby predicted the existence of two strong lines, which should arise in transitions from the states $1s_2$ and $1s_4$ to this ground state, and they were discovered a couple of months later by Gustav Hertz experimentally (Hertz, 1925). Jordan was happy with the result of his excursion into spectroscopy and continued to work on similar questions. In July 1925 he published a further note on atomic structure (Jordan, 1925c), in which he attempted to clarify the relation between the spectroscopic terms of atoms and the configurations of the series electrons beyond what had been given previously by Pauli (1925b) or Heisenberg (1925b). Because of these contributions to spectroscopy, Born praised Jordan together with Heisenberg and Hund in a letter to Einstein: 'On the whole, my young people, Heisenberg, Jordan and Hund, are brilliant. I find that merely to keep up with their thoughts demands at times a considerable effort on my part. Their mastery of the so-called "term zoology" is marvellous' (Born to Einstein, 15 July 1925).

In spite of this success Jordan did not become a spectroscopist, but left these matters to his older colleague Friedrich Hund. Instead, Born engaged Jordan to collaborate with him on more fundamental aspects of atomic theory. In November 1924 Born had turned his attention to the problem of the width of spectral lines, for as he reported at the time: 'I am trying to obtain the finite width of the terms by associating with each state a finite lifetime and then applying perturbation theory to it via a Fourier integral' (Born to Heisenberg, 26 November 1924).

[69] Jordan later remarked about his first spectroscopic paper: 'I had previously [i.e., before the completion of the dissertation] intended to publish another small paper on a question posed by [Born], dealing with neon' (Jordan, AHQP Interview, 17 June 1963, p. 15). Evidently, Jordan had the dates slightly mixed up because the experimental work of Meissner was published in January 1925, months after Jordan had submitted his thesis for publication in *Zeitschrift für Physik*. On the other hand, Born's interest was documented in the acknowledgement of Jordan's paper. There he had stated: 'I am very much obliged to Professors Born and Franck for the friendly interest which they have shown towards this little work' (Jordan, 1925a, p. 881).

Since John H. Van Vleck had treated nondegenerate systems along similar lines (Van Vleck, 1924a, b), Born just concentrated, together with Lothar Nordheim, upon the degenerate case.[70] Now with Jordan he examined an even more general aspect of the theory of interaction between matter, represented by atomic systems, and radiation, denoted by an electromagnetic field which changes with time. Jordan liked the topic Born chose for their first collaboration, for it had to do with his favourite problem of radiation theory. 'Getting into the great trend of development of quantum physics I discussed with my venerable teacher Max Born an article published by Van Vleck, showing that Einstein's famous law of the relation between positive and negative absorption and spontaneous emission by quantized atoms may be understood as a consequence of Bohr's correspondence principle' (Jordan to Mehra, 14 December 1974). Indeed, the joint work with Born allowed him to participate in the forefront of endeavours directed towards the search for a new, consistent quantum theory. The method consisted in reformulating the classical laws describing the interaction between multiply periodic systems and external fields in a language which displayed explicitly the discrete nature of the physical phenomena. In his papers, Van Vleck had discussed the absorption and emission of electromagnetic radiation of a special kind; i.e., he had substituted periodic electric fields and averaged over the phases in order to represent radiation. Born and Jordan were now interested in extending Van Vleck's results to the case of an external field of arbitrary, especially nonperiodic structure, acting on a periodic system. Born explained to Einstein the reason for pursuing this investigation as follows:

> But my principal interest is the rather mysterious difference calculus on which the quantum theory of atomic structure seems to be based. With Jordan I have been examining systematically, though with the minimum of mental effort, every imaginable correspondence relationship between classical, multiply periodic systems and quantum atoms. A paper on this subject in which we examine the effect of nonperiodic fields on atoms will appear soon. This is a preliminary study for an investigation of the processes occurring in atomic collisions (quenching of fluorescence, sensitized fluorescence *à la* Franck, etc.); one can understand, I think the essential characteristics of what goes on. The different behaviour of atoms depends mainly on whether they have an (average) dipole moment, or just a quadrupole moment, or even higher electric symmetry still. (Born to Einstein, 15 July 1925)

In the paper on the quantum theory of aperiodic processes (Born and Jordan, 1925a), which was received by *Zeitschrift für Physik* on 11 June 1925, the authors proceeded in the standard manner. They started from a classical multiply periodic system, which is influenced by an aperiodic field, $\mathbf{E}(t)$, i.e., a field of

[70] Born wrote about this work again to Niels Bohr on 15 January 1925:

> Further, I am working hard on the radiation problem; a part of my considerations has meanwhile been anticipated by Van Vleck. Have you read these papers? They deal with a correspondence-like treatment of absorption [of radiation]. I have enjoyed this very much. Van Vleck deals with non-degenerate systems, while we (Nordheim and myself) seek to investigate precisely the influence of degeneracy. But I am afraid that the result will be the same as what Kramers and Heisenberg have done in the meantime.

arbitrary structure [in time] that exists only in a finite period of time, say from t_0 to t_1. The field is represented mathematically by a Fourier integral

$$\mathbf{E}(t) = \int_{-\infty}^{+\infty} \mathbf{e}(\omega) \exp(2\pi i \omega t)\, d\omega, \tag{65}$$

where $\mathbf{e}(\omega)$ is a vector denoting the Fourier amplitude associated with the frequency ω. The classical part of the problem consisted in applying perturbation theory; that is, one treated the external field as a perturbation of the original system and calculated the change of its properties in some approximation, usually keeping terms up to second order. For their joint work, Born contributed the basic physical ideas and relied on Jordan to carry out the details of the calculation and the mathematical proofs. Thus, Jordan, in a separate paper (Jordan, 1925b), investigated convergence questions; for example, he studied the first-order correction, S_1, to the action function, i.e., the expression

$$S_1 = \frac{i}{2\pi} \mathbf{A}_\tau \int_{-\infty}^{\infty} \frac{\mathbf{e}(\omega)}{\nu\tau + \omega} \exp\left[2\pi i (w^0\tau + \omega t)\right] d\omega, \tag{66}$$

where the vectors $\mathbf{A}_\tau$ denote the Fourier components (labelled τ, which assumes integral values) of the electric moment of the unperturbed system with angle variables w^0 and harmonic frequencies ν.[71] Evidently, for the integral to exist one had to avoid that the path of ω-integrations should go through a point which makes the resonance denominator disappear. As a proper procedure to do so, Jordan suggested to replace the right-hand side of Eq. (66) by the expression

$$S_1 = -\sum_\tau \mathbf{A}_\tau \exp(2\pi i w^0\tau) \int_{t_0}^{t} \mathbf{E}(t') \exp\left[-2\pi i \nu\tau (t - t')\right] dt', \tag{66'}$$

which contained no resonance denominator, and he managed to handle expressions similar to the one for S_1 in the same consistent manner (Jordan, 1925b).

Once they had established the exact classical formulae, such as the ones for the energy absorbed by the periodic system or the fluctuation of the energy, Born and Jordan had to translate them into quantum formulae by using Born's discretization prescription, i.e., by replacing certain classical differential quotients (with respect to the action variables) by the corresponding difference quotients. In this task also Jordan assisted Born effectively because he was well acquainted with the methods of difference calculus. Thus, the two of them developed a systematic method of writing the difference expressions occurring in their theory in close agreement with the ones occurring in mathematical literature, especially in Nörlund's book (Nörlund, 1924), which allowed them to carry out calculations in a consistent manner. They used their formalism to demonstrate some of the

[71] ω refers to the frequency of the external electric field, ν to the (nonperturbed) ground frequency of the atomic system (which is perturbed by the electric field); hence $\nu = \partial w^0/\partial t$. Since the system is multiply periodic, $w^0\tau$ and $\nu\tau$ stand for a sum of terms, each associated with one degree of freedom.

previous results of dispersion theory and to extend the dispersion-theoretic treatment to study the influence of aperiodic fields on atoms. An important example of the latter case occurs when charged particles (electrons or α-particles) are scattered by atoms. In order to account for, say, the angle of deflection of the charged particles, Born and Jordan took the classical formula and replaced in it the classical frequencies and amplitudes by the corresponding quantum-theoretical ones. The result appeared to be in agreement with physical observation, and the authors concluded that the reformulation scheme with the difference calculus worked whenever it could be applied in a correct mathematical way.

Born and Jordan were, of course, aware that their work would develop only slowly into the scheme which Born had already called 'quantum mechanics.' However, they thought that they were marching along the right path for two reasons. First, they were convinced that the difference calculus represented the most adequate mathematical tool to express quantum laws. Second, they believed that they had rediscovered an important guiding principle, which helped them to select that restricted class of previous relations of atomic theory that could be taken over into quantum mechanics: this was the principle of using relations employing only observable quantities. For example, on the basis of this principle, equations containing the unmeasurable phase differences between electron motions of different atoms made no sense; hence one had to average over the phases in such equations before translating them into quantum theory. Jordan liked this principle because it provided a justification for an otherwise arbitrary step in atomic theory. After all, this principle could be traced back to his favourite philosopher Ernst Mach; moreover, it had already been applied successfully by Albert Einstein in developing relativity mechanics and, maybe, it would now help in creating a new atomic mechanics. However, in summer 1925, it was neither Jordan nor Born who took the next step towards the discovery of quantum mechanics, but Werner Heisenberg. And this step was contained in the manuscript which Heisenberg gave to Born in the middle of July to do with it what he saw fit, and which Born forwarded for publication to *Zeitschrift für Physik* (Heisenberg, 1925c).

Jordan had heard of Heisenberg's new ideas on the quantum-theoretical reformulation of kinematic and mechanical relations in private discussions.[72] He certainly found them interesting, not only because he respected Heisenberg as a successful senior colleague (whose results had guided him in his own work on spectroscopy), but also because Heisenberg had applied consistently the principle of avoiding nonobservable quantities in equations of quantum theory. However, he did not think of working along the lines suggested by Heisenberg. Born, on the other hand, was intrigued by the discovery that Heisenberg's ideas required the use of matrices. He quickly wanted to get on with the formulation of

[72] Friedrich Hund recalled that in July 1925 he attended a private seminar, which Heisenberg gave to a small circle of students and friends at Hertha Sponer's home before he left Göttingen to visit Leyden and Cambridge.

quantum mechanics with the help of matrix methods. Upon Pauli's refusal to collaborate, he immediately asked Jordan to help him again. Even more than in the previous work on the extension of dispersion theory and difference methods he needed assistance in the new task, which was not only difficult but seemed to require lots of technical labour.

The first thing which Born asked Jordan to do was to prove a guess. Heisenberg's quantum condition, Eq. (5), had fixed the diagonal elements of the product $\mathsf{pq} - \mathsf{qp}$, with p and q denoting the infinite matrices representing the quantum-theoretical momentum and position variables of a (periodic) system of one degree of freedom. Immediately after his discovery of the possibility of formulating quantum-mechanical equations with matrices, Born tried to rewrite the quantum condition as a matrix equation having the form

$$\mathsf{pq} - \mathsf{qp} = \frac{h}{2\pi i}\,\mathsf{1}, \tag{67}$$

where h denotes Planck's constant and $\mathsf{1}$ the infinite unit-matrix. Equation (67) thus constituted the first equation of the new theory and Born was proud of it. 'But this was only a guess, and my attempts to prove it failed,' he recalled later. 'Jordan . . . after only a few days brought me the solution of the problem' (Born, 1978, p. 218).

In order to establish Eq. (67), Pascual Jordan turned to the example of the one-dimensional anharmonic oscillator, whose equation of motion may be written as

$$\ddot{\mathsf{q}} + \omega_0^2\mathsf{q} + \lambda\mathsf{q}^n = 0, \qquad n = 2, 3, \ldots, \tag{68}$$

where $\ddot{\mathsf{q}}$ denotes the second time derivative of the position coordinate q, ω_0 the frequency constant and λ the parameter of anharmonicity. By multiplying Eq. (68) with $m\mathsf{q}$ (m being the oscillating mass) from the right and from the left, respectively, he obtained two equations whose difference yielded the matrix-equation

$$m(\ddot{\mathsf{q}}\mathsf{q} - \mathsf{q}\ddot{\mathsf{q}}) = 0, \tag{69}$$

where 0 is the infinite zero-matrix. Now the first term on the left-hand side of Eq. (69) could be expressed as the difference of $d(m\dot{\mathsf{q}}\mathsf{q})/dt$ and $m\dot{\mathsf{q}}^2$, where $\dot{\mathsf{q}}$ denotes the first time derivative of the position coordinate; similarly the second term could be expressed as the difference of $d(\mathsf{q}\,m\,\dot{\mathsf{q}})/dt$ and $m\,\dot{\mathsf{q}}^2$. Hence Jordan concluded that Eq. (68) was equivalent to the equation

$$\frac{d}{dt}(\mathsf{pq} - \mathsf{qp}) = 0 \tag{69a}$$

because $m\dot{\mathsf{q}}$ $(= \mathsf{p})$ represents the momentum variable of the oscillator. Now the nondiagonal elements of any matrix describing a property of the anhar-

monic oscillator contain the exponential time-dependent factor, $\exp[(2\pi i/h) \times (E_n - E_m)t]$, where E_n and E_m are the (different) energies of different states denoted by the indices n and m. Hence the nondiagonal elements of the matrix $\mathsf{pq} - \mathsf{qp}$ must be zero, otherwise they could not be constant as required by Eq. (69a). Jordan thus proved Eq. (67) for the case of the one-dimensional anharmonic oscillator.

Jordan's quick performance pleased Born, and they agreed to develop the new theory of matrix mechanics together. After a careful study of Heisenberg's paper Born sent it off to *Zeitschrift für Physik* by the end of July 1925 and left for vacation in Switzerland. Jordan stayed behind in Germany to work on the joint enterprise. Fortunately, he had no difficulty in using matrix methods, because he had—in connection with the assistance he gave to Courant in the writing of *Methoden der mathematischen Physik*—once consulted ('a little bit') Maxime Bôcher's *Einführung in die höhere Algebra* (1911), the German translation of a book in which (finite) matrices were discussed in some detail (Bôcher, 1907). Thus, his mathematical background was just of the right kind for him to become Born's successful principal collaborator in the formulation of matrix mechanics.[73]

A year later Born described his qualities as follows: '[Jordan is] perhaps not so productive as far as practical results are concerned, but [he is] very philosophically minded and mainly interested in fundamental questions. Besides Heisenberg and Pauli, I consider him to be the most gifted of the younger colleagues' (Born to Bohr, 18 December 1926).

[73] Jordan's initial familiarity with matrix methods should not be overestimated. As he himself said about it: 'I knew a little bit, not very much . . . Something on this was contained in Courant's book [Courant and Hilbert, 1924]. And there is [a book on] *Algebra* by [Maxime] Bôcher' (Jordan, AHQP Interview, 17 June 1963, p. 29).

Chapter II
Matching the Tools and the Task

Werner Heisenberg had concluded his paper on the quantum-theoretical re-interpretation of kinematic and mechanical relations by recommending 'a more intensive mathematical investigation of the method which has been very superficially employed here' (Heisenberg, 1925c, p. 893). He had given the completed manuscript to Max Born, expecting that he would look into it cursorily and, if he did not find some serious mistake, send it off for publication. Although Born had not followed Heisenberg's work during the previous several weeks, he soon became absorbed in his results, especially after recognizing the presence of a well-defined mathematical scheme behind them. Towards the end of July 1925 the following situation developed in Göttingen: While Heisenberg was visiting Leyden and Cambridge, his suggestion for the investigation of the mathematical method underlying the re-interpretation of kinematic and mechanical relations was taken up by Born. Born, in turn, recruited Jordan to help him in this investigation, especially in the application of the algebra of matrices to formulate a consistent atomic mechanics.

The prospect of using the advanced methods of matrix algebra made Born happy. Not only did he enjoy mathematical formalism for its own sake, as he had shown often enough in his scientific career, he had actually expected for some years that the solution of problems in atomic physics would require the use of a new mathematical tool. Now he realized that the new mathematical tool was the matrix calculus, and he even happened to be acquainted with it. Born's familiarity with the use of matrices considerably facilitated his task of formulating Heisenberg's ideas in a consistent mathematical language. Indeed, it soon became clear that the quantum-mechanical theory, for which physicists had searched for many years, would be achieved faster than expected. To that end, as a first step, one had to express Heisenberg's quantum-theoretical concepts and relations in term of matrices; then one had to attempt to cast all other dynamical relations into matrix form. Thus, in July 1925 Born proposed the programme of matrix mechanics, i.e., of a reformulation of the complete scheme of classical mechanics (including the dynamics of the electron) as a matrix scheme.[74]

[74]The name 'matrix mechanics' never appeared in the papers of the Göttingen physicists in 1925 and 1926. They always spoke about the 'new mechanics' or 'quantum mechanics.' Other people referred to the matrix scheme of Born, Jordan and Heisenberg as 'Heisenberg's mechanics' or 'Göttingen mechanics.' Even Erwin Schrödinger, who invented the alternative formalism of wave mechanics, followed this nomenclature. However, in fall 1925 some mathematicians in Göttingen denoted the new quantum-theoretical approach with the work '*Matrizenphysik*' ('matrix physics'), as

Together with Pascual Jordan, he carried it out during the following two months. By the end of September they submitted a paper to *Zeitschrift für Physik*, in which they presented the new matrix mechanics (Born and Jordan, 1925b).

II.1 The Programme of Matrix Mechanics

At the outset of their paper, '*Zur Quantenmechanik*' ('On Quantum Mechanics'), which was received by *Zeitschrift für Physik* on 27 September 1925, Born and Jordan stated their programme as follows:

> 'The theoretical approach of Heisenberg [1925c] recently published in this Journal, which aimed at setting up a new kinematical and mechanical formalism in conformity with the basic requirements of quantum theory, appears to us of considerable potential significance. It represents an attempt to render justice to the new facts by setting up a new and really suitable conceptual system instead of adapting the customary conceptions in a more or less artificial and forced manner. The physical reasoning which led Heisenberg to his development has been so clearly described by him that any supplementary remarks appear superfluous. But, as he himself indicates, in its formal, mathematical aspects his approach is but in its initial stages. His hypotheses have been applied only to simple examples without being fully carried through to a generalized theory. Having been in an advantageous position to familiarize ourselves with his ideas throughout their formative stages, we now strive (since his investigations have been concluded) to clarify the mathematically formal content of his approach and present some of our results here. These indicate that it is in fact possible, starting with the basic premises given by Heisenberg, to build up a closed mathematical theory of quantum mechanics which displays strikingly close analogies with classical mechanics, but at the same time preserves the characteristic features of quantum phenomena. (Born and Jordan, 1925b, p. 858)

At the start of their investigation Born and Jordan found themselves in a somewhat difficult situation because Heisenberg, who had initiated the physical ideas, was away from Göttingen and therefore unavailable for discussions. They could not even ask him whether he agreed with their continuation of his work or whether he would prefer to proceed alone in a different direction. Born was conscious of this situation, especially since Pauli had warned him to keep his hands off and not spoil Heisenberg's physical ideas by mathematical formalism. On the other hand, he could, with clear conscience, point to the fact that Heisenberg had given him a finished paper in which he had written down all his results and that this paper would be published before long; hence every physicist

---

Heisenberg mentioned disapprovingly to Pauli (Heisenberg to Pauli, 16 November 1925). In his review of Born's *Problems of Atomic Dynamics* (Born, 1926d), Edwin C. Kemble then spoke explicitly about 'matrix mechanics' (Kemble, 1926b, p. 424). After 1927 the notion of 'matrix mechanics' became standard in the literature (see, e.g., Allen, 1927, p. 78, or Thirring, 1928, p. 405). Later, even Born and Jordan used it in their book *Elementare Quantenmechanik* (1930).

had the right to pick up the ideas and elaborate on them. Why should one blame him, Max Born, for using the advantage of having seen Heisenberg's results first? Moreover, he was convinced that the recognition of the possibility of rewriting quantum-mechanical relations in the language of matrices would take atomic theory a great step forward. Thus, it made no sense to await Heisenberg's return, especially since he would not come back to Göttingen immediately after his summer vacation, but spend a few weeks in the fall at Bohr's Institute in Copenhagen. Finally, his absence did not really prevent Born and Jordan from going ahead in late July 1925, for they believed that they had a well-defined programme of their own.

Immediately after Jordan proved Born's conjecture about the matrix formulation of the quantum conditions as Eq. (67), their collaboration on the new theory began in earnest. As Jordan recalled: 'Born explained something about Heisenberg's ideas to me . . . I realized that the [matrix] multiplication which Born had already recognized constituted the fundamental idea, and that one could use this multiplication in order to formulate equations involving [Heisenberg's] patterns of observable [quantum-theoretical] quantities. Since this idea convinced me at once, it determined already fully the programme of our joint collaboration' (Jordan, AHQP Interview, 18 June 1963, p. 19). However, Born and Jordan did not spend many days together working on the matrix formulation of mechanical relations. After a few days they had to interrupt their work: Born left for Tübingen to deliver a lecture on 30 July, and from there he went on to Switzerland for his vacation. 'I was so tired out that I had a little breakdown and was compelled to keep quiet,' Born remembered later (Born, Conversations). However, the elaboration of matrix mechanics continued, for as Jordan reported:

> During Born's stay at Silvaplana I was in Hanover at my parents' house, and there I thought about a part of the material that was then treated in the paper of Born and myself. I exchanged letters about it with Born, to whom I reported, of course, my progress. I still recall that, after some time, he considered it desirable to interrupt the exchange of correspondence, since the double strain of a fatiguing sanatorium cure and our discussion through letters on this exciting subject did not agree well with him. (Jordan to van der Waerden, 8 October 1964)[75]

Born stayed in Silvaplana, in Engadine, until the middle of September and returned to Göttingen via Zurich.[76] Thus, the main burden of the work fell on Pascual Jordan, who proved himself equal to the task. Besides the advice, which

[75] In his recollections Born gave a different schedule of the origin of the first paper on matrix mechanics. He said: 'After Jordan and I had sent off our paper [Born and Jordan, 1925b] we separated. I went to Silvaplana in the Engadine to join my family' (Born, 1978, p. 219). This story is not confirmed by the facts, for the paper was submitted only later in September, after Born's return from Switzerland to Göttingen.

[76] These dates and facts are confirmed by recollections and letters. As Born wrote: 'My wife remembers that it was very cold and snowed on 14 September, a day before our departure for home, and our children could toboggan' (Born to van der Waerden, 8 June 1962). Born's visit to Zurich, on the other hand, was mentioned by Hermann Weyl in a letter to Jordan, dated 13 November 1925.

he received through letters from Born, he had only little material to base his treatment on. On the one hand, there were the notes which Born and he had taken from Heisenberg's paper before it had been forwarded to *Zeitschrift für Physik*; on the other, he had available some literature about the mathematics he needed. In particular, Jordan consulted *Courant–Hilbert* (1924) and the book of M. Bôcher (1907, 1911) for that purpose. 'What we needed in mathematics,' he recalled later, 'was the fundamentals and, in part, what one knew about quadratic forms, etc.' (Jordan, AHQP Interview, 19 June 1963, p. 2). With these few tools, however, he marched rapidly, as one may conclude from Heisenberg's letters of those days. Heisenberg had been informed by Born early in August about the new collaboration. On 20 August he wrote to Jordan from Munich: 'I heard from Born that you have made great progress in quantum mechanics, and I would naturally be enormously interested to learn something about your calculations. Since, as Born wrote to me, you would like to have galley proofs of my paper [Heisenberg, 1925c], I shall send you the manuscript; unfortunately I do not have the proof-sheets anymore. I shall be very grateful to you if you could briefly report to me about your calculations.' Jordan duly informed Heisenberg, who replied a few weeks later, still from Munich: 'I am very glad that you believe that you have found the proof of the frequency condition, and I request you to write it to me to Copenhagen as soon as you can.' And he added: 'I shall be happy—in case your proof is really all right—if you would publish it' (Heisenberg to Jordan, 10 September 1925).

In August 1925 Jordan had already developed the matrix scheme so far that he could proceed to give a general proof of energy conservation and of the frequency condition in quantum mechanics. Naturally Heisenberg, who, in his example of the anharmonic oscillator had only given the proof in second-order perturbation theory, was extremely interested in these results. He learned that Born and Jordan had conceived a much more ambitious scheme of quantum mechanics than he had considered in his previous paper (Heisenberg, 1925c). There he had employed the Newtonian equation of motion to describe specific examples of mechanical systems, such as the anharmonic oscillator and the rotator. He had not referred, however, to the Hamiltonian scheme, which had provided the basis of the old atomic theory. To Born it was evident that one had to employ the Hamiltonian scheme in the new theory also. 'I had written this book on atomic mechanics [Born, 1925] completely from the Hamiltonian standpoint,' he said later, 'and I was convinced that the best way to present all the fundamental laws [of quantum mechanics] was the Hamiltonian form' (Born, Conversations, p. 39). Therefore the task which he set himself and Jordan was to formulate a new 'quantum mechanics' by re-interpreting the classical dynamical laws of Newton and Hamilton in terms of matrix equations.

Born and Jordan did not wish to confine themselves merely to translating Heisenberg's considerations in the language of matrices, in order 'to clarify the formal content of his approach,' although that task would have served as a valid motivation for their undertaking. Rather, their intention was to obtain the

quantum laws corresponding to *all* the fundamental equations of classical mechanics and electrodynamics, including even relativistic considerations. For this purpose they had to examine carefully the existing matrix methods and see whether they were adequate to do the job. In Hamiltonian mechanics, for example, all the dynamical variables of a system of n degrees of freedom may be expressed as functions of q_k, the generalized position coordinates, and p_k, their conjugate momenta, with $k = 1, \ldots, n$. Hence Born and Jordan, in their re-interpretation, had to consider rather general functions of matrices $\mathbf{q}_k$ and $\mathbf{p}_k$ denoting the dynamical variables in matrix mechanics. In principle, the mathematical tools for doing so were available in matrix algebra, which they adapted for their purposes. The reformulation of the equations of mechanics and electrodynamics in the language of matrices required two other operations which did not seem to be available and had to be defined: these were the differentiation of a dynamical variable (expressed as a matrix) with respect to time, and the differentiation of a matrix function with respect to one of its (matrix) variables. By introducing both of these differential operations Born and Jordan completed the formal tools necessary for the reformulation of the dynamical laws in matrix form.

When Born returned to Göttingen soon after 15 September 1925 he found that Jordan had most of the material ready for their joint paper. All he had to do was to give it, during the following week or so, the finishing touch and to submit it to *Zeitschrift für Physik*, where it was received on 27 September 1925 and published the following November.[77] We shall discuss the essentials of this paper in the following.

II.2 Operations with Matrices

In Maxime Bôcher's *Algebra* (1911), the book which Jordan consulted for mathematical questions, a quadratic matrix was defined as an array consisting of an equal number of rows and columns, which was framed on both sides by vertical double bars. Born and Jordan did not follow this notation; rather they took over the one from *Courant–Hilbert* (see Chapter 1, Section 1.2 of Courant and Hilbert, 1924), thus replacing the double bars by round brackets.[78] They represented an infinite matrix by a bold letter, say $\mathbf{a}$, in order to distinguish it from other quantities, and denoted its coefficients or elements by $a(nm)$, where

[77] This was confirmed by Jordan, for he remarked: 'It might indeed have been so . . . that when we [i.e., Born and Jordan] met again [in September 1925 in Göttingen], I had already written down to a large extent the first draft of the paper. (To be sure, I do not any longer have more accurate recollections concerning the details associated with it.) I think it is probable that the date of reception, i.e., 27 September 1925, represented entirely correctly the arrival of our manuscript at the journal' (Jordan to van der Waerden, 8 October 1964).

[78] Arthur Cayley had written the arrays of coefficients between single vertical lines (Cayley, 1855, p. 282). Later, this notation became standard for determinants (see, e.g., in *Courant–Hilbert* and in Bôcher's book).

the indices n and m referred to the nth row and mth column, respectively, with n, $m = 0, 1, 2, \ldots$.[78a] Thus, the matrix $\mathbf{a}$ could be written as

$$\mathbf{a} = (a(nm)) = \begin{pmatrix} a(00) & a(01) & a(02) & \ldots \\ a(10) & a(11) & a(12) & \ldots \\ a(20) & a(21) & a(22) & \ldots \\ & \vdots & & \end{pmatrix}. \tag{70}$$

By putting the indices within round brackets instead of using lower indices as in mathematics (e.g., in Bôcher's book or in *Courant–Hilbert*), Born and Jordan not only wished to indicate the relationship of their matrix elements with Heisenberg's transition amplitudes, but also to retain the freedom of adding further (upper and lower) indices to the matrix itself should it be necessary—as, for instance, in describing systems having several degrees of freedom.

After introducing this notation they turned to the operations of addition and multiplication, as defined in the mathematical literature. Thus, for example, they got the sum of two matrices, $\mathbf{a}$ and $\mathbf{b}$, by adding the coefficients, $a(nm)$ and $b(nm)$, at the same position, resulting in the coefficient $c(nm)$ of the matrix sum, $\mathbf{c} = \mathbf{a} + \mathbf{b}$. The product $\lambda\mathbf{a}$ of a matrix $\mathbf{a}$ by the numerical factor λ was represented by $(\lambda a(nm))$. These operations satisfy the usual laws of arithmetic; especially, matrix addition is commutative. The commutativity property does not extend to 'matrix multiplication,' the operation introduced by Arthur Cayley under the name of 'composing' (Cayley, 1855, p. 284). In the case of two infinite matrices, $\mathbf{a}$ and $\mathbf{b}$, Born and Jordan wrote the product as

$$\mathbf{ab} = \mathbf{c} = (c(nm)), \qquad n, m = 0, 1, 2, \ldots,$$

with

$$c(nm) = \sum_{k=0}^{\infty} a(nk)b(km). \tag{70a}$$

However, the matrix multiplication obeys the associative and distributive laws; that is, the equations

$$(\mathbf{ab})\mathbf{c} = \mathbf{a}(\mathbf{bc}) \tag{70b}$$

and

$$\mathbf{a}(\mathbf{b} + \mathbf{c}) = \mathbf{ab} + \mathbf{ac} \tag{70c}$$

hold for any matrices $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$. With the help of these definitions and laws one could then define special matrices, like the unit matrix $\mathbf{1}$ (i.e., the matrix whose

[78a] Scalar matrices are denoted here by regular *sans serif* type. Vector matrices will be denoted by bold *sans serif* type.

product with an arbitrary matrix is identical with that matrix), and the inverse, $\mathbf{a}^{-1}$, of a given matrix $\mathbf{a}$.

To achieve their goal of reformulating the laws of classical mechanics in terms of matrices, Born and Jordan also needed the concept of a function of matrices. This notion had already been considered by Arthur Cayley in his fundamental memoir on matrix methods. For example, in *Courant–Hilbert* such matrix functions $\mathbf{f}(\mathbf{a})$ were considered as are polynomials of a given matrix, i.e., a sum of powers $\mathbf{a}^0\ (=\mathbf{1}), \mathbf{a}^1, \mathbf{a}^2, \ldots, \mathbf{a}^n$, where each term is multiplied with a constant coefficient. Even the exponential of a matrix was defined there as

$$\exp(\mathbf{a}) = \sum_{\nu=0}^{\infty} \frac{\mathbf{a}^\nu}{\nu!} = 1 + \mathbf{a} + \frac{1}{2!}\mathbf{a}^2 + \ldots . \tag{71}$$

Thus, for the purpose of quantum-mechanical theory, Born and Jordan felt justified in admitting the use of the most general 'matrix functions' of matrices, $\mathbf{f}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m)$, depending on m matrix variables, $\mathbf{x}_k$ $(k = 1, \ldots, m)$, which can formally be represented as a sum of a finite or infinite number of products of powers of the $\mathbf{x}_k$ weighted by numerical coefficients. However, the functions $f(x_1, \ldots, x_m)$, with x_k $(k = 1, \ldots, m)$ known variables, do not describe the most general situation either in conventional analysis or in classical mechanics. Frequently one finds that the dependence of the unknown variables y_l $(l = 1, \ldots, n)$, on the known variables x_k is given only by a set of equations involving both the x_k and y_l. Therefore, Born and Jordan had to re-interpret the functional dependence given by these equations in terms of matrices. Hence they also defined the functions $\mathbf{y}_l(\mathbf{x}_1, \ldots, \mathbf{x}_m)$ through the equations

$$\mathbf{f}_1(\mathbf{y}_1, \ldots, \mathbf{y}_n; \mathbf{x}_1, \ldots, \mathbf{x}_m) = 0,$$

$$\ldots \tag{72}$$

$$\mathbf{f}_n(\mathbf{y}_1, \ldots, \mathbf{y}_n; \mathbf{x}_1, \ldots, \mathbf{x}_m) = 0.$$

They argued that such matrix functions made sense because of the following reason: if one assumes that the $\mathbf{y}_l$ $(l = 1, \ldots, n)$ can be expressed as a series of products in increasing powers of the $\mathbf{x}_k$ $(k = 1, \ldots, m)$, each term multiplied with an unknown coefficient, then one obtains the coefficients by inserting the power series *Ansatz* into Eqs. (72). They noted, however, that in the above procedure there was formally a difference with respect to classical theory in the manner of calculating the coefficients of the products of powers of $\mathbf{x}_k$:

> Naturally the number of equations and unknowns exceeds that which would ensue from applying the method of undetermined coefficients in the normal type of analysis incorporating *commutative* multiplication. In each of the equations [(72)], upon substituting the series for the $\mathbf{y}_l$ and gathering together correlated terms, one obtains not only a summation term $C'\mathbf{x}_1\mathbf{x}_2$ but also a term $C''\mathbf{x}_2\mathbf{x}_1$ and has therefore to bring both C' and C'' to vanish (not just only $C' + C''$). This is, however, made

possible by the fact that in the expansion of each of the $\mathbf{y}_l$, two terms $\mathbf{x}_1\mathbf{x}_2$ and $\mathbf{x}_2\mathbf{x}_1$ appear with two available coefficients. (Born and Jordan, 1925b, p. 862)[79]

The discussion of (matrix) functions of matrices signified much more than merely a technical point in the work of Born and Jordan. The development of classical mechanics from Newton to Hamilton had yielded the fundamental result that every dynamical variable of a given dynamical system could be expressed as a function of the sets of generalized position and momentum coordinates, q_k and p_k $(k = 1, \ldots, n)$, where n is the number of degrees of freedom. These functions representing the dynamical variables have the general form of series of products of powers of the q_k and p_k. In formulating the new quantum mechanics rationally, Born had decided to stay as close to Hamiltonian mechanics as possible. Hence the recipe he and Jordan followed was to re-interpret all the classical dependences of dynamical variables, say y_l, on the canonical variables, say x_k (i.e., the set of all q_k and p_k of the system under consideration), whose matrix form could be expressed by Eqs. (72). The important result of their investigation was that the reformulation could be achieved by taking into account the fundamental property of matrices, i.e., their noncommutativity, which had been known since the work of Cayley. The noncommutativity had been first discovered by Heisenberg to hold for dynamical variables in atomic systems. Born and Jordan went beyond Heisenberg by asking the principal question: Given any classical dynamical variable, is it possible to obtain the corresponding matrix expression by means of well-defined operations? In order to answer this question they had to go beyond the known rules of matrix algebra because the classical laws relating physical quantities or variables also involve other mathematical processes, such as the differentiation with respect to a parameter or a variable. What corresponded to such operations in the matrix theory?

One of the most important, if not *the* most important, parameters in classical dynamics is the time variable. Born and Jordan assumed that it plays the same role in quantum theory as well, and that time derivatives of matrices correspond to time derivatives of classical variables. They defined the time derivative of a matrix $\mathbf{a}$, just as in classical theory, by the equation

$$\frac{d\mathbf{a}}{dt} = \dot{\mathbf{a}} = \lim_{\tau \to 0} \frac{\mathbf{a}(t+\tau) - \mathbf{a}(t)}{\tau}, \tag{73}$$

where $\mathbf{a}(t+\tau)$ and $\mathbf{a}(t)$ denote the matrices for neighbouring values of the time parameter, t and $t+\tau$, respectively.[80] Then, by definition of the matrix product,

[79] In their paper Born and Jordan did not go into the problem of convergence of infinite power series of matrices, which, for instance, was discussed in *Courant–Hilbert* (Chapter I, Section 2) in connection with the series $\sum_{\nu=0}^{\infty} \mathbf{a}^{\nu}$.

[80] The definition of the derivative of a matrix with respect to a parameter can be copied from the definition of the derivative of a vector with respect to a parameter. Hence, although it was not contained in the books on matrix algebra, which Born and Jordan consulted, they had no difficulty in writing down Eq. (73).

Eq. (70a), its time derivative follows the usual rule of differentiation, that is,

$$\frac{d}{dt}(\mathsf{ab}) = \dot{\mathsf{a}}\mathsf{b} + \mathsf{a}\dot{\mathsf{b}}. \tag{74}$$

With these relations one can easily define the time derivative of a matrix function, which is expressed as a finite or infinite sum of terms, each of the latter being a product of powers of matrices x_k ($k = 1, \ldots, n$). In other words, the time derivative of an arbitrary analytic matrix function—i.e., a function which can be expanded, at least in certain domains of its variables, in power series—obtains a well-defined meaning. Evidently, the derivatives of matrices and matrix functions with respect to parameters other than the time may be handled in the same manner.

While the differentiation of matrices with respect to a parameter presented no difficulties (because one could proceed in conformity with the usual method of analysis), the differentiation of a matrix with respect to another matrix appeared to be a different matter altogether. For this purpose, Born and Jordan introduced 'symbolic differentiation,' noting right away that 'only in a few respects does this process display similarity to that of differentiation in ordinary analysis. For example, the rules of differentiation of a product or of a function of a function no longer apply here in general' (Born and Jordan, 1925b, p. 862). Only in case that all matrices, which occur in the expressions whose derivative one wants to obtain, do commute with one another, the rules of conventional analysis should still hold. Born and Jordan explicitly considered the differentiation of a product of matrices,

$$\mathsf{y} = \prod_{m=1}^{s} \mathsf{x}_{l_m} = \mathsf{x}_{l_1}\mathsf{x}_{l_2} \ldots \mathsf{x}_{l_s}, \tag{75}$$

with respect to one of its factors. In Eq. (75), the indices $l_1, \ldots, l_s$, may denote any of the indices $1, \ldots, n$, of a set of n matrices x_k; that is, the function y is a general product of powers of the matrices x_k, embracing also products of the type $\mathsf{x}_1^2\mathsf{x}_2\mathsf{x}_1\mathsf{x}_3$. The partial derivative of y with respect to x_k is then defined as

$$\frac{\partial \mathsf{y}}{\partial \mathsf{x}_k} = \sum_{r=1}^{s} \delta_{l_r k} \prod_{m=r+1}^{s} \mathsf{x}_{l_m} \prod_{m=1}^{r-1} \mathsf{x}_{l_m}, \tag{76}$$

where δ_{jk} is the Kronecker symbol, which is unity for identical indices ($j = k$) and zero otherwise ($j \neq k$).[81]

[81] At first sight, the definition of the matrix differentiation of a product, given by Eq. (76), looks unnecessarily complicated. One could surely imagine—as the theoreticians did invent later—a different and simpler definition. This definition was probably due to Jordan, for as Born remarked: 'This was certainly merely an invention of Jordan. Perhaps I helped, but I had the feeling that doing such things was not in my line' (Born, Conversations, p. 38). In inventing the symbolic matrix differentiation, Jordan was also motivated by physical considerations, as we shall discuss in connection with the formulation of the variational principle.

The complicated prescription according to Eq. (76) follows the usual rules of differentiation in analysis only if the product $\mathbf{y}$, Eq. (75), has a special form; e.g., if $\mathbf{y}$ is simply a power of one variable, say $\mathbf{x}_1^s$. Already in the case that $\mathbf{y}$ is a product of powers of two variables, say $\mathbf{x}_1^n\mathbf{x}_2^m$, the rule leads to the result

$$\frac{\partial}{\partial \mathbf{x}_1}(\mathbf{x}_1^n\mathbf{x}_2^m) = \mathbf{x}_1^{n-1}\mathbf{x}_2^m + \mathbf{x}_1^{n-2}\mathbf{x}_2^m\mathbf{x}_1 + \ldots + \mathbf{x}_2^m\mathbf{x}_1^{n-1}, \tag{76a}$$

which is different from the result $n\mathbf{x}_1^{n-1}\mathbf{x}_2^m$ that would be expected from the usual analysis with commuting quantities. However, for a sum of matrices the usual differentiation rule holds; in particular,

$$\frac{\partial}{\partial \mathbf{x}_k}(\mathbf{y}_1 + \mathbf{y}_2) = \frac{\partial \mathbf{y}_1}{\partial \mathbf{x}_k} + \frac{\partial \mathbf{y}_2}{\partial \mathbf{x}_k}, \tag{77}$$

where $\mathbf{y}_1$ and $\mathbf{y}_2$ are any two products of the form given by Eq. (75).

The two rules of symbolic matrix differentiation, stated in Eqs. (76) and (77), uniquely determine the derivative of the most general analytic matrix function. Born and Jordan pointed out that the apparently complicated rules can be expressed in a simple way by using the concept of the 'diagonal sum,' i.e., of the trace of a matrix, which is defined as the sum of all diagonal elements of a matrix, that is,

$$D(\mathbf{y}) = \sum_{n=0}^{\infty} y(nn), \tag{78}$$

where $y(nn)$ denote the diagonal elements of the matrix $\mathbf{y}$. Then the matrix element in the mth *row and* nth *column* of the derivative $\partial\mathbf{y}/\partial\mathbf{x}_k$ becomes identical with the derivative of $D(\mathbf{y})$ with respect to the element $\mathbf{x}_k(nm)$ in *the* nth *row and* mth *column* of the matrix $\mathbf{x}_k$, or

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}_k}(mn) = \frac{\partial D(\mathbf{y})}{\partial x_k(nm)}, \qquad n, m = 0, 1, \ldots. \tag{79}$$

Born and Jordan proved this result with the help of the multiplication rule, Eq. (70a), for matrices having the product form given in Eq. (75); it may easily be extended because of the sum-rule, Eq. (77), to all analytic functions. The definition of the matrix derivative through Eq. (79) not only looks more compact than the equivalent definition through Eqs. (76) and a subsequent application of Eq. (77); it also possesses two further advantages. First, it helps occasionally in actual calculations on account of the special properties of the trace. For example, the trace of a product of two matrices, $\mathbf{x}_1$ and $\mathbf{x}_2$, has the following property: if $D(\mathbf{x}_1\mathbf{x}_2)$ $(= \sum_{n,k} x_1(nk)x_2(kn))$ assumes a finite value, then this value does not depend upon the order of the factors $\mathbf{x}_1$ and $\mathbf{x}_2$—because in that case the two sums over k and n in the trace of the matrix product may be exchanged, which

implies the identity of $D(\mathsf{x}_1\mathsf{x}_2)$ and $D(\mathsf{x}_2\mathsf{x}_1)$. A similar property holds for the trace of a product matrix y having the form given by Eq. (75); $D(\mathsf{x}_{l_1}\mathsf{x}_{l_2}\ldots\mathsf{x}_{l_s})$ does not change its value if the factors in the matrix product are permuted cyclically, i.e., $D(\mathsf{x}_{l_1}\mathsf{x}_{l_2}\ldots\mathsf{x}_{l_s})$ is equal to $D(\mathsf{x}_{l_r}\mathsf{x}_{l_{r+1}}\ldots\mathsf{x}_{l_s}\mathsf{x}_{l_1}\ldots\mathsf{x}_{l_{r-1}})$, $r = 1, \ldots, s$. Hence the partial derivative of the product matrix $\mathsf{x}_{l_1}\mathsf{x}_{l_2}\ldots\mathsf{x}_{l_s}$ with respect to x_k remains the same if one permutes the factors in a cyclic manner. The second advantage of the definition of the derivative through Eq. (79) lay in the fact that it allowed Born and Jordan to formulate the variational principle of classical mechanics in a simple and compact way also in quantum mechanics.

Besides concerning themselves with the symbolic matrix differentiation and some of its implications, Born and Jordan drew attention to another important property of matrix equations, which they expected to be extremely useful in quantum mechanics. They noted, in particular, that every matrix equation involving the matrices $\mathsf{x}_1, \mathsf{x}_2, \ldots, \mathsf{x}_r$ '*remains valid if in all the matrices* x_j *one and the same permutation of all rows and all columns is undertaken*' (Born and Jordan, 1925b, p. 865). In order to prove this assertion they defined the matrix operation by means of which different rows and columns are permuted. Especially, they considered the 'permutation' matrix $\mathbf{\Pi}$, with elements,

$$\prod(nm) = \begin{cases} 1 & \text{if } m = k_n \\ 0 & \text{otherwise,} \end{cases} \tag{80}$$

where k_n assumes the values $0, 1, 2, \ldots$. If they now multiplied a matrix x with $\mathbf{\Pi}$ from the left-hand side, they obtained the result

$$\mathbf{\Pi}\,\mathsf{x} = \left(\sum_{k=0}^{\infty} \prod(nk)\,x(km)\right) = (x(k_n m)); \tag{81}$$

i.e., the product matrix $\mathbf{\Pi}\,\mathsf{x}$ becomes identical to the original matrix x in which the nth row and the k_nth row have been permuted. Similarly, if x is multiplied from the right-hand side by $\mathbf{\Pi}^T$ (where $\prod^T(nm) = \prod(mn)$), the transpose of the permutation matrix, then the columns m and k_m in x are permuted. From this result Born and Jordan drew two conclusions. First, they noted the following property of matrix methods: 'It is thus apparent that from matrix equations one can never determine any sequence or order of rank of matrix elements' (Born and Jordan, 1925b, p. 866). This property implies that the organization of matrix elements in a matrix describing dynamical variables in quantum mechanics must be fixed by additional considerations, as we shall see below in the case of the harmonic oscillator. Second, they speculated that, since matrix equations remain invariant if multiplied from the left by $\mathbf{\Pi}$ and from the right by $\mathbf{\Pi}^{-1}$ $(= \mathbf{\Pi}^T)$, many matrix equations may remain even valid if multiplied from the left-hand side by an arbitrary (nonsingular) matrix b and from the right-hand side by b^{-1}. An invariant property of that kind would later play a role in the theory of canonical transformations of quantum-mechanical systems.

II.3 Dynamical Laws and Energy Conservation

In order to formulate quantum mechanics in matrix form, Born and Jordan first assembled the known facts of matrix algebra and added a few new definitions, such as the one concerning matrix differentiation.[82] With the help of these purely mathematical concepts and methods they then proceeded to develop matrix mechanics. As a first step, they looked for the proper expressions for Heisenberg's ideas in matrix language; after that they generalized these results. They discovered immediately that, instead of using the Newtonian equations of motion as Heisenberg had done, they could equally well start from the Hamiltonian scheme of dynamics. They showed, in particular, that the quantum dynamical laws followed, like the corresponding classical ones, from a fundamental variational principle. Moreover, they found that one of the basic conservation laws of classical mechanics, the one ensuring energy conservation, also held in quantum mechanics, and that the quantum-theoretical frequency condition followed automatically from their matrix scheme.[83]

Of these steps, the translation of Heisenberg's ideas about transition amplitudes, the multiplication rule, etc., came about most easily. Born and Jordan started right away from the concept of transition amplitudes, i.e., the assembly of terms occurring in the quantum-theoretical analogue of classical Fourier series describing multiply periodic systems. According to Heisenberg these terms could be written as $q(nm)\exp[2\pi i\nu(nm)]$ $(n, m = 0, 1, 2, \ldots)$, where the quantum-theoretical frequency $\nu(nm)$ is associated with the transition from the stationary state of the system denoted by the index n to the stationary state denoted by m.[84] Heisenberg had further required that the coefficients, $q(nm)$ and $q(mn)$, which describe the amplitudes of the transition and its inverse, are complex conjugates of each other, and their absolute values, $|q(nm)|$ and $|q(mn)|$, should be the same. Born and Jordan now recognized that Heisenberg's arrays of quantum-theoretical amplitudes could be written in matrix form as

$$\mathsf{q} = \left(q(nm)\exp\left[2i\pi\nu(nm)t\right]\right). \tag{82a}$$

The matrix q was thus supposed to represent the position coordinate of a quantum-mechanical system. Similarly, the canonically conjugate momentum variable had to be given by a matrix p, where

$$\mathsf{p} = \left(p(nm)\exp\left[2\pi i\nu(nm)t\right]\right). \tag{82b}$$

They noted that 'the matrices have to be Hermitian, i.e., on transposition of the

[82] This material is contained in Chapter I, Sections 1 and 2, entitled '*Matrizenrechnung*,' in the published paper (see Born and Jordan, 1925b, pp. 860–866).

[83] The basic formulation and results of quantum mechanics are contained in Chapter II, Sections 3 and 4, entitled *Dynamik*, in Born and Jordan, 1925b, pp. 866–875.

[84] Heisenberg had denoted the ensemble of transition amplitudes as $a(n, n-\alpha)\exp[i\omega(n, n-\alpha)t]$.

matrices, each element is to go over into its conjugate complex value, a condition which should apply for all real t' (Born and Jordan, 1925b, p. 866). With these words they introduced into physics a new concept, that of Hermitian matrices.

What Born and Jordan called Hermitian matrices go back indeed to Charles Hermite, who introduced in the theory of quadratic forms certain substitutions in order to reduce, say, a quadratic form of n variables to a sum of squares (Hermite, 1854).[85] In case the coefficients of these substitutions are real, the related matrix becomes identical with its transpose; in case the coefficients are complex quantities, the real part constitutes a symmetric matrix and the imaginary part an antisymmetric matrix. Hermitian matrices or Hermitian (quadratic or bilinear) forms of finitely many variables had been applied in the second half of the nineteenth century to various problems of mathematics, especially in the theory of invariants and in number theory. Later, David Hilbert, in the last section of his fourth communication to the theory of linear integral equations, had also discussed quadratic Hermitian forms of infinitely many variables (Hilbert, 1906a, Section XII). He had proved, for instance, that compact forms of that type—i.e., Hermitian forms whose real and imaginary parts are separately compact—possess a discrete spectrum of eigenvalues and can be reduced to a sum of squares. Now in quantum mechanics Hermitian forms, or the related matrices or operators, would find their most important application and become known to all theoretical physicists.

Born and Jordan observed that many quantum-theoretical facts fitted naturally into their matrix scheme. For example, the empirical frequency sum-rule

$$\nu(jk) + \nu(kl) + \nu(lj) = 0, \tag{83}$$

which expresses the fact that the frequencies associated with transitions from a given stationary state (index j) via intermediary states (indices k and l) back to the initial state add to zero, could be obtained by considering the diagonal elements of a product of three matrices of the type of Eq. (82a). Further, they incorporated the frequency condition

$$h\nu(nm) = W_n - W_m, \tag{84}$$

[85] Charles Hermite, born in Dieuze, Lorraine, on 24 December 1822, received his education at the *Lycée Louis-le-Grand* in Paris (1840–1842) and at the *École Polytechnique* (1842–1843), but graduated only in 1848 from the latter because he did poorly in examinations. After holding minor teaching positions at the *École Polytechnique* and at the *Collège de France*, he finally became professor at the *Ecole Normale* in 1869 and at the Sorbonne in 1870. Hermite retired in 1897 and died on 14 January 1901 in Paris.

Mathematically productive since the age of twenty, Hermite was one of the principal developers of the theory of algebraic forms, especially quadratic forms, and he contributed importantly to the theory of elliptic and Abelian functions. He was a master of elegant methods and his proofs had artistic quality. He succeeded in giving the solution of a general quintic equation by means of elliptic modular functions, and in showing that the number e (the base of natural logarithms) is transcendental. Hermite was a generous man, who encouraged many young mathematicians in their first work. He was honoured by a foreign membership of the Royal Society (1873) and election to the Academy of Sciences, Paris (1892).

where W_n and W_m denote the energies of the states (denoted by the indices n and m) of the system under consideration, into their scheme in the following way. They assumed that for a quantum-mechanical system, having the discrete energy values W_n ($n = 0, 1, 2, \ldots$), there exists a diagonal matrix W ($= (\delta_{nm} W_n)$) with the following properties: for any dynamical variable, g, of the system—where g is a function of the position and momentum variables and possesses therefore the structure given in Eqs. (82)—the time derivative, $\dot{\mathsf{g}}$, is given by the equation

$$\dot{\mathsf{g}} = \frac{2\pi i}{h} (\mathsf{Wg} - \mathsf{gW}) = \frac{2\pi i}{h} (W_n - W_m)\, g(nm). \tag{85}$$

With the help of these few definitions and concepts, Born and Jordan went on to establish the fundamental equations of matrix mechanics. They started with systems which could be described by the *Hamiltonian matrix* of the form

$$\mathsf{H} = \frac{1}{2m} \mathsf{p}^2 + \mathsf{U}(\mathsf{q}^2). \tag{86}$$

Here m denotes the mass (which is a given number), p and q represent the momentum and position variables, and $\mathsf{U}(\mathsf{q})$ is a matrix function representing the potential energy (which depends only on the position variable of the system under investigation). The Hamiltonian matrix H, Eq. (86), was obtained by reformulating the classical Hamiltonian according to Born and Jordan's programme, i.e., by keeping the formal relations of the classical theory and substituting matrices and matrix functions for the corresponding classical quantities. In the case of such systems (of one degree of freedom) the quantum-theoretical equations of motion possess exactly the same formal structure as the ones in classical dynamics, that is,

$$\dot{\mathsf{q}} = \frac{\partial \mathsf{H}}{\partial \mathsf{p}} = \frac{1}{m} \mathsf{p}, \tag{87a}$$

$$\dot{\mathsf{p}} = -\frac{\partial \mathsf{H}}{\partial \mathsf{q}} = -\frac{\partial \mathsf{U}}{\partial \mathsf{q}}, \tag{87b}$$

because the potential energy matrix U does not depend on the momentum matrix p. In other words, Heisenberg's treatment of the anharmonic oscillator—the only successful example one had at the time—also followed from a Hamiltonian matrix with the help of the equations of motion in a quantum-theoretically reformulated Hamiltonian scheme.

However, Born and Jordan did not stop at this point. They decided to deal also with systems other than the ones described by the Hamiltonian matrix function (86). As they noted:

> We shall now use a correspondence consideration to try to determine the equations of motion belonging to an arbitrary Hamiltonian function. This is required in relation to relativistic mechanics and in particular for the treatment of electron

> motion under the influence of magnetic fields. For in this latter case, the function **H** cannot in a Cartesian coordinate system any longer be represented by a sum of two functions of which one depends only on the momenta and the other on the coordinates. (Born and Jordan, 1925b, p. 868)

In treating such systems a difficulty arises because the symbolic matrix differentiation, as defined by Eq. (76), seems to spoil the Hamiltonian scheme completely. For a given matrix **H**, it yields equations of motion which have a different structure from the ones in classical theory, although in the correspondence limit the quantum-theoretical equations would approach again the classical ones. The difficulty could be traced, of course, to the noncommutativity of the matrices **p** and **q**, which leads to an ambiguity in the products consisting of factors **p** and **q**.

And yet, the differentiation rule (76) from which the ambiguity arose also provided its resolution, and it was for this reason that Born and Jordan invoked their 'correspondence argument.' In particular, they referred to the fact that in classical dynamics the equations of motion of a given system can be derived from an action principle, i.e., the condition that the time integral L, the Langrangian function of the system, assumes an extremum when the actual path during the time interval between t_0 and t_1 is taken, or

$$\int_{t_0}^{t_1} L\,dt = \int_{t_0}^{t_1} (p\dot{q} - H(p,q))dt. \tag{88a}$$

To reformulate condition (88a) in quantum mechanics, Born and Jordan introduced the Lagrangian matrix function **L**, which is defined formally as in classical theory. They further argued that in the case of a sufficiently large time interval, $t_1 - t_0$, only constant terms of **L** contribute to the integral. Hence they suggested that the integration condition (88a) could be replaced by the summation condition

$$D(\mathsf{L}) = D(\mathsf{p}\dot{\mathsf{q}} - \mathsf{H}(\mathsf{p},\mathsf{q})) = \text{extremum}, \tag{88b}$$

where D denotes the diagonal sum and the matrices **p** and **q** have to be varied for fixed frequencies $\nu(nm)$ of the system. This variational principle in quantum mechanics then yielded immediately the Hamiltonian equations of motion. Since the derivatives of $D(\mathsf{L})$ with respect to any element of the actual matrix variables, **p** and **q**, of the system must be zero, the equations

$$2\pi i\nu(nm)q(nm) = \frac{\partial D(\mathsf{H})}{\partial p(mn)} \tag{89a}$$

and

$$2\pi i\nu(mn)p(nm) = -2\pi i\nu(nm)p(nm) = \frac{\partial D(\mathsf{H})}{\partial q(mn)} \tag{89b}$$

followed for all indices n and m $(= 0, 1, 2, \ldots)$. By referring to the matrix differentiation rule, Eq. (79), Born and Jordan immediately noticed that Eqs. (89a, b) led to the usual equations of motion of the Hamiltonian scheme, even for the matrix variables q and p of quantum-mechanical systems that have arbitrary Hamiltonian matrix functions H $(= \mathsf{H}(\mathsf{p}, \mathsf{q}))$. That is, the equations of motion assumed the canonical form,

$$\dot{\mathsf{q}} = \frac{\partial \mathsf{H}}{\partial \mathsf{p}} \tag{90a}$$

and

$$\dot{\mathsf{p}} = -\frac{\partial \mathsf{H}}{\partial \mathsf{q}}\,, \tag{90b}$$

provided the variational principle, Eq. (88b), held.

That a variational principle should be valid also in matrix mechanics seemed evident to Born and Jordan. Indeed, the need for variational principles to govern the laws of physics was clearly recognized by Born throughout his work on atomic problems, although David Hilbert had already held this belief before him.[86] Even when Born had encountered serious difficulties concerning the validity of the application of classical mechanics to such problems of atomic theory as the calculation of the helium states, he nevertheless had continued to think that 'the real quantum mechanics must retain the [complete] Hamiltonian formalism' (Born, Conversations, p. 40). It was from this point of view that he had written his *Vorlesungen über Atommechanik* (Born, 1925), which contained, so he believed, those parts of the old atomic theory which had the possibility of surviving in a future theory. And the variational principle was the very foundation of the Hamiltonian formalism; hence Born believed that the most satisfactory formulation of the new quantum theory of atoms would start from a variational principle. Now he was glad that, with Jordan's help, he had succeeded in expressing the variational principle in matrix mechanics in the simple form of Eq. (88b), and he regarded the derivation of the Hamiltonian equations of motion, Eqs. (90a, b), as a triumph of the variational methods. However, one might argue that the existence of the variational principle is not a *necessary* prerequisite of the dynamical equations in quantum mechanics, for one could just as well postulate the Hamiltonian equations of motion as the foundation of the theory.[87]

[86] David Hilbert had discussed the application of the variational principle to various problems of physics, especially to the theory of gravitation and matter (Hilbert, 1915, 1917). Also in his lectures on the foundations of quantum theory in winter semester 1922–1923 (see footnote 63), he had attributed a leading role to the variational principle in atomic theory.

[87] The preference for deriving the dynamical relations from a variational principle was very much connected with Hilbert and his school in Göttingen and was not shared by theoreticians in general. Thus, Wolfgang Pauli, in his *Encyklopädie* article on relativity, already had criticized Hilbert's derivation of the field equations of gravitation by means of postulating a variational principle (see Pauli, 1921b, p. 705, footnote 277). In 1925 he attributed Born's insistence on a variational principle in quantum mechanics to what he called 'Göttingen's deluge of formal learning' ('*Göttinger Gelehrsamkeitsschwall*', Pauli to Kronig, 9 October 1925).

Quantum mechanics was, of course, not determined completely by reformulating the equations of motion of the corresponding classical theory in matrix form. In addition, the dynamical matrix variables, q and p, of a quantum system (with one degree of freedom) had to satisfy the re-interpreted quantum condition, Eq. (67). This matrix equation, which they called the 'sharpened' ('*verschärfte*,' Born and Jordan, 1925b, p. 871) quantum condition, yielded important information about the structure of matrices which represent the variables in quantum mechanics. By taking the trace on both sides of Eq. (67), Born and Jordan showed that the matrices p and q could not be finite because for two finite matrices the difference of the traces $D(\mathsf{p},\mathsf{q})$ and $D(\mathsf{q},\mathsf{p})$ must be zero. Even more, the matrix products did not even have finite traces and the same held for their difference. Hence Born and Jordan concluded that 'they do not belong to the class of "bounded" infinite matrices hitherto almost exclusively investigated by mathematicians' (Born and Jordan, 1925b, p. 870, footnote (1)). This observation confronted the physicists with serious mathematical difficulties because they were not allowed to take over the well-known mathematical results of the theory of infinite matrices—such as Hilbert's (1906a)—but had to examine carefully the theorems used in each case separately. Their situation resembled somewhat the one which G. W. Hill had encountered fifty years earlier in dealing with infinite determinants in lunar theory; and, like Hill, Born and Jordan had now to justify their mathematical procedure by pointing to the reasonableness of the physical results obtained therewith. In the early steps of the matrix formulation of quantum-mechanical relations, they could only hope that their methods would be justified *a posteriori*. Their lack of caution in running around the minefield of the theory of infinite matrices, picking and choosing what might be useful, only enhanced their confidence as pioneers in the new theory of quantum mechanics.[88]

While Born and Jordan did not concern themselves with a detailed justification of their mathematical methods, they were eager to derive consequences from their quantum-theoretical scheme. As one of the most important they considered the proof of energy conservation, in particular of the fact that the Hamiltonian matrix of a mechanical system (whose analogue in classical theory was a conserved quantity) did not change in time. For this purpose they first restricted themselves to a periodic system having one degree of freedom, described by the Hamiltonian H,

$$\mathsf{H} = \mathsf{H}_1(\mathsf{p}) + \mathsf{H}_2(\mathsf{q}), \tag{91}$$

where the first term on the right-hand side depends on the momentum variable p and the second on the position coordinate q. The functions $\mathsf{H}_1(\mathsf{p})$ and $\mathsf{H}_2(\mathsf{q})$ were assumed to be power series of their variables, say $\sum_s a_s \mathsf{p}^s$ and $\sum_s b_s \mathsf{q}^s$, with constant coefficients a_s and b_s. Many systems in mechanics indeed possess such a

[88] It should be noted that the mathematical parts of the Born–Jordan paper were left mainly to Jordan, who did not have a very deep knowledge of the theory of infinite matrices and did not care for 'mathematical subtleties' (see Jordan, AHQP Interview, 19 June 1963, p. 2).

Hamiltonian function, which is called a 'separable' Hamiltonian. A particular example of such a system was the anharmonic oscillator treated by Heisenberg: in that H_1 had the form $(1/2m)\mathsf{p}^2$, m being the oscillating mass, and H_2 was either $(m/2)\omega_0^2\mathsf{q}^2 + (\lambda/3)\mathsf{q}^3$ or $(m/2)\omega_0^2\mathsf{q}^2 + (\lambda/4)\mathsf{q}^4$, with ω_0 the frequency constant and λ the parameter of anharmonicity.[89] All Hamiltonians of the type (91) share the property that differentiation with respect to the dynamical variables can be carried out easily because only one of the terms, H_1 or H_2, has to be considered and the differentiation rule of classical theory holds.

For the purpose of proving energy conservation Born and Jordan proceeded as follows. First, they showed that the Hamiltonian equations of motion, Eqs. (90a, b), can be replaced by the equivalent algebraic equations,

$$\dot{\mathsf{q}} = \frac{2\pi i}{h}(\mathsf{Hq} - \mathsf{qH}) \tag{92a}$$

and

$$\dot{\mathsf{p}} = \frac{2\pi i}{h}(\mathsf{Hp} - \mathsf{pH}), \tag{92b}$$

involving only the differences of the products of the variables q and p with the Hamiltonian matrix. They derived these equations by replacing the partial derivatives, $\partial\mathsf{H}/\partial\mathsf{p}$ and $\partial\mathsf{H}/\partial\mathsf{q}$, by the terms $\partial\mathsf{H}_1(\mathsf{p})/\partial\mathsf{p}$ and $\partial\mathsf{H}_2(\mathsf{q})/\partial\mathsf{q}$, and taking into account the relations

$$\frac{\partial\mathsf{H}_1(\mathsf{p})}{\partial\mathsf{p}} = \frac{2\pi i}{h}\left[\mathsf{H}_1(\mathsf{p})\mathsf{q} - \mathsf{q}\mathsf{H}_1(\mathsf{p})\right] \tag{93a}$$

and

$$\frac{\partial\mathsf{H}_2(\mathsf{q})}{\partial\mathsf{q}} = -\frac{2\pi i}{h}\left[\mathsf{H}_2(\mathsf{q})\mathsf{p} - \mathsf{p}\mathsf{H}_2(\mathsf{q})\right], \tag{93b}$$

which followed from repeated application of the quantum condition, Eq. (67). For separable Hamiltonians the right-hand sides of Eqs. (93a, b) do not change if H_1 and H_2 are replaced by the full Hamiltonian matrix H; thus Eqs. (92a, b) follow at once.

From the new equations of motion, Born and Jordan concluded that for any analytic function g $(= \mathsf{g}(\mathsf{p}, \mathsf{q}))$ of the momentum and position variables, the time derivative is also given by the equation[90]

$$\dot{\mathsf{g}} = \frac{2\pi i}{h}(\mathsf{Hg} - \mathsf{gH}). \tag{94}$$

[89] It is reasonable to assume that Jordan first proved energy conservation in the case of the anharmonic oscillator and reported about it to Heisenberg in late August 1925. (See Heisenberg to Jordan, 10 September 1925.)

[90] Born and Jordan introduced the notation $|^{\mathsf{H}}_{\mathsf{g}}|$ for the difference of the matrix products, $\mathsf{Hg} - \mathsf{gH}$ (Born and Jordan, 1925b, p. 872). Just a few weeks later, Paul Dirac in Cambridge made a fundamental use of this difference, identifying it—up to the factor $h/2\pi i$—with the quantum-theoretical analogue of the classical Poisson bracket (Dirac, 1925d).

Finally, if they substituted the Hamiltonian matrix H, Eq. (91), for the variable g in Eq. (94), they obtained the equation

$$\dot{\mathsf{H}} = \frac{2\pi i}{h}(\mathsf{HH} - \mathsf{HH}) = 0. \tag{95}$$

Thus, energy conservation was proved to be satisfied *for all quantum systems whose Hamiltonian matrix is separable*. Moreover, Born and Jordan had succeeded, almost without any calculation, in what Heisenberg had attempted to show in the case of the one-dimensional anharmonic oscillator in perturbation approximation. Since, according to the definition of Hermitian matrices (see Eqs. (82)), the Hamiltonian H must then be a diagonal matrix, they derived from Eqs. (92) the relations

$$h\nu(nm) \quad q(nm) = \left[H(nn) - H(mm) \right] q(nm) \tag{96a}$$

and

$$h\nu(nm) \quad p(nm) = \left[H(nn) - H(mm) \right] p(nm). \tag{96b}$$

A similar result holds for any dynamical variable because of Eq. (94). Thus, Born and Jordan also arrived at another important goal: a proof of Bohr's frequency condition, which assumed the form,

$$h\nu(nm) = H(nn) - H(mm), \qquad \text{for } n, m = 0, 1, 2, \ldots, \tag{97}$$

in matrix mechanics.[91]

What was still lacking was the treatment of systems described by a Hamiltonian function which is not separable in the variables p and q, i.e., it contains such products as $\mathsf{p}^s\mathsf{q}^r$. Born and Jordan did not succeed in proving that such a Hamiltonian must always be a diagonal matrix in their scheme. Instead they proposed the following procedure. For an arbitrary nonseparable Hamiltonian, say H^*, they constructed another matrix, H, which contained, say, instead of a term $\mathsf{p}^2\mathsf{q}$, the symmetrized sum $\frac{1}{3}(\mathsf{p}^2\mathsf{q} + \mathsf{pqp} + \mathsf{qp}^2)$. Then, by direct calculation, they proved the theorem: '*To each function* $\mathsf{H}^* = \mathsf{H}(\mathsf{p}, \mathsf{q})$ *there can be assigned a function* $\mathsf{H} = \mathsf{H}(\mathsf{p}, \mathsf{q})$ *such that as Hamiltonians* H^* *and* H *yield the same equations of motion and that for these equations of motion* H *assumes the role of an energy which is constant in time and fulfills the frequency condition*' (Born and Jordan, 1925b, pp. 872–873).

Born and Jordan had thus absorbed all the necessary physical conceptions, including the re-interpreted quantum conditions, into their scheme of matrix mechanics. Moreover, they had been able to state explicitly the fundamental

[91] By comparing Eqs. (94) and (85), one easily finds that the diagonal elements of the Hamiltonian matrix, $H(nn)$, and the energy terms, W_n, are identical; hence the matrix H may be identified with the diagonal matrix W defined earlier.

postulates of the new theory: such as the use of Hermitian matrices for representing dynamical variables; the existence of a time-independent, i.e., diagonal, Hamiltonian matrix H; and the fact that the matrix which expresses the difference of the products pq and qp in the dynamical problem (whose momentum and position variables are p and q) should be diagonal simultaneously with H. By giving these formal requirements, which had to be satisfied in treating any dynamical problem—at least any which possessed one degree of freedom—Born and Jordan had replaced Heisenberg's intuitive approach by a systematic method which was applicable generally.

II.4 An Example of Discrete Mechanics: The Oscillator

In formulating matrix mechanics Born and Jordan had embarked upon the ambitious goal of 'constructing an entirely self-contained theory, without the need to invoke assistance from classical theory on the basis of the correspondence principle' (Born and Jordan, 1925b, p. 876). This meant in particular that for calculations with quantum systems one had to rely entirely on the matrix reformulation of the dynamical equations and the 'sharpened' quantum condition—at least as long as the matrix reformulation involved no ambiguity. Any further reference to the analogous classical system, which had previously been used again and again together with correspondence arguments, should be superfluous because the real content of Bohr's correspondence principle had already been fully absorbed in the reformulation prescription. Born and Jordan chose the anharmonic oscillator to illustrate how the matrix methods worked in a simple practical example. They proceeded in two steps: first, they constructed the sequence of energy states of the harmonic oscillator; then they developed a matrix mechanical perturbation method for obtaining the energy states of the anharmonic oscillator.

The Hamiltonian matrix of the one-dimensional harmonic oscillator, whose mass is normalized to unity for convenience, may evidently be written as

$$\mathsf{H} = \tfrac{1}{2}\mathsf{p}^2 + \tfrac{1}{2}\omega_0^2\mathsf{q}^2, \tag{98}$$

where ω_0 is a parameter determining the strength of the restoring force. The momentum matrix, p, is given by the time derivative of q, the position matrix, and the equation of motion becomes

$$\ddot{\mathsf{q}} + \omega_0^2\mathsf{q} = 0. \tag{99}$$

This *matrix equation* is equivalent to *infinitely many algebraic equations* for $q(nm)$, the infinitely many elements of the Hermitian matrix q, which have the form given by Eq. (82a). Each of these equations, or explicitly,

$$\left[\nu^2(nm) - \nu_0^2\right]q(nm) = 0 \tag{99a}$$

(where $\nu_0 = \omega_0/2\pi$ is a constant of the Hamiltonian describing the system, and $n, m = 0, 1, 2, \ldots$) involves one particular matrix element $q(nm)$.[92] Thus it follows that either this matrix element is zero or the absolute value of the transition frequency connected with it, which is assumed to be related to the energy states through Eq. (84), is given by the constant value ν_0. However, from the application of quantum condition, Eq. (67), it follows that not all matrix elements $q(nm)$ can be zero. In the case of finite $q(nm)$, then, Eq. (99a) implies a strict condition for the absolute value of the energy differences, $W_n - W_m$, of the harmonic oscillator that can be stated as follows: For each index n (denoting a row in the infinite matrix representing a quantum-mechanical variable) there must exist a corresponding index n' (denoting a column) such that $q(nn') \neq 0$ and

$$|W_n - W_{n'}| = h\nu_0. \tag{100}$$

Assuming further that *all energy eigenvalues are different*, or $W_n \neq W_m$ if $m \neq n$ —which appears to be reasonable since the classical one-dimensional harmonic oscillator is a nondegenerate system—one may in fact conclude that, for a given n, there are in general two indices, n' and n'', which satisfy Eq. (100). That is, for any given W_n exactly two different values, $W_{n'}$ and $W_{n''}$, satisfy Eq. (100), and the corresponding transition frequencies are

$$\nu(nn'') = -\nu(nn') = \nu_0. \tag{101}$$

Turning now to the elements of the Hamiltonian matrix (98), written as

$$H(nm) = 2\pi^2 \sum_{k=0}^{\infty} q(nk)q(km)\left[\nu_0^2 - \nu(nk)\nu(km)\right], \tag{102}$$

one can see that, in view of Eqs. (99a), only the diagonal elements, $H(nn)$, are different from zero. Each of these diagonal elements, that is,

$$H(nn) = 4\pi^2\nu_0^2\left[|q(nn')|^2 + |q(nn'')|^2\right], \tag{103}$$

consists of two terms at most, both of which are larger than or equal to zero. Hence the Hamiltonian matrix of the harmonic oscillator is positive definite, and the energy eigenvalues, W_n $(= H(nn))$, have at least a lower bound zero. Moreover, since Born and Jordan had assumed all energy eigenvalues to be different, there had to exist one *uniquely defined lowest or ground state*, say W_{n_0}.

The existence of the lowest energy state could be used for two purposes: first, to calculate—as Heisenberg had proposed previously (Heisenberg, 1925c)—the absolute values of the transition amplitudes $q(nm)$; second, to establish a *natural*

[92] Instead of the quantum frequencies $\nu(nm)$, Born and Jordan used the frequencies $\omega(nm) = 2\pi\nu(nm)$ in order to save on writing. We shall, however, stick to the usual definition of frequencies to avoid confusion. (See, e.g., Born and Jordan, 1925b, p. 883, the last equation of Chapter III.)

order of the elements in the infinite matrix **q**. The key for doing so was provided by the quantum condition (67) which, in the case of the harmonic oscillator, assumed the form of equations

$$|q(nn')|^2 - |q(nn'')|^2 = \frac{h}{8\pi^2 \nu_0} \,. \tag{104}$$

For the index n_0, associated with the ground state, the second term on the left-hand side of Eq. (104) becomes zero because the state n'' is supposed to have a lower energy than the state n which is excluded. The value of $|q(n_0 n_0')|^2$ is then substituted in Eq. (103), yielding the energy $\frac{1}{2}h\nu_0$ for the ground state. Born and Jordan now denoted n_0' by n_1 and obtained from the quantum condition the equation

$$|q(n_1 n_1')|^2 - |q(n_1 n_1'')|^2 = \frac{h}{8\pi^2 \nu_o} \,. \tag{104a}$$

The second term on the left-hand side of this equation was already known to be equal to $|q(n_1 n_0)|^2$ $(= |q(n_0 n_0')|^2)$; hence it was possible to calculate the value of the first term, and with it the energy value of the first excited state. Next, denoting n_1' by n_2, etc., Born and Jordan continued this procedure and obtained the general formulae for the absolute square of the matrix elements,

$$|q(n_k, n_{k+1})|^2 = \frac{h}{8\pi^2 \nu_0}(k+1), \qquad k = 0, 1, 2, \ldots, \tag{105}$$

as well as for the energy terms,

$$W_{n_k} = h\nu_0\left(k + \tfrac{1}{2}\right), \qquad k = 0, 1, 2, \ldots \,. \tag{106}$$

Equation (106) gives a clear order to the energy states of the harmonic oscillator. Any state n must be contained in the set of states n_k, the reason being that the energy states W_{n_k} are nondegenerate and W_{n_0} is uniquely fixed. 'This proves,' Born and Jordan concluded, 'that the indices $0, 1, 2, 3, \ldots$ can be arranged into a new sequence $n_0, n_1, n_2, n_3, \ldots$ such that the formulae [(105)] and [(106)] apply' (Born and Jordan, 1925b, p. 878). With these new indices the solution for the matrix **q** assumed the form

$$q = \begin{pmatrix} 0 & q(01) & 0 & 0 & 0 & \ldots \\ q(10) & 0 & q(12) & 0 & 0 & \ldots \\ 0 & q(21) & 0 & q(23) & 0 & \ldots \\ & \vdots & & & & \end{pmatrix}, \tag{107}$$

which Born and Jordan called the 'normal form.' In this standard arrangement

of indices the energy values, W_n $(= H(nn))$, increase with increasing n. 'If inversely, one requires that $W_n = H(nn)$ should always increase with n,' they noted, 'then it necessarily follows that $n_k = k$ . . . . But thereby only the notation becomes fixed and the calculation more transparent: nothing new is conferred *physically*' (Born and Jordan, 1925b, pp. 878–879).

Thus, two facts made it possible to bring the elements of the matrix $\mathbf{q}$ (and thereby also the elements of the matrices representing other dynamical variables like $\mathbf{p}$, $\mathbf{H}$, etc.) into the 'normal form': the quantum condition, Eq. (104), which ensured the discreteness of the states, and the positivity of the Hamiltonian matrix elements, Eq. (103), leading to the existence of a ground state. No further postulates were necessary to establish the matrix $\mathbf{q}$, Eq. (107), which had been obtained earlier by Heisenberg by guessing from classical theory, and the full solution of the harmonic oscillator, especially its stationary states, could be calculated. Born and Jordan summarized the achievements of their new matrix scheme as follows:

> Therein lies a profound difference between this and the previously adopted semi-classical determination of the stationary states. The classically calculated orbits merge into one another continuously; consequently the quantum orbits selected at a later stage have a particular sequence right from the outset. The new [quantum] mechanics presents itself as a truly discontinuous theory in that herein there is no question of such a sequence of quantum states defined by the physical process; rather the quantum numbers are indeed no more than distinguishing indices which can be ordered and normalized according to any practical standpoint whatsoever (e.g., according to increasing energy W_n). (Born and Jordan, 1925b, p. 879)

In the solution of the harmonic oscillator problem Born and Jordan demonstrated the conceptual differences between a consistent quantum theory and the earlier semiclassical recipes. After this demonstration of principles they were able to deal with more sophisticated systems, especially the anharmonic oscillator.

In the case of the anharmonic oscillator, described by the equation of motion,

$$\ddot{\mathbf{q}} + \omega_0^2 \mathbf{q} + \lambda \mathbf{q}^2 = 0, \tag{108}$$

where $\lambda \mathbf{q}^2$ is a small perturbation term, Born and Jordan obtained the solution in the following way. They expanded the position matrix $\mathbf{q}$ in a series of terms multiplied by increasing powers of the parameter λ, which implied the subsequent expansions of the matrix elements $q(nm)$, that is,

$$q(nm) = q^{(0)}(nm) + \lambda q^{(1)}(nm) + \lambda^2 q^{(2)}(nm) + \ldots , \tag{109a}$$

and of the related transition frequencies $\nu(nm)$, that is,

$$\nu(nm) = \nu^{(0)}(nm) + \lambda \nu^{(1)}(nm) + \lambda^2 \nu^{(2)}(nm) + \ldots . \tag{109b}$$

They did the same with the second time derivative, $\ddot{\mathbf{q}}$, and $\mathbf{q}^2$, the square of the

position matrix. Then they substituted these series into Eq. (108), or into the equivalent algebraic equations for the matrix elements. By gathering together the terms multiplied by the same power of λ they arrived at sets of algebraic equations, each associated with the same power of λ, which could be solved in successive approximations: in zeroth-order approximation one considered only the equations between terms independent of λ; in the first-order approximation one considered the equations between terms multiplied by λ, etc. This procedure had already been used previously by Heisenberg; but now it appeared as an immediate consequence of the matrix methods applied to an oscillator system whose anharmonicity played the role of the small perturbation.

In zeroth-order approximation the matrix elements of the position matrix of the anharmonic oscillator must be identical with those of the harmonic oscillator, in particular,

$$q^{(0)}(n, n+1) = \sqrt{C(n+1)}\, \exp(i\phi_n) \qquad \text{with } C = \frac{h}{4\pi\omega_0}\,, \tag{110}$$

where ϕ_n denotes an undetermined phase ($n = 0, 1, 2, \ldots$). Moreover, the frequencies, $\nu^{(0)}(n, n+1)$, become equal to $-\omega_0/2\pi$, the values of the harmonic oscillator. Besides these quantities, only the matrix elements, $q^{(0)}(n+1, n)$, which are given by $\sqrt{C(n+1)}\, \exp(-i\phi_n)$, and the associated frequencies, $\nu^{(0)}(n+1, n)$ $(= \omega_0/2\pi)$, are different from zero. Proceeding to the next approximation, Born and Jordan derived two sets of equations, that is,

$$\left[\omega_0^2 - 4\pi^2\nu^{(0)}(nm)^2\right]q^{(1)}(nm) - 8\pi^2\nu^{(0)}(nm)\nu^{(1)}(nm)q^{(0)}(nm)$$
$$+ \sum_{k=0}^{\infty} q^{(0)}(nk)q^{(0)}(km) = 0 \tag{111a}$$

and

$$\sum_{k=0}^{\infty} \left\{ \nu^{(0)}(nk)\left[q^{(0)}(nk)q^{(1)}(kn) + q^{(1)}(nk)q^{(0)}(kn)\right]\right.$$
$$\left. + \nu^{(1)}(nk)q^{(0)}(nk)q^{(0)}(kn)\right\} = 0, \tag{111b}$$

where $n, m = 0, 1, 2, \ldots$. Now taking those matrix elements in Eqs. (111a), in which n is arbitrary and $m = n - 1$ and $n - 2$, and observing the relation $2\pi\nu^{(0)}(n, n-2) = 2\omega_0$, which follows from the zeroth-order approximation, they obtained the following contributions to the matrix elements from the first-order

perturbation calculations,[93]

$$q^{(1)}(n,n) = -\frac{C}{\omega_0^2}(2n+1), \tag{112a}$$

$$q^{(1)}(n,n-1) = 0, \tag{112b}$$

$$q^{(1)}(n,n-2) = \frac{C}{3\omega_0^2}\sqrt{n(n-1)}\ \exp\left[i(\phi_{n-1}+\phi_{n-2})\right], \tag{112c}$$

and no contribution to the transition frequency $\nu(n,n-1)$. Similarly they solved the equations of the second-order approximation in a straightforward manner. The latter yielded the finite transition amplitudes $q(n,n-3)$ $(n=4,5,\ldots)$, as well as the corrections to the amplitudes $q(n,n-1)$ $(n=1,2,\ldots)$ and to the related frequencies $\nu(n,n-1)$.

With one exception all first- and second-order contributions to the position matrix of the anharmonic oscillator thus calculated were expressed in terms of the known quantities of the harmonic oscillator solution. The exception occurred in the matrix element $q(n,n-1)$, whose second-order correction, that is,

$$q^{(2)}(n,n-1) = \frac{11}{18\omega_0^4}(Cn)^{3/2}\exp(i\psi_{n-1}), \tag{113}$$

included a phase ψ_{n-1} which was not related to the phase ϕ_{n-1} of the harmonic oscillator occurring in Eq. (110). Born and Jordan expressed the hope that this arbitrariness—because the extra phase seemed to have no explanation—would be removed in the next order of approximation (which they did not calculate). However, at this point they noted: 'We have to leave this question open here but we should like to indicate its significance as a point of principle in regard to the completeness of the entire theory. All questions of statistics invariably depend finally upon whether or not our supposition that of the phases of the $q(nm)$ *one* in each row (or each column) of the matrix remains undetermined be valid' (Born and Jordan, 1925b, p. 881). In other words, the problem of determining the phases ψ_n had to be solved later on by means of a systematic perturbation theory.[94]

Assuming for the moment the identity of ψ_n and ϕ_n, Born and Jordan had arrived at the matrix mechanical solution of the anharmonic oscillator problem up to terms of order λ^2; they had calculated the elements of the position matrix $\mathbf{q}$

[93] From the first-order contributions to $q(n,n-1)$ and $q(n,n-2)$, the contributions for the matrix elements $q(n-1,n)$ and $q(n-2,n)$ can be found from the hermiticity of the matrix $\mathbf{q}$; hence $q^{(1)}(n-1,n)$ and $q^{(1)}(n-2,n)$ are identical with the complex conjugates of $q^{(1)}(n,n-1)$ and $q^{(1)}(n,n-2)$, respectively.

[94] The necessary proof was given in the following paper on matrix mechanics. (See Born, Heisenberg and Jordan, 1926, Chapter 2, Section 2.)

including the exponential functions. This matrix was in the normal form from the zeroth-order approximation onwards because it appeared evident that the perturbation term did not change the order of the states as obtained for the harmonic oscillator. The energy matrix, which is of course diagonal, could be calculated from q and the canonically conjugate momentum p $(= \dot{\mathsf{q}})$ by adding a term $(\lambda/3)\mathsf{q}^3$ to the Hamiltonian (98). The diagonal elements W_n $(= H(nn))$ were then given by the expression

$$W_n = \frac{h\omega_0}{2\pi}\left(n + \frac{1}{2}\right) - \lambda^2 \frac{5}{3}\left(\frac{h}{4\pi\omega_0^2}\right)^2\left[n(n+1) + \frac{17}{30}\right]. \tag{114}$$

This result deviated, as Born and Jordan realized immediately, from the one obtained from the second-order perturbation calculation in the previous semi-classical theory; in the latter, the first term was proportional to the quantum number n and the second to n^2, with the same order of factors as in Eq. (114). Formally it was possible to remove the discrepancy between the old and the new mechanics by introducing half-integral quantum numbers in the old theory, provided one neglected a constant term of order λ^2 in the energy. However, with the advent of quantum mechanics such tricks could be dispensed with. For the purpose of comparing the new results with the known ones, Born and Jordan just pointed out that the quantum-mechanical frequencies, $\nu(n, n-1)$, of the anharmonic oscillator—at least in the second-order approximation—agreed with the corresponding classical frequencies, i.e., with $(1/h)(\partial W_n^{(\mathrm{cl})}/\partial n)$. Finally, they showed that the frequencies $\nu(n, n-1)$ satisfied the frequency condition.[95] This, together with the observation that the quantum-mechanical energy expression, Eq. (114), might also be obtained with the help of the previously reformulated classical perturbation scheme (in agreement with the correspondence principle—see Born, 1924b), concluded Born and Jordan's quantum-mechanical treatment of the anharmonic oscillator.

In spite of the limitation of their success to a simple example, Born and Jordan felt very encouraged by their results. They were convinced that in the matrix scheme they had formulated a quantum theory that surpassed the shaky guesswork of earlier years. The new theory indeed seemed to contain all the requirements which one had tried to establish in atomic theory by a careful application of correspondence arguments to specific problems. These requirements were guaranteed by the very fact that the classical Hamiltonian scheme as well as the quantum condition for multiply periodic systems were taken over in the matrix language. But matrix mechanics signified much more than a clever

[95] The equations for the frequency conditions in the published paper contain several mistakes. For example, the equation involving $\nu(n, n-1)$ exhibits a wrong λ^2-term: a factor $\frac{5}{3}$ must be added there. The frequency condition for $\nu(n, n-2)$ is also incorrect because the λ^2-terms are not included on both sides; actually, Born and Jordan could not determine the λ^2-correction to the frequency $\nu(n, n-2)$ in their second-order approximation.

reformulation of existing approaches: it provided a mathematically consistent and well-defined description of all atomic systems from a unified point of view. Born and Jordan therefore firmly believed that the matrix scheme could and should be applied also to problems that had not been treated before: as, for example, radiation theory. In the last part of their paper (Chapter 4) they went on to deal with this problem.

II.5 Preliminary Remarks on Radiation

Born and Jordan's first goal had been to give a mathematically secure foundation to the concepts of quantum theory. Among the fundamental assumptions in his paper on the reformulation of kinematic and mechanical relations, Heisenberg had proposed to interpret $|q(nm)|^2$, the square of the transition amplitudes associated with the position coordinate of a periodic system, as the probability for a quantum jump in the system (Heisenberg, 1925c, Section 1). In quantum mechanics this assumption meant, of course, a bold step which required justification. In classical theory the emission of radiation from an accelerated electron could be calculated according to the laws of electrodynamics. Especially, there existed a well-defined relation between the quantities characterizing the motion of the electron and the electromagnetic fields characterizing the intensity of the emitted radiation. In the new quantum-mechanical formalism classical quantities, such as the acceleration of an electron, were simply represented by the corresponding matrix quantities. Was one permitted to perform the same reformulation with the variables of the electromagnetic field? Born and Jordan answered this question in the affirmative, but they immediately added that 'the considerations advanced here have only a provisional character' (Born and Jordan, 1925b, p. 883).

One of the main reasons for their cautious attitude was the fact that, in early September 1925, they did not know how to formulate the quantum condition for systems having several degrees of freedom. For the moment they bypassed this difficulty by restricting themselves to free cavity radiation, which may be treated as a system of uncoupled oscillators, each possessing a fixed frequency. As they argued:

> There is hardly any doubt as to how to treat this system. For this purpose the fact, that the fundamental equations of electrodynamics are linear (superposition principle), is of special significance; because from that it follows that the substitute ("*Ersatz*") oscillators [representing the cavity radiation] are *harmonic*, and just in the case of harmonic oscillators—in contrast to other systems—energy conservation is valid independently of the quantum condition . . . One would therefore expect that one may similarly obtain the integral theorems of the electrodynamics of vacuum (energy and momentum conservation) quite generally by solely reformulating Maxwell's equations in matrix language, without taking into account the

> quantum condition. By demonstrating this we simultaneously gain the means of justifying Heisenberg's assertion about the significance of $|q(nm)|^2$. (Born and Jordan, 1925b, p. 884)[96]

In order to formulate Maxwell's equations in matrix language, the meaning of the operations with vectors had to be redefined. For instance, it was necessary to indicate how the scalar and vector products, $\mathbf{A} \cdot \mathbf{B}$ and $\mathbf{A} \times \mathbf{B}$, respectively, of two vector matrices, $\mathbf{A} = (\mathsf{A}_x, \mathsf{A}_y, \mathsf{A}_z)$ and $\mathbf{B} = (\mathsf{B}_x, \mathsf{B}_y, \mathsf{B}_z)$, had to be expressed, because their components would not commute in general.[96a] Born and Jordan decided to retain the usual definitions of the products also in matrix theory (i.e., $\mathbf{A} \cdot \mathbf{B} = \mathsf{A}_x\mathsf{B}_x + \mathsf{A}_y\mathsf{B}_y + \mathsf{A}_z\mathsf{B}_z$, and $(\mathbf{A} \times \mathbf{B})_x = \mathsf{A}_y\mathsf{B}_z - \mathsf{A}_z\mathsf{B}_y$, etc.). Then they wrote for the equations describing the change in time of an electric field matrix $\mathbf{E}$ and a magnetic field matrix $\mathbf{H}$ the matrix equations

$$\operatorname{curl} \mathbf{H} - \frac{1}{c}\dot{\mathbf{E}} = 0 \tag{115a}$$

and

$$\operatorname{curl} \mathbf{E} + \frac{1}{c}\dot{\mathbf{H}} = 0, \tag{115b}$$

where c is the velocity of light *in vacuo*. In composing the (scalar) energy density matrix W, that is,

$$\mathsf{W} = \frac{1}{8\pi}(\mathbf{E}^2 + \mathbf{H}^2), \tag{116}$$

the question of the commutativity did not arise, because only the products of the same factors (e.g., $\mathbf{E}^2 = \mathsf{E}_x^2 + \mathsf{E}_y^2 + \mathsf{E}_z^2$) occurred. As Born and Jordan noted, this W satisfied the equation of continuity in matrix form, that is,

$$\dot{\mathsf{W}} + \operatorname{div} \mathbf{S} = 0, \tag{117}$$

with the vector matrix $\mathbf{S}$, where

$$\mathbf{S} = \frac{c}{8\pi}(\mathbf{E} \times \mathbf{H} - \mathbf{H} \times \mathbf{E}) \tag{117a}$$

represents the quantum-theoretical analogue of the classical Poynting vector. The vector matrix $\mathbf{S}$ should describe, as in classical electrodynamics, the transport of

[96] As we have discussed in Section II.3, the proof of energy conservation (for systems having a separable Hamiltonian) involved the use of the quantum condition. In the case of the harmonic oscillator, however, one can do without it. The Hamiltonian consists of two terms, $\frac{1}{2}\mathbf{p}^2$, where $\mathbf{p}$ is the momentum matrix ($= \dot{\mathbf{q}}$, since the mass is normalized to unity), and $\frac{1}{2}\omega_0^2\mathbf{q}^2$, where $\mathbf{q}$ is the position matrix and ω_0^2 the force constant. The time derivative of the first term gives $\frac{1}{2}\dot{\mathbf{p}}\mathbf{p} + \frac{1}{2}\mathbf{p}\dot{\mathbf{p}}$, which is identical to $-\frac{1}{2}\omega_0^2(\mathbf{qp} + \mathbf{pq})$. The time derivative of the second term gives $+\frac{1}{2}\omega_0^2(\dot{\mathbf{q}}\mathbf{q} + \mathbf{q}\dot{\mathbf{q}})$ $= \frac{1}{2}\omega_0^2(\mathbf{pq} + \mathbf{qp})$. Both time derivatives add up to zero; hence the Hamiltonian matrix is constant in time.

[96a] Vector matrices are denoted here by bold *sans serif* type.

electromagnetic energy; that is, apart from the factor c^{-2}, it should be identical with **g**, the momentum density vector matrix of the radiation. Born and Jordan also went on to prove momentum conservation in matrix electrodynamics, i.e., they succeeded in relating the time derivative of **g** to the divergence of the redefined Maxwell stresses.[97]

The equations of vacuum electrodynamics in matrix form involve well-defined operations on the vector matrices **E** and **H**, provided they describe periodic motions: say, **E** represents a plane wave moving in the x-direction and having the form

$$\mathbf{E} = \left\{ \mathbf{E}(nm) \exp\left[2\pi i \nu(nm) \left(t - \frac{x}{c} \right) \right] \right\}, \tag{118}$$

where $\mathbf{E}(nm)$ is the electric vector amplitude associated with the frequency $\nu(nm)$. Born and Jordan hoped, however, that the matrix equations of electrodynamics would remain valid even if the frequencies $\nu(nm)$ in Eq. (118) were to assume continuous values, so that aperiodic fields could be treated as well. They also did not expect any particular difficulty in casting relativistic equations, such as those describing mechanical and electrodynamic phenomena, into the matrix language. However, for the moment they did not elaborate on these questions; instead they treated the problem of radiation emitted from a Hertzian dipole, especially since it allowed them to answer the question concerning the interpretation of the matrix elements of the position matrix of a periodic system.

In classical electrodynamics an oscillating electron of charge e (which can be considered as a Hertzian dipole), whose position vector is **q** relative to the equilibrium position, loses per second the amount of energy

$$-\frac{dW}{dt} = \frac{2}{3} \frac{e^2}{c^3} (\ddot{\mathbf{q}}^2)_{\text{av}}, \tag{119}$$

through the surface enclosing it; $(\ddot{\mathbf{q}}^2)_{\text{av}}$ denotes the time average of the square of the acceleration vector. Born and Jordan repeated the calculation of the energy loss with matrices: they replaced the vector **q** by the vector matrix **q** (representing the electron's position in quantum theory) and the time average of $\ddot{\mathbf{q}}^2$ by the diagonal sum of the matrix $\ddot{\mathbf{q}}^2$.[98] Assuming that the electron oscillates along a

[97] The components of Maxwell's stress tensor are defined as: $\mathsf{T}_{xx} = 1/8\pi[(\mathsf{E}_x^2 - \mathsf{E}_y^2 - \mathsf{E}_z^2) + (\mathsf{H}_x^2 - \mathsf{H}_y^2 - \mathsf{H}_z^2)]$, etc.; $\mathsf{T}_{yz} = (\mathsf{E}_y\mathsf{E}_z + \mathsf{E}_z\mathsf{E}_y + \mathsf{H}_y\mathsf{H}_z + \mathsf{H}_z\mathsf{H}_y)$, etc., analogous to classical theory. Born and Jordan showed by direct calculation that $\dot{\mathsf{g}}_x$ was equal to the divergence of the vector composed of the three matrices T_{xx}, T_{xy} and T_{xz}, etc.

[98] **Z**, the Hertzian vector matrix of the dipole system, was defined as $\mathbf{Z} = (eq(nm)\{\exp[2\pi i \nu(nm) \times (t - r/c)]\}/r)$, where r is the distance of the point of observation from the equilibrium position of the electron. From **Z**, the electric and magnetic vector matrices **E** and **H** (of the radiation emitted from the electron) could be obtained via matrix formulae corresponding to the ones in classical electrodynamics, and from **E** and **H** the Poynting vector **S** was constructed. (Note that in Born and Jordan, 1925b, p. 887, Eq. (104) is misprinted; the right-hand side should be replaced by $(e/4\pi c^3) \times (\mathbf{r}/r^4) \cdot (\mathbf{r} \times \ddot{\mathbf{q}})^2$, where **r** is the vector denoting the point of observation relative to the equilibrium position—it commutes with all matrices.) The energy loss, $-dW/dt$, was then obtained by integrating **S** over a surface enclosing the oscillating electron.

fixed direction, determined by a scalar matrix $\mathbf{q} = (q(nm)\exp[2\pi i\nu(nm)t]$, they obtained the quantum-theoretical expression for the energy loss as

$$-\frac{dW}{dt} = \frac{64\pi^4 e^2}{3c^3}\sum_{n=0}^{\infty}\left[\sum_{k<n}\nu(nk)^4|q(nk)|^2\right]. \tag{120}$$

In this expression, which had been formulated in such a way that the sum included only the spontaneous radiation from a state n of higher energy to a state k of lower energy, the complete analogy to the classical result could be pointed out: in fact, each transition had the intensity corresponding to the one calculated from the classical electrodynamics.[99]

Equation (120) established a direct numerical relationship between the radiation emitted from an oscillating electron (i.e., a quantum-theoretical dipole) and $q(nm)$, the elements of its position matrix. Born and Jordan had thus proved Heisenberg's interpretation of the $q(nm)$ as transition amplitudes—whose absolute square is proportional to the intensity of the emitted radiation—to be correct.[100] Their result also provided a secure basis for formulating dispersion theory in matrix mechanics. In spite of this success they were not yet able to give a method for computing the matrix elements $q(nk)$ for electrons bound in atoms; such a calculation seemed to involve the particular coupling between the electromagnetic field and the motion of the electron. Thus, Jordan, looking back on the last part of the matrix mechanics paper, remarked: 'What is written about radiation in the first paper of Born and myself is not very profound' (Jordan to van der Waerden, 10 April 1962). However, in September 1925, he and Born were quite content with their preliminary investigation into radiation theory. Since Heisenberg announced that he would join their collaboration, they expected to proceed even faster with their programme of developing a full-fledged scheme of matrix mechanics. The three of them would be able to work out the theory of systems having several degrees of freedom, for as Born and Jordan added a remark in proof to their paper a few weeks later: 'The extension of the theory to several degrees of freedom, [which we have] announced here, has been carried out meanwhile together with W. Heisenberg and will be published as a continuation of this paper. There we shall also discuss several points, already touched upon here, which have been clarified further in the meantime' (Born and Jordan, 1925, p. 888). Among these points belonged, besides the treatment of degenerate systems, a detailed discussion of cavity radiation leading to a derivation of Planck's law from the quantum-mechanical scheme.

[99] Since Einstein's work on the statistical theory of radiation (Einstein, 1916d), it was well known that including spontaneous radiation alone did not yield formulae which agreed with the classical ones in the correspondence limit. The diagonal sum, $D(\ddot{\mathbf{q}}^2)$ $(= \sum_{n=0}^{\infty}\sum_{k=0}^{\infty}[2\pi\nu(nk)]^4 \cdot |q(nk)|^2)$, correctly includes terms other than those associated with spontaneous emission (i.e., terms with $k < n$). Thus, a factor of two is obtained on the right-hand side of Eq. (120).

[100] Born and Jordan argued further that this interpretation remained valid also for systems whose energy states were degenerate.

Chapter III
Completion of the Matrix Scheme

Matrix mechanics was formulated at Göttingen during the University's summer vacation in 1925. Born and Jordan presented the programme of the theory and worked out its most obvious consequences in August and early September; they were then joined by Werner Heisenberg, who, since the end of July, had taken a complete leave from work on physical problems and had spent nearly a month and a half in Munich and on hiking tours in the Alps.[101] In September, Heisenberg did not return to Göttingen; rather he went to Copenhagen, but from there he contributed to the elaboration of the scheme of matrix mechanics in letters to Born and Jordan. Two months later Born, Heisenberg and Jordan submitted a long paper, entitled '*Zur Quantenmechanik. II.*' ('On Quantum Mechanics II'), to *Zeitschrift für Physik*, which covered fifty-nine pages of its volume 35 (Born, Heisenberg and Jordan, 1926; abbreviated as BHJ, 1926, in the following).

It may seem remarkable how this long memoir containing the theory in its most general form—including periodic systems of several, even infinitely many, degrees of freedom—could be composed in such a short period of time, especially since the three co-authors spent little more than a week together.[102] However, several circumstances combined to yield fast progress of the work. First, Heisenberg, on hearing about the early success of the matrix methods—such as the proof of energy conservation—became very enthusiastic about applying them himself. With his active participation Born and Jordan's programme received a big push forward, for he immediately thought of using matrix mechanics in those physical problems with which he had been concerned for some time. Born and Jordan were gratified to have Heisenberg on their side, for they could rely on his physical intuition. Second, in extending the matrix scheme to general multiply periodic systems, all three of them were able to make use of their thorough knowledge of the classical theory, especially of the perturbation schemes and the theory of canonical transformations. Hence they experienced no difficulties in searching for the appropriate classical equations that could be

[101] From Zell-am-Ziller, a little mountain village in Tyrol, Austria, Heisenberg wrote to Bohr on 31 August 1925: 'To be sure, I have not thought at all about physics during the entire past month, and I do not know whether I still understand anything about it.'

[102] Heisenberg returned from Copenhagen to Göttingen around 20 October 1925; his first letter to Niels Bohr from Göttingen was dated 21 October 1925. Born left for his American trip on 28 October 1925 (Born to van der Waerden, 10 July 1962).

reformulated in quantum mechanics. Third, in studying the matrix theory of canonical transformations, Born discovered that mathematically it became identical with the theory of principal-axes transformations of infinite quadratic forms. Thus, the theory developed by his teacher David Hilbert twenty years earlier provided the clue to solving the most important problem of quantum mechanics: the calculation of the energy states of atomic systems. Moreover, it seemed to remove another fundamental difficulty of the old atomic theory that lay in different treatments of discrete and continuous states.

The endeavours of Born, Heisenberg and Jordan led to the development of the theory of matrix mechanics, which was applicable to all types of multiply periodic systems, to nondegenerate and degenerate ones, and in principle even to aperiodic systems. In addition, the authors realized that the matrix equations had a simpler structure than the corresponding classical equations. Thus, for example, the set of Hamilton–Jacobi partial differential equations of classical mechanics was replaced by a set of algebraic equations; their solution established a unitary transformation of the quantum–mechanical Hamiltonian matrix to a diagonal matrix. Now the discussion of conservation laws also appeared to be considerably more elementary. The three co-authors succeeded in presenting the theory in a finished and compact form. Four years later Born and Jordan took it over unchanged into their book *Elementare Quantenmechanik* (1930). This book reflected on every page the debt to the mathematical foundation of quantum mechanics that Hilbert had provided in his theory of integral equations. But it also emphasized the other, the physical foundation of the new atomic theory, which had come from Niels Bohr's correspondence approach, for as Born and Jordan remarked in their preface:

> Indeed, we see in the new mechanics the rigorous realization of Bohr's programme. We believe that the new mechanics does not represent an approximation to classical ideas, nor any departure from Bohr's conceptions (stationary states, quantum jumps, transition probabilities, etc.), but rather their [i.e., of Bohr's conceptions] systematic extension to a logically consistent system. We have expressed this by asking the founder of modern atomic theory [i.e., Bohr] to allow us to dedicate this book to him. (Born and Jordan, 1930, p. VI)

On receiving a copy of the book, Heisenberg wrote to his former collaborators: 'I have been especially happy about the dedication to Bohr. One must at least once publicly emphasize that all the essential ideas of Bohr have also been taken over into the new quantum mechanics' (Heisenberg to Born and Jordan, 20 January 1930).

III.1 The Three-Man Collaboration

During his long vacation in summer 1925 Heisenberg had received the news about Born and Jordan's work on the matrix formulation of the quantum-theoretical ideas contained in the paper which he had completed in Göttingen

before leaving for visits to Leyden and Cambridge. He became very excited with the progress achieved by his colleagues and immediately expressed the desire to use the new formalism, too. For this purpose, on 20 August, he asked Jordan for further information about the details of calculations in matrix mechanics; and he repeated the same request three weeks later in a postcard, asking that the material should be sent to him to Copenhagen where he was going. 'Because,' he wrote, 'from the day after tomorrow I shall begin to work on physics again and I do not, of course, want to repeat your considerations all over again' (Heisenberg to Jordan, 10 September 1925). Thus, Heisenberg also began to work on matrix mechanics in Copenhagen and at once reported about his results to Jordan: 'In the two days that I have been here I have not been able to think about much. However, a very simple derivation of Kramers' dispersion theory and of the formulae of Kramers and myself has occurred to me on the basis of the new formalism' (Heisenberg to Jordan, 13 September 1925). Evidently, the information, which he had meanwhile received from Jordan about matrix mechanics, enabled him to perform calculations with the new scheme himself. Heisenberg was quite satisfied with his initial results, for he remarked: 'I am very glad, for after all I have written everything quite correctly in my paper [Heisenberg, 1925c], provided that this theory makes any sense at all; gradually I begin to believe that this is really true' (Heisenberg to Jordan, 16 September 1925).

With this expression of hope and confidence Heisenberg began his work on matrix mechanics. At this stage, however, he had no familiarity with matrix methods. As he confessed later:

> I had not realized that those things which I had used in my paper were matrices. I had never taken a course on matrix theory. Of course, I knew how to solve linear equations in a trivial way as one learns in school, but I had never learnt the general scheme of matrices and I did not know that matrices can represent groups. So when I learnt from Born that what I had done was really an example of matrix multiplication, I was very interested, but it was new to me. Then, of course, I started to read up on it and learnt it. (Heisenberg, Conversations, p. 163; also AHQP Interview)[103]

The reason why Heisenberg accepted the unusual matrix formalism at once was that it seemed to provide perhaps the only means of handling the greatest difficulty in his quantum-theoretical scheme: the noncommutativity of the re-

[103] Heisenberg's own testimony contradicts some of the legends which are told about his mathematical education. Thus, there is the story that, when Heisenberg returned from Copenhagen to Göttingen in fall 1925, somebody said to him: 'Now what were you doing, sleeping in Hilbert's lectures?' (Heisenberg, Conversations, pp. 162–163). Jordan recalled another story: 'Once, later on, I heard in conversations that there was a legend that Heisenberg, in order to be able to do physics, had learned matrix theory from Emmy Noether. But this is not true. Emmy Noether represented [i.e., taught] a kind of mathematics which Heisenberg did not like at all, did not esteem at all' (Jordan, AHQP Interview, 18 June 1963).

interpreted Fourier series. As he recalled later on:

> In my paper [Heisenberg, 1925c], the fact that [the product of quantum-theoretical variables] XY was not equal to YX was very disagreeable to me. I felt this was the only point of difficulty in the whole scheme, otherwise I would be perfectly happy. But this difficulty had worried me and I was not able to solve it. Only for the example of the anharmonic oscillator that I had worked out was I able to get rid of this difficulty. I had written down the Thomas–Kuhn sum-rule as a quantization rule [see Eq. (5)], but I had not recognized that this was just the **pq** − **qp** condition [in matrix mechanics]. (Heisenberg, Conversations, p. 163; also AHQP Interview)

Now in September 1925 Heisenberg was glad that he had not abandoned the use of the re-interpreted Fourier series just because he had not known how to handle it. The matrix methods of Born and Jordan provided the tool for dealing with noncommuting quantities, and Heisenberg learned an important lesson: 'If one finds a difficulty in a calculation which is otherwise quite convincing, one should not push the difficulty away; one should rather try to make it the centre of the whole thing' (Heisenberg, Conversations, pp. 163–164). In any case, he quickly learned the rules of matrix calculus and turned to physical problems.[104]

The first questions in matrix mechanics that occupied Heisenberg were the proof of the dispersion formulae and the formulation of a perturbation theory for systems of one degree of freedom. He completed both tasks within a few days, reporting his conclusions to Jordan in a letter and a postcard, dated 13 and 16 September, respectively. Then, in a long letter to Wolfgang Pauli on 18 September, he wrote about the calculations in detail: 'Quantum mechanics, which is now very dear to my heart, has meanwhile, mainly through Born and Jordan, made decisive progress, which I shall report to you in the following, even to get a clear idea myself as to what I should believe. One is now really in a position to prove energy conservation, frequency condition and the dispersion formula in a general way on the basis of the previous assumptions [i.e., the assumptions of Heisenberg, 1925c].' After sketching the results of Born and Jordan, Heisenberg presented to Pauli a matrix reformulation of the classical perturbation scheme for multiply periodic systems (Born and Pauli, 1922). 'The equations,' he concluded, 'contain both Kramers' dispersion theory [Kramers, 1924a] and Born's formula for the energy W_2 [i.e., the second-order correction to the energy (see Born, 1924b, p. 389, Eq. (33)] and the formulae established by Kramers and myself [Kramers and Heisenberg, 1925]. I believe that is everything that one can demand for the moment' (Heisenberg to Pauli, 18 September 1925).

[104] For the calculations, which Heisenberg performed in Copenhagen, the information received from Jordan on matrix methods was sufficient. On his return to Göttingen, however, he studied the relevant portions of *Courant & Hilbert* (1924) and the translation of M. Bôcher's book on algebra (Bôcher, 1911). As he recalled: 'I must have looked into the library in Göttingen and tried to learn something. I think then I read some textbooks about matrices, and then I tried to figure out myself how matrices worked' (Heisenberg, Conversations, p. 318). However, already in a letter to Pauli, written on 18 September 1925, Heisenberg mentioned the book of Courant and Hilbert. At that time his knowledge of matrix terminology was not very deep; for example, he said 'unit determinant' instead of 'unit matrix.'

Heisenberg did not content himself with these results, however he went on without delay to introduce another important generalization of the matrix scheme. In a postcard to Jordan dated 21 September 1925, he stated the quantum conditions for systems of several degrees of freedom described by the momentum coordinates p_k and the position coordinates q_k, $k = 1, \ldots, f$, that is,

$$\mathsf{p}_k\mathsf{q}_l - \mathsf{q}_l\mathsf{p}_k = \begin{cases} \dfrac{h}{2\pi i}\,\mathsf{1} & \text{for } k = l \\ \mathsf{0} & \text{for } k \neq l, \end{cases} \tag{121a}$$

and

$$\mathsf{p}_k\mathsf{p}_l - \mathsf{p}_l\mathsf{p}_k = 0, \tag{121b}$$

$$\mathsf{q}_k\mathsf{q}_l - \mathsf{q}_l\mathsf{q}_k = 0. \tag{121c}$$

He added:

> That these conditions yield the frequency conditions, can be demonstrated easily according to your method . . . In case of the separation [of the variables p_k and q_k], the Eqs. [(121)] reduce to your and Born's formulae, but in case of non-separability one cannot understand Eqs. [(121)] from the classical theory without further arguments. However, I still believe, because of the frequency condition, that Eqs. [(121)] are also always true for non-periodic systems . . . I also believe that, if the conditions [(121)] are correct, the perturbation theory can be taken over word for word in the case of several degrees of freedom; of course, in the case of degeneracy, something completely different results. (Heisenberg to Jordan, 21 September 1925)

The key thus seemed to be provided to deal with a large class of quantum-mechanical systems.

Meanwhile Born had returned from his vacation in Switzerland shortly after the middle of September. He quickly completed with Jordan the joint paper on the formulation of matrix mechanics (Born and Jordan, 1925b). At the same time he learned that not only did Heisenberg agree with their work, but had also started to concern himself with the matrix formalism. He was happy to see that Heisenberg had joined his and Jordan's efforts and that an intensive collaboration on the new theory was developing. He regretted, of course, that Heisenberg could participate in the work only by mail from Copenhagen. After some time he felt that all of them needed closer contact, and he wrote to Niels Bohr: 'Now I must ask you not to take it amiss that I have suggested to Heisenberg to come to Göttingen a little earlier, so that we may still be able to work together for a few days. It is unfortunate that I have to go to America just at this time, but I cannot turn down the invitation which I have once accepted . . . We want to release a three-man paper: Heisenberg, Jordan, and myself; this I want to complete at least roughly before leaving' (Born to Bohr, 10 October 1925). Before leaving at the end of October for the United States, where he had accepted to lecture in winter 1925–1926 at the Massachusetts Institute of Technology, Born wanted to

discuss the theory in detail with Heisenberg and Jordan. 'In this way' he wrote to Bohr, 'much useless work will be saved and the thing will be advanced' (Born to Bohr, 10 October 1925).

Born had good reason to request Heisenberg's presence in Göttingen at that moment.[105] It was necessary to coordinate as much as possible the ideas and intentions of the three collaborators. Heisenberg and Jordan had never worked together before; their interests had been different and, when Jordan completed his doctorate under Born in fall 1924, Heisenberg had been away in Copenhagen for a long stay. But even Born and Heisenberg had not collaborated actively for more than a year; the two of them had treated different problems of atomic theory and followed ideas which seemed to go in different directions. Now in fall 1925 the situation had changed: Born discovered that Heisenberg's and his own ideas in quantum mechanics could be formulated systematically in terms of matrices, and Jordan had proved to be his most able helper in doing so. The success of Born and Jordan had convinced Heisenberg into joining the bandwagon of matrix mechanics. However, in spite of their common interest, the three of them had somewhat different goals in mind. Jordan frequently supported Born's intention of establishing a quantum-mechanical scheme to replace classical mechanics (augmented by quantum conditions) that would be as complete and general as possible. Born and Jordan thus wished to fulfill the aim of their first paper (Born and Jordan, 1925b), which was to erect the matrix formulation of quantum mechanics on the basis of a minimum set of assumptions and to extend the formalism to include also (nonrelativistic and relativistic) systems of several, even infinite, degrees of freedom. On the other hand, Heisenberg, once his call for a consistent mathematical scheme had been answered, was less interested in the mathematical foundations of matrix mechanics and much more in its applications to actual problems. In doing so he felt that one should not leave for the moment the familiar domain of previously observed and treated quantum phenomena, like the ones connected with the emission and absorption of spectral lines and the dispersion of light by atoms. Apart from verifying old results, such as the Kramers–Heisenberg dispersion formulae, by means of the new formalism, he wanted to move closer to the ambition he had entertained earlier in May of that year: that of working out a theory of the hydrogen spectrum. Thus he remarked to Pauli: 'To be sure, the most important thing is still missing, namely the actual integration [of the equations of matrix mechanics] in the case of hydrogen' (Heisenberg to Pauli, 18 September 1925). It was for this purpose that he was eager to establish the quantum conditions (121) for systems of several degrees of freedom. Hence he reproached Jordan, who thought of attacking also the problems of electrodynamics—which implied systems of infinitely many degrees of freedom—by writing: 'As for the matrix

[105] Heisenberg's renewed stay at Copenhagen had already been arranged in spring 1925, when he had returned to Göttingen for the summer semester. The reason was that in September 1924 he had received a Rockefeller Fellowship for nine months, but had remained in Copenhagen for only seven months.

calculation in electrodynamics, I would much prefer that one thinks for the moment only about mechanics' (Heisenberg to Jordan, 7 October 1925).

Indeed, there were many points of disagreement between the collaborators that needed straightening out. These concerned both matters of principle and procedure. Born and Jordan had constructed the new theory by setting up a formal mathematical scheme from which the physical results, such as the conservation of energy, could be deduced. In doing so they had emphasized, in particular, the fundamental difference between their approach and the earlier one based on the correspondence principle. Heisenberg, who had followed the Copenhagen tradition of employing the correspondence principle in his own work, did not fully agree with this attitude. He was convinced that even after the discovery of matrix mechanics one could not dispense with Bohr's correspondence principle, and he continued to use it in establishing matrix relations. A more specific point, on which Heisenberg disagreed with Born and Jordan, was the question of the differentiation of matrix functions with respect to a matrix variable.

In their paper, Born and Jordan had proposed a seemingly complicated definition for the partial derivative of a matrix function y with respect to one of its matrix variables, x_k, through Eq. (76). They had, however, argued in favour of this definition because the differential quotient, $\partial\mathsf{y}/\partial\mathsf{x}_k$, could be reformulated in terms of the diagonal sum, $D(\mathsf{y})$, and this alternative definition, Eq. (79), looked much simpler. Moreover, it had enabled them to suggest a variational principle in quantum mechanics, Eq. (88b), from which followed the Hamiltonian matrix equations of motion, similar to the corresponding equations in classical dynamics. Now, in October, Heisenberg raised objections and he wrote to Born:

> I have now thought in more detail about the question of the definition of $\partial\mathsf{f}/\partial\mathsf{q}$ and $\partial\mathsf{f}/\partial\mathsf{p}$, respectively, and I am now convinced that your definition is *not* correct—at any rate, it is not useful. That, by itself, it is a question of suitability as to which definition one chooses, becomes evident from the fact that in the case of Cartesian coordinates—that is, for coordinates in which the problem is defined from the very beginning—the difference between the two definitions does not play a role. However, what appears to me to be a decisive objection against your definition is that with it $(h/2\pi i)[1] = \mathsf{pq} - \mathsf{qp} = \mathsf{PQ} - \mathsf{QP}$ would *not* represent in general a canonical tranformation. For in the case of such a transformation [the Hamiltonian matrix function] H will, in general, pass over into a complicated function of [the transformed momentum and position variables] P and Q, which also contains mixed terms $\mathsf{P}\cdot\mathsf{Q}$, etc.; then, however, the equations of motion are not satisfied. (Heisenberg to Born, 5 October 1925)

Thus, Heisenberg argued that in the case of a Hamiltonian which is a separable function in Cartesian matrix variables—such as the Hamiltonian of the anharmonic oscillator or of the hydrogen atom—a transformation of the variables, which leaves the quantum condition, Eq. (67), invariant, and which he defined as a 'canonical transformation' in quantum mechanics, would lead to a Hamilto-

nian matrix function $\mathsf{H}(\mathsf{PQ})$ that does not satisfy the equations of motion.[106] Heisenberg did not find this conclusion to be acceptable; hence he went on to prove as the 'only practical definition' that one, in which the reformulated canonical equations of motions held for all pairs of dynamical variables admitted in quantum mechanics. This definition was expressed by the equation

$$\frac{\partial}{\partial \mathsf{x}_k}\left(\mathsf{x}_1^{l_1} \cdot \mathsf{x}_2^{l_2} \ldots \mathsf{x}_k^{l_k} \ldots \mathsf{x}_s^{l_s}\right) = l_k\left(\mathsf{x}_1^{l_1} \cdot \mathsf{x}_2^{l_2} \ldots \mathsf{x}_k^{l_k - 1} \ldots \mathsf{x}_s^{l_s}\right) \tag{122}$$

rather than by Eq. (76).

Heisenberg's new definition of the matrix derivative formally came much closer to the usual differentiation procedure. Thus, in the general case of a function $\mathsf{f}(\mathsf{x}_1, \ldots, \mathsf{x}_s)$, the partial differential coefficient $\partial \mathsf{f}/\partial \mathsf{x}_1$ could be written as

$$\frac{\partial \mathsf{f}}{\partial \mathsf{x}_1} = \lim_{\alpha \to 0} \frac{1}{\alpha}\left[\mathsf{f}(\mathsf{x}_1 + \alpha \mathsf{1}, \mathsf{x}_2, \ldots, \mathsf{x}_s) - \mathsf{f}(\mathsf{x}_1, \mathsf{x}_2, \ldots, \mathsf{x}_s)\right], \tag{123}$$

where $\mathsf{1}$ is the unit-matrix and α represents a (real) number going to zero in the limit considered. Heisenberg noticed immediately that now the usual rule of the differentiation of a product of two matrix functions, f and g, i.e., the equation

$$\frac{\partial}{\partial \mathsf{x}_k}(\mathsf{f} \cdot \mathsf{g}) = \frac{\partial \mathsf{f}}{\partial \mathsf{x}_k}\mathsf{g} + \mathsf{f}\frac{\partial \mathsf{g}}{\partial \mathsf{x}_k}, \tag{124}$$

was satisfied. From this rule he derived an important consequence. If F is a function of p_k and q_k, the dynamical variables of a quantum-mechanical system —which satisfy the commutation relations (121)—then the differential quotients, $\partial \mathsf{F}/\partial \mathsf{p}_k$ and $\partial \mathsf{F}/\partial \mathsf{q}_k$, are given by the equations

$$\frac{\partial \mathsf{F}}{\partial \mathsf{p}_k} = \frac{2\pi i}{h}(\mathsf{F}\mathsf{q}_k - \mathsf{q}_k\mathsf{F}) \tag{125a}$$

and

$$-\frac{\partial \mathsf{F}}{\partial \mathsf{q}_k} = \frac{2\pi i}{h}(\mathsf{F}\mathsf{p}_k - \mathsf{p}_k\mathsf{F}). \tag{125b}$$

F may represent any variable of the system, including the Hamiltonian $\mathsf{H}(=\mathsf{H}(\mathsf{p}_k, \mathsf{q}_k))$. Hence the equations of motion assume the particular form

$$\dot{\mathsf{q}}_k = \frac{\partial \mathsf{H}}{\partial \mathsf{p}_k} = \frac{2\pi i}{h}(\mathsf{H}\mathsf{q}_k - \mathsf{q}_k\mathsf{H}) \tag{126a}$$

[106] The reason why the Hamiltonian equations of motion for the transformed Hamiltonian matrix $\mathsf{H}(\mathsf{P}, \mathsf{Q})$ were not satisfied with Born and Jordan's differentiation definition, Eq. (76), was that for a complicated function $\mathsf{f}(\mathsf{P}, \mathsf{Q})$ the differential quotients $\partial \mathsf{f}/\partial \mathsf{Q}$ and $\partial \mathsf{f}/\partial \mathsf{P}$ could not be expressed as differences, $\mathsf{HP} - \mathsf{PH}$ and $\mathsf{HQ} - \mathsf{QH}$, respectively. This was demanded, however, by the reformulated equations of motion, Eqs. (90) and (92).

and

$$\dot{\mathsf{p}}_k = -\frac{\partial \mathsf{H}}{\partial \mathsf{q}_k} = \frac{2\pi i}{h}(\mathsf{H}\mathsf{p}_k - \mathsf{p}_k\mathsf{H}). \tag{126b}$$

From these equations Heisenberg deduced immediately the time independence of the Hamiltonian matrix and the frequency conditions for systems having several degrees of freedom. For this purpose one has to observe only that the time derivative of the momentum variable p_k can be expressed in two ways, that is

$$\dot{\mathsf{p}}_k = \frac{2\pi i}{h}(\mathsf{W}\mathsf{p}_k - \mathsf{p}_k\mathsf{W}) = \frac{2\pi l}{h}(\mathsf{H}\mathsf{p}_k - \mathsf{p}_k\mathsf{H}), \tag{127}$$

where W is the diagonal matrix of the energy eigenvalues of the system under investigation. Hence the frequency conditions followed. In addition, the difference of the matrices, $\mathsf{W} - \mathsf{H}$, must commute with all functions, F, of the dynamical variables, in particular with the Hamiltonian matrix H itself. Therefore the difference $\mathsf{WH} - \mathsf{HW}$ must be zero, and this gives the energy conservation due to Eq. (85) which defines the matrix W.

Evidently, Heisenberg had excellent arguments in favour of his new definition of the matrix derivative. Still he was aware of the fact that it did not agree in general with the definition, Eq. (79), involving the diagonal sum, and that this led to a certain disadvantage. Thus he wrote to Born: 'What appears to me certain from the above is that a nice formulation of the canonical transformations, on one hand, and of the minimum principle, on the other, are not immediately consistent with each other. It would be good if you could perhaps reformulate the minimum principle in such a way that it remains valid when one introduces the new definition of $\partial/\partial\mathsf{p}$' (Heisenberg to Born, 5 October 1925). However, he added in a postscript: 'It just occurred to me that the reason, which speaks in favour of your definition of the differentiation is still evidently the minimum principle. Now I must ask: Is the principle valid at all? For, suppose I make an arbitrary transformation of p, q, to the P, Q, such that $\mathsf{PQ} - \mathsf{QP} = \mathsf{pq} - \mathsf{qp}$, then the equations of motion *do not* hold with your definition of $\partial\mathsf{H}/\partial\mathsf{p}$. This must mean, however, that the minimum principle is not valid in these coordinates (i.e., in P and Q). But if this is so, then what does it mean at all?' (Heisenberg to Born, 5 October 1925). Evidently, Heisenberg preferred to drop the variational principle, which Born and Jordan had been so proud of.

Whether it was the definition of the derivative and the related validity of the Born–Jordan variational principle, or the exact form of the perturbation scheme, all these questions made it necessary that Heisenberg should return to Göttingen, which he did shortly after mid-October.[107] A tough time began, for as Born

[107] Heisenberg wrote to Jordan on 29 September: 'I definitely intend to come to Göttingen between 18 and 20 October, in order to be able to work with Born for a week.' He must have been in Göttingen a few days before 21 October 1925, for on that day he wrote a letter to Bohr, thanking him for his hospitality and mentioning the fact that he was already deeply involved in the work on matrix mechanics with Born and Jordan.

recalled: 'When Heisenberg came it was rather hectic. I had to leave on 2 November [1925] to America, and I wanted to have the paper finished when I left. Heisenberg was very keen on it, too, and so we worked very hard' (Born, Conversations, p. 43).[108] The work was somewhat simplified by the fact that Jordan had already prepared a preliminary manuscript of a paper which served as the basis for discussion. 'Heisenberg's most important contribution, namely the commutation rules for several degrees of freedom, was also contained in this draft' (Jordan to van der Waerden, 8 October 1964). After some debating, the authors also agreed to present both definitions of the matrix derivative, i.e., the definition through Eq. (123) *and* the original Born–Jordan definition, Eq. (76), in their joint paper, but to use only the former definition.[109] The main reason for this procedure was that with Heisenberg's definition one could remove the ambiguity, which Born and Jordan had observed earlier in reformulating non-separable Hamiltonians (see Section II.3). As Jordan remarked many years afterwards: 'The problem, which we treated first, arose primarily from the fact that we took the complicated form of the derivation. After we dropped the latter, the problem disappeared totally' (Jordan, AHQP Interview, 18 June 1963).

Besides this point the three collaborators had many more questions to settle.[110] At times they became discouraged with the amount of work to be done in the writing of the paper. Heisenberg, especially, disagreed with Born and Jordan's intention to include too many items, and he complained to Pauli soon after returning to Göttingen: 'At the moment I am almost exclusively occupied with quantum mechanics, and I have serious doubts as to whether the problem of writing this three-man paper has a solution at all in a finite time' (Heisenberg to Pauli, 23 October 1925). And yet, the authors finally did agree on what to put into the paper. All three of them contributed their share, which Jordan summarized as follows: 'The *S*-transformations [i.e., the canonical transformations] and

[108] As we have mentioned earlier (see footnote 102), Born left Göttingen already on 28 October 1925.

[109] Born recalled later on about this discussion as follows: 'I remember only that I insisted on a certain simplicity, while Jordan had an inclination to make things algebraically involved, as he always does. I didn't want that; we had some little fighting about that. And Heisenberg was even more in my line' (Born, Conversations, p. 41). Actually, as we have mentioned, Heisenberg had proposed the simplification of the definition of the derivative, but probably Jordan defended the old rule, Eq. (76), longer than Born. Jordan himself gave the following account: 'True, in the first paper we had a still more complicated definition . . . We then replaced this definition by a more suitable, more adequate one. After making an effort to show that under certain circumstances both definitions yielded the same results, we then dropped the first definition. It was, however, appropriate to work with the other definition from the beginning' (Jordan, AHQP Interview, 18 June 1963, p. 22).

[110] It should be mentioned that, independently of Heisenberg, Hermann Weyl also arrived at the same definition of the matrix derivative. He informed Born in a letter, dated 27 September 1925, about it. In the same letter he proposed Eqs. (121) as the quantum conditions for systems having several degrees of freedom.

Several months later Fritz London in Stuttgart picked up again the question of the matrix derivative. In a note he described a general definition, from which the two definitions presented in the three-man paper followed (London, 1926a). He further demonstrated that with any definition the Hamiltonian matrix was a constant in time.

their use for perturbation theory was Born's idea. The generalized commutation rules for *several* degrees of freedom came from Heisenberg. The angular momentum investigations were mostly carried out by me, with an important supplement from Heisenberg. The considerations concerning radiation theory—the first *Ansatz* for the later quantization of wave fields—came from me' (Jordan to van der Waerden, 8 October 1964).[111] Each of them concentrated on his part, besides checking the calculations of the others. 'And the writing down of it,' Born recalled, 'we also did in parts and then we put it together like a puzzle' (Born, Conversations, p. 43). Thus, in spite of Heisenberg's anxiety, the main part of the work was accomplished within a few days. And Born could leave Göttingen with a good conscience because the joint paper was basically in hand, though it was completed only afterwards by his collaborators. In Born's absence, Heisenberg became responsible for the final editing. A few weeks later, on 16 November 1925, *Zeitschrift für Physik* received the paper, entitled '*Zur Quantenmechanik. II.*' ('On Quantum Mechanics II'); it was published in one of the issues of February 1926 of Volume 35.

In the three-man paper, Born, Heisenberg and Jordan set out 'to develop further the theory of a general quantum-theoretical mechanics whose physical and mathematical basis has been treated in two previous papers [Heisenberg, 1925c; Born and Jordan, 1925b] by the present authors' (BHJ, 1926, p. 557). They summarized the content of the paper as follows: 'The quantum mechanics developed in Part I of this paper [Born and Jordan, 1925b] from Heisenberg's approach [Heisenberg, 1925c] is here extended to systems having arbitrarily many degrees of freedom. Perturbation theory is carried through for nondegenerate and for a large class of degenerate systems, and its connection with the eigenvalue theory of Hermitian forms is demonstrated. The results so obtained are employed in the derivation of momentum and angular momentum conservation laws, and of selection rules and intensity formulae. Finally, the theory is applied to the statistics of eigenvibrations of a black-body cavity' (BHJ, 1926, p. 557).[112]

After Born's departure from Göttingen, Heisenberg had to coordinate the

[111] Jordan's recollection is substantiated by the information obtained from letters. It also agrees with the recollection of his co-authors, for as Born remarked: 'In the paper we divided our competences. I took the chapters on perturbation theory and on Hilbert space. Heisenberg and Jordan did the angular momentum things—I hardly took any part in it' (Born, Conversations, p. 43).

[112] Apart from the introduction, the Born–Heisenberg–Jordan paper contained four major chapters, divided into thirteen sections: Chapter 1. Systems Having One Degree of Freedom (§1. Fundamental Principles, §2. The Canonical Equations, Energy Conservation and Frequency Condition, §3. Canonical Transformations, §4. Perturbation Theory, §5. Systems for which Time Enters Explicitly into the 'Energy Function'); Chapter 2. Fundamentals of the Theory for Systems Having an Arbitrary Number of Degrees of Freedom (§1. The Canonical Equations of Motion; Perturbation Theory for Non-Degenerate Systems, §2. Degenerate Systems); Chapter 3. Connection with the Theory of Eigenvalues of Hermitian Forms (§1. General Method, §2. Application to Perturbation Theory, §3. Continuous Spectra); Chapter 4. Physical Applications of the Theory (§1. Laws of Conservation of Momentum and Angular Momentum; Intensity Formulae and Selection Rules, §2. The Zeeman Effect, §3. Coupled Harmonic Resonators; Statistics of Wave Fields).

various sections written by the authors separately, and had to rewrite the introduction.[113] As he recalled: 'It was actually a new introduction which used some older parts which Born had written. Born had written an introduction but he was not too satisfied, so I did write a new introduction' (Heisenberg, Conversations). Comparing the content of this introduction with the motivation of the earlier paper by Born and Jordan, it is evident how Heisenberg altered Born's original proposal. He did so especially with respect to the correspondence principle. During Heisenberg's recent stay in Copenhagen he and Bohr had arrived at a common opinion about the new quantum mechanics. Bohr not only accepted the results obtained by Heisenberg and Born and Jordan as 'a step probably of fundamental inportance,' but he regarded the whole apparatus of quantum mechanics 'as a precise formulation of the tendencies embodied in the correspondence principle' (Bohr, 1925b, p. 852). Even though the emphasis in 'Quantum Mechanics I' (Born and Jordan, 1925b) had been to replace the correspondence arguments completely by the matrix formalism, Heisenberg, in agreement with Bohr, now regarded the three-man paper—the first comprehensive exposition of the foundations of modern quantum mechanics in matrix language—as the final culmination of the correspondence principle. He expressed it as follows:

> If one keeps in mind the fundamental differences between quantum and classical theory, differences which stem from the basic quantum-theoretical postulates, then the formalism proposed in the two above-mentioned publications [Heisenberg, 1925c; Born and Jordan, 1925b] and in this paper, if proved to be correct, would appear to represent a theory of quantum mechanics as close to that of classical theory as could possibly be hoped under any reasonable circumstances. In this context we merely recall the validity of energy and momentum conservation laws and the form of the equations of motion (Chapter 1, §2). This similarity of the new theory with classical theory also precludes any question of a separate correspondence principle outside the new theory; rather, the latter can itself be regarded as an exact formulation of Bohr's correspondence considerations. In the further development of the theory, an important task will be to investigate more closely the nature of this correspondence and to describe how the symbolic quantum geometry goes over into the intuitive classical geometry. (BHJ, 1926, p. 558)

While, in the introduction, Heisenberg still expressed a cautious attitude and tried to link the new quantum-mechanical theory as tightly as possible with the erstwhile procedures of Niels Bohr, the main thrust and spirit of the three-man paper were quite different. Especially, in the presentation of the theory of canonical transformations and the perturbation scheme, there prevailed a great optimism concerning the prospects of matrix mechanics. These sections, which

[113] As Jordan recalled later: 'It is my very definite recollection . . . I can say with certainty that the paper, so far as the material it should cover was concerned, was indeed completed before Born's departure; however, it was finally formulated and put together only after his departure. For the final write-up [and arrangement] Heisenberg was responsible' (Jordan to van der Waerden, 8 October 1964).

formed the principal part of the paper, were mainly contributed by Max Born. His programme and intentions ultimately dominated the content of the three-man paper. As he remarked several decades later: 'I think the style [of the three-man paper] is essentially mine' (Born, Conversations, p. 43).

III.2 Towards a New Perturbation Theory

In developing the formalism of matrix mechanics, by means of which actual problems in atomic physics had then to be solved, Born, Heisenberg and Jordan appealed to two guidelines. First, they expected that *all* results, which had been obtained during the previous year on dispersion and related phenomena by systematic guessing involving the correspondence principle, such as Kramers' dispersion formula, should now follow directly from the matrix scheme. Second, they believed firmly that the tried and true methods of classical mechanics could be translated successfully into the new formalism. As Heisenberg recalled: 'The analogy with the classical Hamilton–Jacobi technique was the beginning of all efforts; what we did first was just to try to imitate the old methods as closely as we could' (Heisenberg, Conversations, p. 284). However, in carrying out this programme of imitation, the collaborators arrived initially at quite different-looking results, caused by their separation from each other and their different attitudes towards quantum mechanics. Born and Jordan in Göttingen studied the fundamental papers of Hilbert and his disciples on the theory of infinite quadratic forms and thought of applying the mathematical methods developed in these papers to matrix mechanics. Heisenberg, by temperament and training, was not so familiar with and less interested in this mathematical background; he still attributed to the correspondence principle a guiding role, and regarded the work on matrix mechanics as only a more systematic formulation of his earlier guesswork in atomic physics that had proved to be so successful. Hence their different approaches to perturbation theory in matrix mechanics provide insight into the respective scientific thinking of the three collaborators.

On receiving information about the matrix methods employed by Born and Jordan, Heisenberg in Copenhagen began to apply them himself: he immediately derived, within the new formalism the dispersion formulae of Kramers and of Kramers and himself (Kramers, 1924a, b; Kramers and Heisenberg, 1925), developing for this purpose a first-order perturbation theory which he sketched in a letter to Jordan (Heisenberg to Jordan, 13 September 1925).[114] A couple of days later he announced proudly: 'Going beyond what I have written in my last letter, I have now obtained a complete perturbation theory for problems with one degree of freedom up to an arbitrarily high approximation' (Heisenberg to Jordan, 16 September 1925). He informed Born and Pauli about the details of his

[114] Since Jordan's letters to Heisenberg, written at that time, are not available, we do not know the details of the information which Heisenberg received on matrix theory. However, we may assume with certainty that they did not go beyond what is contained in the Born–Jordan paper.

perturbation theory in long letters.[115] Born had just returned from his vacation to Göttingen and was engaged in completing the final write-up of the first paper on matrix mechanics (Born and Jordan, 1925b). Simultaneously he began to formulate his own perturbation scheme, which appeared to him to be quite different from the one contained in Heisenberg's letter.

Heisenberg, in his formulation of the quantum-mechanical perturbation theory, followed the classical scheme as closely as possible. He considered a periodic system of one degree of freedom, described by the Hamiltonian H, such that H is given by a Hamiltonian H_0 plus terms proportional to powers of a small parameter λ, that is,

$$H = H_0 + \lambda H_1 + \lambda^2 H_2 + \ldots \quad . \tag{128}$$

The solution described by H_0, the unperturbed Hamiltonian, is supposed to be known. $H_0 = H_0(p_0, q_0)$, where p_0 and q_0 are the momentum and coordinate variables, respectively, is identical with the constant energy W_0, which is a function only of the action variable of the unperturbed system (given by the action integral, $\int p_0 \, dq_0$, over one period of the motion). In classical theory, then, the energy of the perturbed system, described by the full Hamiltonian H, Eq. (128), was obtained with the help of a canonical transformation of its dynamical variables, p and q. That is, one had to solve the Hamilton–Jacobi partial differential equation

$$H\left(p = \frac{\partial S}{\partial q}, q\right) = W, \tag{129}$$

where S is the action function and W, the energy of the perturbed system, which depends only on J, the constant action variable of the system. S is assumed to be a function of J and the position variable q; it may be expanded in terms of the perturbation parameter λ, that is,

$$S = q_0 p_0 + \lambda S_1 + \lambda^2 S_2 + \ldots \quad . \tag{130}$$

Hence it was possible to solve Eq. (129) in successive approximations. For this purpose one expressed also the variables of the perturbed system as power series in λ, say $p = p_0 + \lambda p_1 + \lambda^2 p_2 + \ldots$ and $q = q_0 + \lambda q_1 + \lambda^2 q_2 + \ldots$, and substituted these expressions, as well as the expansions (128) and (130), into Eq. (129); then one found the zeroth-order approximation by keeping only terms independent of λ, the first-order approximation by keeping terms proportional to λ, etc.

In trying to reformulate the classical perturbation scheme, Eqs. (128)–(130), in matrix mechanics Heisenberg observed that, in agreement with the matrix

[115] Heisenberg's letter to Born on perturbation theory in matrix mechanics is not available. From the information we have got (see, e.g., Heisenberg's postcard to Jordan, dated 21 September 1925, where he wrote: 'I have already written a sketch of perturbation theory to Born, hence I do not write it again to you; you can read it all in [the letter to] Born . . . '), we can be certain that the perturbation theory discussed in the letters to Pauli (Heisenberg to Pauli, 18 September 1925) and to Born was identical.

differentiation rule (i.e., Eq. (125b) for one degree of freedom), the classical expression $\partial S/\partial q$ had to be replaced—apart from constant factors—by the matrix expression $\mathbf{Sp} - \mathbf{pS}$, involving the transformation matrix $\mathbf{S}$ and the canonically conjugate matrix variable $\mathbf{p}$. This fact allowed him to write down a matrix-theoretical perturbation scheme for a quantum system described by a Hamiltonian matrix $\mathbf{H}$, which, in analogy with the classical Hamiltonian of Eq. (128), can be expanded as a series involving terms multiplied by increasing powers of λ (i.e., $\mathbf{H} = \mathbf{H}_0 + \lambda\mathbf{H}_1 + \lambda^2\mathbf{H}_2 + \ldots$). To begin with, he assumed that $\mathbf{H}_0$, the Hamiltonian of the unperturbed system, was diagonal in the dynamical variables $\mathbf{p}_0$ and $\mathbf{q}_0$. Then he expanded all quantities occurring in the matrix-equation, which corresponds to the Hamilton–Jacobi equation (129), in power series in λ. He found that the position matrix $\mathbf{q}$ assumed the form

$$\begin{aligned}\mathbf{q} = \mathbf{q}_0 &+ \lambda(\mathbf{S}_1\mathbf{q}_0 - \mathbf{q}_0\mathbf{S}_1) \\ &+ \lambda^2\left[(\mathbf{S}_1^2\mathbf{q}_0 - 2\mathbf{S}_1\mathbf{q}_0\mathbf{S}_1 + \mathbf{q}_0\mathbf{S}_1^2) + (\mathbf{S}_2\mathbf{q}_0 - \mathbf{q}_0\mathbf{S}_2)\right] + \ldots,\end{aligned} \tag{131}$$

and that a similar equation held for $\mathbf{p}$, the momentum matrix. 'You will immediately recognize the law for forming the $\mathbf{q}_i$ and $\mathbf{p}_i$, respectively,' Heisenberg wrote to Pauli. 'One always deals with iterated forms of the type $\mathbf{Sq} - \mathbf{qS}$' (Heisenberg to Pauli, 18 September 1925). With the help of these equations he found for $\mathbf{H}$ an expression having the same structure as the right-hand side of Eq. (131), except that $\mathbf{q}_0$ was replaced everywhere by $\mathbf{H}(\mathbf{p}_0, \mathbf{q}_0)$. By substituting the expression for the Hamiltonian matrix into the reformulated Eq. (129)—i.e., the equation $\mathbf{H}(\mathbf{p}, \mathbf{q}) = \mathbf{W}$—and by equating the terms with equal powers of λ, he finally obtained the following set of equations:

$$\begin{aligned}\mathbf{H}_0 &= \mathbf{W}_0, \\ \mathbf{S}_1\mathbf{H}_0 - \mathbf{H}_0\mathbf{S}_1 + \mathbf{H}_1 &= \mathbf{W}_1, \\ \mathbf{S}_2\mathbf{H}_0 - \mathbf{H}_0\mathbf{S}_2 + \tfrac{1}{2}(\mathbf{S}_1^2\mathbf{H}_0 - 2\mathbf{S}_1\mathbf{H}_0\mathbf{S}_1 + \mathbf{H}_0\mathbf{S}_1^2) + \mathbf{S}_1\mathbf{H}_1 - \mathbf{H}_1\mathbf{S}_1 + \mathbf{H}_2 &= \mathbf{W}_2,\end{aligned} \tag{132}$$

etc. (All matrices such as $\mathbf{H}_0$, $\mathbf{H}_1$, $\mathbf{S}_1$, etc., in Eqs. (132) depend on $\mathbf{p}_0$ and $\mathbf{q}_0$, the zeroth-order of the dynamical variables of the system under investigation.)

Heisenberg was very happy with the new scheme which, as he confessed to Pauli, imitated exceedingly well the 'famous Born–Pauli model' (Heisenberg to Pauli, 18 September 1925), i.e., the classical scheme which Pauli had worked out together with Born and Göttingen several years earlier (Born and Pauli, 1922).[116]

[116] For example, the classical equation corresponding to the second equation in Eqs. (132) was $(\partial H_0/\partial J)(\partial S_1/\partial w^0) + H_1 = W_1$, where J and w^0 denote the action variable of the perturbed and the angle variable of the nonperturbed system, respectively (see Born and Pauli, 1922, p. 140, Eq. (7_1)). Now, in the first term the differential quotient $\partial H_0/\partial J$ may be replaced by ν_0, the frequency of the unperturbed system, while the differential quotient $\partial S_1/\partial w^0$ becomes identical with $2\pi i S_1$ because S_1 is assumed to be periodic in w^0. In matrix theory this product, or $\nu_0 h(2\pi i/h)S_1$, then indeed corresponds to the difference, $\mathbf{S}_1\mathbf{H}_0 - \mathbf{H}_0\mathbf{S}_1$.

As he recalled much later: 'The whole thing was that I wanted to make the ideas of Bohr's correspondence principle more and more precise and, finally, I came out with something which was simpler than classical mechanics. That, of course, was a great surprise and a great encouragement. And the perturbation theory was just a very nice point where it happened' (Heisenberg, Conversations, p. 285). His satisfaction, however, was not shared by Born who had meanwhile worked out a different scheme. Born recalled the situation as follows:

> All of us posed the problem [of developing perturbation theory in quantum mechanics]. Then Heisenberg wrote us a letter with the formalism. But before his letter arrived we had already found the transformation of a matrix into another coordinate system [i.e., $\mathbf{f} \rightarrow \mathbf{f}'$] as $\mathbf{f}' = \mathbf{SfS}^{-1}$. And then I looked at Heisenberg's scheme and thought: "That's pure nonsense; it is so complicated, when all he has to do is to develop [the transformation matrix] $\mathbf{S}$ into a power series with respect to a given function." We did this and it came out beautifully. I remained in the belief all the time that Heisenberg's suggestion was wrong. (Born, Conversations, p. 41)

Already in their initial investigations on matrices Born and Jordan had noticed the possibility that every matrix equation may be invariant if one transforms the matrices involved, say $\mathbf{a}$, into $\mathbf{a}' = \mathbf{bab}^{-1}$, where $\mathbf{b}$ is an arbitrary singular matrix and $\mathbf{b}^{-1}$ is its inverse (Born and Jordan, 1925b, p. 866). At the same time, however, they had added that this transformation property did not seem to hold always for equations involving matrix differentiation. Now it seemed to be profitable to apply the transformation invariance in the problem of calculating the eigenvalues of $\mathbf{H}$, the Hamiltonian matrix of a given quantum system. All one had to do was to introduce a transformation matrix $\mathbf{S}$ in such a way that the transformed Hamiltonian became identical with the diagonal energy matrix $\mathbf{W}$, or

$$\mathbf{H}(\mathbf{p}, \mathbf{q}) = \mathbf{S}\mathbf{H}(\mathbf{p}_0, \mathbf{q}_0)\mathbf{S}^{-1} = \mathbf{W}. \tag{133}$$

Evidently, Eq. (133) represented the matrix reformulation of the classical Hamilton–Jacobi equation (129), for it served the same purpose: that is, it provided a method of performing a transformation of the dynamical variables, $\mathbf{p}_0$ and $\mathbf{q}_0$ (by which one initially described the system under investigation), to new variables, $\mathbf{p}$ $(= \mathbf{Sp}_0\mathbf{S}^{-1})$ and $\mathbf{q}$ $(= \mathbf{Sq}_0\mathbf{S}^{-1})$, in which the Hamiltonian is a constant of the time. The beautiful simplicity of this equation and its connection with the mathematical theory of principal-axes transformation of quadratic forms convinced Born and Jordan immediately that it had to be *the* matrix reformulation of Eq. (129).

Born and Jordan proceeded right away to develop their perturbation scheme from Eq. (133). They expanded the transformation matrix $\mathbf{S}$ as the classical analogue in Eq. (130), and obtained the expansion of the inverse matrix $\mathbf{S}^{-1}$ as

$$\mathbf{S}^{-1} = 1 - \lambda \mathbf{S}_1 + \lambda^2(\mathbf{S}_1^2 - \mathbf{S}_2) + \ldots . \tag{134}$$

On substituting the expressions for $\mathbf{S}$ and $\mathbf{S}^{-1}$ into Eq. (133) they derived, by direct calculation, the following set of equations:

$$\mathbf{H}_0 = \mathbf{W}_0,$$

$$\mathbf{S}_1\mathbf{H}_0 - \mathbf{H}_0\mathbf{S}_1 + \mathbf{H}_1 = \mathbf{W}_1,$$

$$\mathbf{S}_2\mathbf{H}_0 - \mathbf{H}_0\mathbf{S}_2 + \mathbf{H}_0\mathbf{S}_1^2 - \mathbf{S}_1\mathbf{H}_0\mathbf{S}_1 + \mathbf{S}_1\mathbf{H}_1 - \mathbf{H}_1\mathbf{S}_1 + \mathbf{H}_2 = \mathbf{W}_2, \qquad (135)$$

$$\mathbf{S}_r\mathbf{H}_0 - \mathbf{H}_0\mathbf{S}_r + \mathbf{F}_r(\mathbf{H}_0, \ldots \mathbf{H}_r; \mathbf{S}_0, \ldots \mathbf{S}_{r-1}) = \mathbf{W}_r,$$

etc., where $\mathbf{H}_0, \mathbf{H}_1, \ldots$, are to be taken throughout as having the arguments $\mathbf{p}_0$ and $\mathbf{q}_0$, and $\mathbf{F}_r$ are matrix functions as indicated, with $\mathbf{F}_1 = \mathbf{H}_1$, $\mathbf{F}_2 = \mathbf{H}_0\mathbf{S}_1^2 - \mathbf{S}_1\mathbf{H}_0\mathbf{S}_1 + \mathbf{S}_1\mathbf{H}_1 - \mathbf{H}_1\mathbf{S}_1 + \mathbf{H}_2$, etc. The perturbation scheme, Eqs. (135), appeared to be simpler than Heisenberg's competing scheme, Eqs. (132); in addition, it turned out to be at least as close to the classical scheme.[117]

From the third equation (second-order approximation) onwards, the Born–Jordan scheme disagreed with Heisenberg's proposal. Since Eqs. (135) had been derived rigorously, Born and Jordan thought that Heisenberg, who was not as familiar with matrix methods at that time as they were, had simply made an error.[118] At first, Heisenberg was bothered by this disparity. However, he soon discovered that his scheme was not really wrong; it just differed from Born and Jordan's by the assumption he had made for the perturbation expansion of the transformation function $\mathbf{S}$. He wrote to Jordan:

> Initially I was very much ashamed of the "computational error" in my theory of the solution, and I wrote about it to Born—especially, since it was no computational error, but only a wrongly-guessed formula. Meanwhile, however, I have thought it over, and found that my carelessness had not led me astray at all this time, and that

[117] In classical theory the first three equations of the perturbation scheme in the case of a periodic system having one degree of freedom and a Hamiltonian H $(= H_0 + \lambda H_1 + \lambda^2 H_2 + \ldots)$ were (see Born and Pauli, 1922, p. 140, Eqs. (7_0)–(7_2)):

$$H_0 = W_0,$$

$$\frac{\partial H_0}{\partial J}\frac{\partial S_1}{\partial w^0} + H_1 = W_1,$$

$$\frac{\partial H_0}{\partial J}\frac{\partial S_2}{\partial w^0} + \tfrac{1}{2}\frac{\partial^2 H}{\partial J^2}\left(\frac{\partial S_1}{\partial w^0}\right)^2 + \frac{\partial H_1}{\partial J}\frac{\partial S_1}{\partial w^0} + H_2 = W_2,$$

where S_1 and S_2 are the terms of the action function S $(= Jw^0 + \lambda S_1 + \lambda^2 S_2 + \ldots)$, w^0 is the angle variable (in the zeroth order), and J is the action variable of the system.

[118] In his recollections, Born remarked: 'One point sticks in my memory which illustrates the fact that Heisenberg was at that time not familiar with matrix calculus. He sent us a suggestion how to treat perturbations which I recognized was wrong. The correct methods followed easily from the theory of canonical transformations which Jordan and I had meanwhile developed; Heisenberg at once accepted our method' (Born, 1978, p. 220).

> my formulae were as correct as Born's (only in Born's case the justification is much nicer and more transparent, while in my case one must prove something "easily" by complete induction). (Heisenberg to Jordan, 29 September 1925)

Indeed, if one wrote the expansion of the transformation matrix $\mathbf{S}$ as

$$\mathbf{S} = 1 + \lambda \mathbf{S}_1 + \lambda^2\left(\tfrac{1}{2}\mathbf{S}_1^2 + \mathbf{S}_2\right) + \ldots, \tag{136}$$

then one arrived at Heisenberg's perturbation scheme, Eqs. (132).[119] However, he frankly admitted that 'Born's form for the canonical transformation is almost the high point of the present theory' (Heisenberg to Jordan, 29 September 1925).

The definition of the transformations, such as those contained in Eq. (133), as 'canonical' transformations in matrix mechanics, appears to be rather bold. In classical mechanics, canonical transformations of the dynamical variables are determined by the fact that they leave the (canonical) Hamiltonian equations of motion invariant; that is, these equations hold for all dynamical variables which are related through canonical transformations. If one tried to make use of this definition in quantum mechanics, one had to take into account the fact that the corresponding matrix equations, Eqs. (90), involved matrix differentiation (i.e., the differentiation of matrices with respect to matrices). Hence it was not clear from the very beginning whether—especially for nonseparable Hamiltonians—these equations remained invariant under the transformation of the variables $\mathbf{p}$ and $\mathbf{q}$ into $\mathbf{P}$ and $\mathbf{Q}$, with

$$\mathbf{P} = \mathbf{S}\mathbf{p}\mathbf{S}^{-1} \qquad \text{and} \qquad \mathbf{Q} = \mathbf{S}\mathbf{q}\mathbf{S}^{-1}. \tag{137}$$

However, if the differential quotients, $\partial \mathbf{H}/\partial \mathbf{p}$ and $\partial \mathbf{H}/\partial \mathbf{q}$, could be replaced by algebraic expressions—and this was always the case when one used Heisenberg's definition of the matrix differentiation, contained in Eq. (123)—then the transformation (137) indeed represented a canonical transformation. By the end of September 1925, Born, Heisenberg and Jordan agreed to use the algebraic form, Eqs. (126), of the Hamiltonian equations in matrix mechanics. Hence they were perfectly allowed to consider the transformations (137) as canonical transformations.

In this context another question arose: namely, whether the transformations (137) represented the *most general* canonical transformations. The authors approached this question in a slightly different form. They noticed that the transformations (137) left the quantum condition invariant; that is, Eq. (67) remained the same for all sets of variables $\mathbf{p}$ and $\mathbf{q}$, related by Eqs. (137). Since the quantum conditions were supposed to play the crucial role in the theory,

[119] Jordan later recalled the following about the discussion concerning perturbation theory:

> Heisenberg's original perturbation formula did not yield—in contrast to Born's, which was based on the canonical transformation $\mathbf{P} = \mathbf{S}\mathbf{p}\mathbf{S}^{-1}$—a procedure which would allow one to find the full power series in the perturbation parameter. Heisenberg's formula gave only the first and second terms of this power series. While Born first believed that Heisenberg's formula for the second [-order] term was incorrect, since it did not agree with the corresponding term of his [Born's] formula, Heisenberg discovered later that the matrix which he denoted as $\mathbf{S}$ was the logarithm of Born's quantity $\mathbf{S}$. In this sense Heisenberg's formula still provided the correct second-order approximation. (Jordan to van der Waerden, 23 July 1962)

Born, Heisenberg and Jordan stated a new definition of the canonical transformations.[120] They said: 'By a "canonical transformation" of the variables p, q, into new variables P, Q, we understand a transformation in which $\mathsf{pq} - \mathsf{qp} = \mathsf{PQ} - \mathsf{QP} = (h/2\pi i)[1]$; for then the same canonical equations apply to P, Q as to p, q' (BHJ, 1926, p. 565). However, with this definition also one could not answer the question concerning the most general form of a canonical transformation. 'We would surmise,' the authors stated, 'that Eq. [(137)] represents in fact the *most general* canonical transformation' (BHJ, 1926, p. 565).

In spite of this uncertainty, the three collaborators accepted and further developed the Born–Jordan transformation theory and the perturbation scheme derived from it. They used it to formulate, especially, the matrix mechanics of systems having several degrees of freedom, both degenerate and nondegenerate. Thus, they provided, at least in principle, the tools to solve the problem of integrating the canonical equations of motion for an arbitrary quantum-mechanical system, which they had recognized to be equivalent to the problem of diagonalizing the Hamiltonian matrix by introducing canonically transformed variables.

III.3 Several Degrees of Freedom and Degeneracy

In their initial formulation of matrix mechanics Born and Jordan had restricted themselves to systems having only one degree of freedom; they expected, in due course, to generalize the theory to systems having an arbitrary number of degrees of freedom that normally occur in atomic physics. The first step towards such a generalization was taken up by Heisenberg in Copenhagen during the second half of September 1925, at a time when his collaborators in Göttingen were still concerned with writing up their paper (Born and Jordan, 1925b). In a postcard to Jordan he stated the quantum conditions, Eqs. (121), for systems having several degrees of freedom. Although Born and Jordan had also discussed the possibility that similar conditions might apply to these systems, they had not succeeded in deriving them from first principles in matrix theory—as they had done with the quantum condition, Eq. (67)—and therefore remained in doubt about the entire question.[121]

[120] The definition of canonical transformations with the help of the quantization condition first occurred in Heisenberg's letters. (See, especially, Heisenberg to Born, 5 October 1925.)

[121] In his postcard, after giving the formulae (121), Heisenberg remarked: 'So far as I recall, you wrote similar conditions in a letter to me, but later you raised objections' (Heisenberg to Jordan, 21 September 1925).

As we have mentioned in footnote 110 Hermann Weyl, independently of Heisenberg, arrived at the same quantum conditions (Weyl to Born, 27 September 1925). Moreover, as is evident from a letter of Heisenberg to Pauli, dated 24 September 1925, Pauli also obtained them at the same time. There Heisenberg wrote: 'Your conjecture about the quantum conditions for [systems with] several degrees of freedom is right, of course.' Finally, Paul Dirac in Cambridge, who worked out the consequences from Heisenberg's quantum-mechanical ideas independently of the Göttingen physicists, presented Eqs. (121) in his symbolic language in his paper on the fundamental equations of quantum mechanics (Dirac, 1925d, p. 648, Eqs. (10)).

However, their doubts and objections did not convince Heisenberg. 'Your reservations concerning the quantum conditions,' he wrote to Jordan, 'do not appear to me really convincing, though I must also admit that I have no compelling reasons for the form [of Eqs. (121)], except only those of formal beauty' (Heisenberg to Jordan, 29 September 1925). And he added the remark: 'Compare, for example, the bracket symbols (the Jacobians [Heisenberg wrote "*die Jakobische*"]) of classical mechanics.' Though with this remark he had hit upon an interesting idea, which was used later in a modified form by Paul Dirac to justify the quantum conditions (121) for systems having several degrees of freedom, he did not elaborate on it further. Heisenberg suggested that for the moment one should not pay particular attention to the question whether or not one could *derive them rigorously* from the equations of motion, Eqs. (126). This point of view was then adopted more or less by all collaborators; they adopted Eqs. (121) finally as *the* quantum conditions for systems having several degrees of freedom. In fact they changed their name to 'commutation relations' ('*Vertauschungsrelationen*') in their paper (BHJ, 1926, p. 573).[122] They also agreed to take Eqs. (126) in all cases, i.e., for separable *and* nonseparable Hamiltonian matrices, as *the* canonical equations of motion. With respect to the question of deriving the commutation relations they remarked:

> Equations [(121)] would, together with Eqs. [(126)], obviously entail too extensive a set of requirements for the $\mathbf{p}_k, \mathbf{q}_k$, if all these equations were *independent* of one another. As an interesting mathematical problem must rank the derivation of Eqs. [(121)] using the minimum number of independent and mutually consistent assumptions; nevertheless, this question will not be handled here. We shall content ourselves with mentioning that $d\sum_k(\mathbf{p}_k\mathbf{q}_k - \mathbf{q}_k\mathbf{p}_k)/dt = 0$ is a general outcome of the equations of motion [(126)]. On the other hand, it will be shown that in general the equations [(121)] together with the equations of motion [(126)] . . . *can* be satisfied (apart, of course, from singular exceptions). (BHJ, 1926, p. 575)[123]

[122] The name 'commutation relations,' which occurred here for the first time, would later replace use of the earlier notation (quantum conditions) completely. The latter referred, of course, to the particular quantum orbits of the old atomic theory. Actually it would turn out that Eqs. (121) are also compatible with motions which were earlier considered to be not quantized. Hence the new name appeared to be more accurate.

The word 'commutation' stemmed from the mathematical language. Hermann Weyl, for example, spoke about the 'commutator $\mathbf{PQP}^{-1}\mathbf{Q}^{-1}$ of the two infinitesimal transformations $\mathbf{P} = 1 + \delta\mathbf{p}$, $\mathbf{Q} = 1 + \epsilon\mathbf{q}$,' which he found to be equal to $1 + \delta\epsilon[(\mathbf{pq} - \mathbf{qp}) + \ldots]$ (Weyl to Born, 27 September 1925).

Paul Dirac also introduced the word 'commute.' As he recalled:

> Another question of notation—it often happens that $uv = vu$ is a special case. When that is so, the mathematicians who had been handling noncommutative algebra said that u permutes with v. It seemed to me that the word "permute" was not really very appropriate. One thinks of permutations as rearranging the order of several quantities, and here we are concerned only with two quantities. So I invented the word "commute." I do not think it had been previously used in mathematics. I said that, when $uv = vu$, u and v commute with each other. That again is a notation which everyone has accepted since then. (Dirac, 1977, p. 129)

[123] From Born's mathematical education in David Hilbert's school it would seem to be quite natural that he attributed an important role to the question of establishing a theory on a minimum number of compatible assumptions or postulates. Heisenberg, on the other hand, did not care for

To prove the compatibility of the commutation relations and the equations of motion, the authors had to generalize perturbation theory to quantum-mechanical systems with several degrees of freedom.

In the formal description of such systems Born, Heisenberg and Jordan again followed the analogy with the previous semiclassical description. In the latter, a multiply periodic system having f degrees of freedom depended on f 'quantum numbers,' $n_1, \ldots, n_f$. In the matrix theory discussed so far a quantum number of the old theory became associated with a pair of indices denoting the rows and columns of infinite quadratic matrices. When one treats systems of f degrees of freedom according to the same prescription, one has to use matrices characterized by f pairs of indices. In their paper the authors therefore represented dynamical variables in terms of '$2f$-dimensional matrices,' writing the position and momentum matrices as

$$\mathsf{q}_k = (q_k(n_1, \ldots, n_f; m_1, \ldots, m_f)) \tag{138a}$$

and

$$\mathsf{p}_k = (p_k(n_1, \ldots, n_f; m_1, \ldots, m_f)), \tag{138b}$$

respectively, $k = 1, \ldots, f$ (BHJ, 1926, p. 572). The significance of Eqs. (138) is that, instead of two-dimensional arrays which describe periodic quantum systems of one degree of freedom, one has now to take $2f$-dimensional arrays.[124]

Although this form of visual representation of the matrices for multiply periodic systems may be considered as being suggestive or convenient in many cases, it is by no means the only possible representation of the variables. In fact, the authors noted explicitly:

> Nevertheless this representation, albeit under certain circumstances very convenient and clear, is by no means essential. Even for several degrees of freedom the fundamental dynamical equations assume the form of *matrix equations*, but these matrices can as before also be written in two-dimensional form. It became apparent already for *one* degree of freedom that the *sequence* of the stationary states, as given by the ordering of the matrix rows, is (in contrast to the erstwhile theory) purely fortuitous and is not governed by any intrinsic property of the system. This

---

such matters. As he remarked: 'I don't care how much of it [i.e., the basic assumptions] one can prove or not, if one assumes [the validity of] the equation of motion' (Heisenberg to Jordan, 29 September 1925).

It is interesting to note that Dirac showed a little later that Eqs. (121)—or the corresponding ones in his theory—followed from a reformulation of the classical Poisson brackets; hence they were naturally consistent with the equations of motion in quantum mechanics (Dirac, 1925d).

[124]Born was accustomed to the representation of systems having several degrees of freedom by multidimensional mathematical objects already from his investigation of three-dimensional lattices of atoms. (See, e.g., Born & von Kármán, 1912.) Note that in Eqs. (138a, b) the matrix elements on the right-hand side consist of amplitudes and periodic factors in time; usually Born, Heisenberg and Jordan denoted only the amplitudes as '$q_k(n_1, \ldots, n_f; m_1, \ldots, m_f)$.'

> observation can now be directly transferred to many-dimensional matrices as well; one can carry out any arbitrary rearrangements and, in particular, transform the $2f$-dimensional matrices into two-dimensional ones. The reason is that the basic definitions of addition and multiplication, as also of differentiation with respect to time, are evidently totally independent of any *ordering relations* existing between the basis systems of indices $n_1, n_2, \ldots, n_f$, which, if taken *singly*, specify the *states* and, if taken *in pairs*, specify the *transitions.* (BHJ, 1926, p. 573)

The question as to which of the two equivalent representations of quantities (accounting for the properties of periodic systems having f degrees of freedom) to use—the one with $2f$-dimensional matrices or the one with two-dimensional matrices—was then determined only by practical considerations. In the applications made by Born, Heisenberg and Jordan in the three-man paper, no preference was given to one or the other; both were used.

The fundamental problem in quantum mechanics, namely, the determination of the energy states of a Hamiltonian matrix H describing a system of f (> 1) degrees of freedom, could now be formulated in the same way as for systems with one degree of freedom. Again the authors called those transformations 'canonical' that leave the commutation relations (121) unaltered; that is, if two sets of variables, $\mathsf{p}_k, \mathsf{q}_k$, and $\mathsf{P}_k, \mathsf{Q}_k$, are connected by transformations of the type given in Eqs. (137) involving a transformation matrix, S, and its inverse, S^{-1}—both of which are either $2f$- or 2-dimensional in agreement with the representation chosen for the dynamical variables—then the same commutation rules are satisfied by $\mathsf{p}_k, \mathsf{q}_k$ and $\mathsf{P}_k, \mathsf{Q}_k$. The diagonalization of the Hamiltonian may be obtained in this case by a canonical transformation of the original variables, say p_k^0 and q_k^0, in the matrix H, according to Eq. (133) generalized to periodic systems of f degrees of freedom.

Similarly Born, Heisenberg and Jordan began to develop a perturbation scheme for multiply periodic systems, which would be applicable, provided their Hamiltonian could be expanded in a series of terms multiplied by increasing powers of a small perturbation parameter λ. They assumed that in the zeroth-order approximation the system consists of f uncoupled systems, each having one degree of freedom, described by the diagonalized Hamiltonians, $\mathsf{H}^{(k)}(\mathsf{p}_k^0, \mathsf{q}_k^0)$, for zeroth-order variables, $\mathsf{p}_k^0, \mathsf{q}_k^0$, $k = 1, \ldots, f$. Hence the zeroth-order energy matrix, W_0, is a direct sum of these individual Hamiltonians, and the perturbation terms give rise to the coupling. In order to diagonalize the total Hamiltonian of the system one can then formally use the same scheme as in the case of one degree of freedom, i.e., Eqs. (135), in which the appropriate matrices for systems of f degrees of freedom are inserted.[125] Thus, the expectation, which Heisenberg had already expressed in his postcard to Jordan on 21 September 1925, that the perturbation scheme for systems of one degree of freedom could be extended to

[125] For example, the dynamical variables, q_k^0 and p_k^0, which diagonalize the Hamiltonian, assume the form given by Eqs. (138); the elements $q_k^0(n_1, \ldots, n_f; m_1, \ldots, m_f)$ are identical with $q_k^0(n_k, m_k)$, the matrix elements of the kth uncoupled system, if for all $j \neq k$ the indices m_j and n_j are identical and zero otherwise.

systems with several degrees of freedom, was fully confirmed, as long as one considered nondegenerate systems.

Everybody, who was familiar with the problems of atomic theory, expected that the so-called degenerate systems had to be treated with special care. As 'degenerate' systems one had denoted those systems which, at least in some approximation, exhibited several identical energy states. Only by the action of a small perturbation, usually provided by a not-too-strong external electric or magnetic field (as in the case of the Stark and Zeeman effects of the hydrogen atom) or by the mutual interaction of electrons (as in the case of many-electron atoms), did each degenerate energy state of the system under investigation split into a pattern of slightly shifted terms. In principle, also in systems of one degree of freedom (with an appropriate potential function) can the energy states become identical, but in the usual situations that occurred in atomic physics the degeneracy was, as a rule, associated exclusively with systems having several degrees of freedom. Among the latter it was easy to find degenerate ones. One had just to assume that in the zeroth-order approximation two or more of the decoupled systems were identical: such as the two electrons moving around a helium nucleus, which represents the zeroth-order approximation to the helium problem. Hence Born, Heisenberg and Jordan had to pay particular attention to studying perturbation theory in the degenerate cases, for they quickly realized that their perturbation scheme represented by Eqs. (135) broke down in such cases. This result followed immediately from the calculation of the elements of the transformation matrix $\mathbf{S}$ according to Eqs. (135); for $S_r(nm)$, the elements of the rth approximation, one obtained

$$S_r(nm) = \frac{F_r(nm)}{h\nu_0(nm)}(1 - \delta_{nm}), \qquad r = 1, 2, \ldots, \tag{139}$$

where $F_r(nm)$ are the elements of the matrix function $\mathbf{F}_r$, and $\nu_0(nm)$ is the frequency of transition from the state denoted by the quantum number n to the state denoted by the quantum number m. (For systems with f degrees of freedom both n and m are sets of f quantum numbers, $n_1, \ldots, n_f$, and $m_1, \ldots, m_f$.[126]) These equations became meaningless when the unperturbed system under consideration possessed degenerate eigenvalues because then certain $\nu_0(nm)$ were zero for $n \neq m$. Thus, the problem arose to develop a perturbation scheme for degenerate systems.

In degenerate systems there occurred many characteristic features, and these features were reflected in the matrix treatment. Thus, for example, Born, Heisenberg and Jordan noted the following fact: while one could still derive energy conservation, i.e., the fact that $\dot{\mathbf{H}}$, the time derivative of the Hamiltonian matrix vanished, this did not lead to a diagonal matrix $\mathbf{H}$. Therefore, in this case, they

[126] To derive Eqs. (139) one observes that $\mathbf{S}$, like the classical action function, has no constant term. Hence all diagonal matrix elements $S_r(nn)$ must be zero. On the other hand, the periodic terms of $\mathbf{S}_r\mathbf{H}_0 - \mathbf{H}_0\mathbf{S}_r$ and $-\mathbf{F}_r$ must be identical.

suggested a 'sharpening of the fundamental equations' by requiring that: '*As the basic equations, one should generally be able to choose the commutation relations and the property* $\mathsf{H} = \mathsf{W} =$ *diagonal matrix*' (BHJ, 1926, p. 577). It was only with the latter requirement that the validity of the frequency condition and the uniqueness of the energy for degenerate systems seemed to be established. However, the authors immediately noted that if the energy is uniquely given, the dynamical variables, p_k and q_k, in which the Hamiltonian is diagonal, are not uniquely defined, but that either of the sets, $\mathsf{p}_k, \mathsf{q}_k$, and $\mathsf{p}'_k, \mathsf{q}'_k$, connected by a constant transformation matrix S, would do the same job.[127]

The situation described above also occurred in the case of nondegenerate systems. However, there the transformation matrix S (relating different systems of variables) must be diagonal, with matrix elements $S(nn)$ equal to $\exp(i\phi_n)$, where ϕ_n denotes a constant phase. The matrix elements of two different sets of dynamical variables, $\mathsf{p}_k, \mathsf{q}_k$, and $\mathsf{p}'_k, \mathsf{q}'_k$, which diagonalize a nondegenerate Hamiltonian matrix, were then related to each other by means of phase differences, that is,

$$p'_k(nm) = p_k(nm) \exp\big[\, i(\phi_n - \phi_m) \big] \tag{140a}$$

and

$$q'_k(nm) = q_k(nm) \exp\big[\, i(\phi_n - \phi_m) \big]. \tag{140b}$$

Hence they could not be determined more accurately than up to a phase.[128] For degenerate systems, on the other hand, the situation was quite different. In particular, it was not permitted to infer from the constancy of the transformation matrix S relating two sets of dynamical variables that it is diagonal; hence very different sets of variables could diagonalize the Hamiltonian of a degenerate system. The authors commented on this uncertain situation as follows:

> This indeterminancy seems to lie in the very nature of things. Apparently, degenerate systems possess a *lability* by virtue of which arbitrarily small perturbations can bring about finite changes in coordinates, and this finds its mathematical expression in that in complete absence of perturbations, the solution of the dynamic equations remains partly indeterminate. Naturally, for every single real atom the coordinates which specify the physical properties of the system, in particular the transition probabilities, are always fixed uniquely either by external perturbations or by the previous history of the system. (BHJ, 1926, pp. 578–579)

[127] The proof of the constancy of S is straightforward. If $\mathsf{H}(\mathsf{p}', \mathsf{q}') = \mathsf{SWS}^{-1}$ is identical with W, then W commutes with S, and $\dot{\mathsf{S}}$ becomes zero.

[128] Equation (140a) follows immediately from the equation $p'_k(nm) = p(nm) S_{nn} S_{mm}^{-1}$. Since the matrices p'_k and p_k are assumed to be Hermitian, one obtains $|S_{nn}| = |S_{mm}|$. This result then leads to Eqs. (140), provided one normalizes all dynamical variables.

The statement contained in Eqs. (140) had been anticipated already earlier by Born and Jordan in connection with the treatment of the anharmonic oscillator (Born and Jordan, 1925b, p. 881). It was independently emphasized by Heisenberg for general systems: 'As far as the phases are concerned, it is trivial in any case that a phase constant of the form $\phi_{ik} = \phi_i - \phi_k$ *always* remains arbitrary' (Heisenberg to Jordan, 7 October 1925).

That a corresponding situation existed in the classical dynamics was known to Born, Heisenberg and Jordan. There, if one considered a periodic system of f degrees of freedom described by a Hamiltonian H, such that it could be expanded in terms of a small perturbation parameter λ (i.e., $H = H_0 + \lambda H_1 + \lambda^2 H_2 + \ldots$) and the zeroth-order term did not depend on one of the action variables, say J_f^0, then the unperturbed system was called degenerate; this implied that the frequency ν_f^0, the time derivative of the angle variable w_f, was zero.[129] Evidently the zeroth-order action variable J_f^0 was not uniquely determined; it could, however, be determined by the perturbation in the following way. The original (zeroth-order) action and angle variables of the system, J_k^0 and w_k^0 ($k = 1, \ldots, f$), were transformed to new variables, $\bar{J}_k^0$ and $\bar{w}_k^0$, with the help of a canonical transformation function S_0; this function then satisfied the special Hamilton–Jacobi partial differential equation

$$\int \ldots \int dw_1^0 \ldots dw_{f-1}^0 H_1\left(\frac{\partial S_0}{\partial w_k^0}, w_k^0 \right) = W_1(\bar{J}_1^0, \ldots, \bar{J}_f^0). \tag{141}$$

Equation (141) expresses an averaging procedure over $w_1^0, \ldots, w_{f-1}^0$, the nondegenerate angle variables of the system; the right-hand side, i.e., $W_1(\bar{J}_1^0, \ldots, \bar{J}_f^0)$, is identical with the time average of the first-order perturbation term, $\bar{H}_1$. The transformation provides well-determined new variables only if $W_1(\bar{J}_1^0, \ldots, \bar{J}_f^0)$ does not vanish; if it does one has to go to a higher-order approximation, the necessary requirement being the existence of a nonzero averaged term in the Hamiltonian. Let us assume for a moment that already $\bar{H}_1$ is different from zero; then the canonical transformation mediated by the action function S_0 changes the variables J_f^0 and w_f^0 considerably, but leaves the nondegenerate variables basically unaltered.

With these reflections on the (classical) background, Born, Heisenberg and Jordan tried to use the analogous method for treating degenerate problems in quantum mechanics. Thus, they started from the approximate Hamiltonian matrix, H_0, which possessed one or several identical eigenvalues, $W_0(nn) = W_0(mm)$, for $m \neq n$. Then they introduced a canonical matrix transformation involving a matrix S_0, which they defined by the equation

$$\mathsf{S}_0^{-1} \mathsf{W}_1 \mathsf{S}_0 = \bar{\mathsf{H}}_1, \tag{141'}$$

where W_1 denotes the first-order correction to the energy matrix, and $\bar{\mathsf{H}}_1$ is the matrix containing only the diagonal elements of H_1, the first-order correction to the Hamiltonian matrix of the quantum system under investigation. Equation (141′) evidently represented the quantum-theoretical reformulation of the classical Eq. (141). The analogue of the classical action function, S_0, i.e., the transfor-

[129] The argumentation that follows may be extended to cases in which the Hamiltonian H_0 does not depend on several of the action variables. Born and Pauli had discussed the classical perturbation theory of degenerate systems in detail in 1922 (Born and Pauli, 1922, p. 144, *passim*).

mation matrix $\mathbf{S}_0$, was finite even in the limit of vanishing perturbation ($\lambda \to 0$). Again, as in the classical case, one had to assume, of course, that $\bar{\mathbf{H}}_1$, the time average, was finite.[130] The transformation of the original variables, $\mathbf{p}_k^0$, $\mathbf{q}_k^0$, to new variables, $\bar{\mathbf{p}}_k^0$ ($= \mathbf{S}_0 \mathbf{p}_k^0 \mathbf{S}_0^{-1}$) and $\bar{\mathbf{q}}_k^0$ ($= \mathbf{S}_0 \mathbf{q}_k^0 \mathbf{S}_0^{-1}$), now brought well-determined quantum variables of the problem in the zeroth-order approximation. $\mathbf{H}_0$ was supposed to remain diagonal in the new variables, assuming the same eigenvalues as before.[131] After performing this extra transformation, Born, Heisenberg and Jordan could go on to apply their perturbation scheme, Eqs. (135), for nondegenerate systems. That is, the perturbation scheme for degenerate systems (given in BHJ, 1926, p. 579, Eqs. (26)–(29)) differed from the one for nondegenerate systems only by the additional transformation involving the matrix $\mathbf{S}_0$; hence the total transformation in the degenerate case is composed as a matrix product, $\mathbf{S}_0\mathbf{S}$, with the factors $\mathbf{S}_0$ and $\mathbf{S}$, the latter indicating a transformation matrix for a nondegenerate system.

The procedure for solving the perturbation equations for a degenerate system was found to be analogous to that in the classical case. From the first equation, that is,

$$\mathbf{S}_0 \mathbf{H}_0 \mathbf{S}_0^{-1} = \mathbf{W}_0, \tag{142}$$

the authors concluded that all matrix elements $S_0(nm)$ were zero, except those for which the associated frequencies $\nu_0(nm)$ vanished. As a result the matrix $\mathbf{S}_0$ reduced to blocks of finite submatrices along the diagonal, the elements outside the blocks being zero. The numbers of rows and columns of these (quadratic) submatrices were determined by the degree of degeneracy of the different eigenvalues of the diagonal matrix $\mathbf{W}_0$; that is, for two identical eigenvalues a 2×2 submatrix appeared in $\mathbf{S}_0$. The nonzero elements of $S_0(nm)$ now had to be chosen such that the second equation of the degenerate perturbation scheme was satisfied.[132] As a result of this procedure, however, those elements $S_1(nm)$ of the first-order perturbation matrix $\mathbf{S}_1$ (the matrix $\mathbf{S}$ was assumed to be expanded as $1 + \lambda\mathbf{S}_1 + \lambda^2\mathbf{S}_2 + \ldots$), for which $\nu_0(nm)$ vanished, could not be determined by the first-order perturbation equation for degenerate systems.[133] The authors fixed them only by going over to the second-order approximation in the following way: they took the 'time average' over the third second-order perturbation equation, arriving at the equation

$$\overline{(\mathbf{H}_0\mathbf{S}_1^2 - \mathbf{S}_1\mathbf{H}_0\mathbf{S}_1 + \mathbf{S}_1\mathbf{H}_1 - \mathbf{H}_1\mathbf{S}_1 + \mathbf{H}_2)} = \mathbf{S}_0^{-1}\mathbf{W}_2\mathbf{S}_0, \tag{143}$$

where the bar on top of the bracket denotes the diagonal elements of the

[130] The definition of the transformation matrix $\mathbf{S}_0$ in terms of $\mathbf{H}_1$ (with the help of Eq. (141′)) is straightforward as long as no diagonal elements of $\bar{\mathbf{H}}_1$ are zero. If this is not so, one has to replace $\mathbf{H}_1$ by $\mathbf{H}_2$ and go on to the next approximation. This procedure will only change Eq. (141′).

[131] The proof of this assertion could be given only later on in the mathematical theory of principal-axes transformation of quadratic forms. (See Section III.5.)

[132] Evidently this equation is equivalent to the equation $\mathbf{S}_1\mathbf{H}_0 - \mathbf{H}_0\mathbf{S}_1 + \mathbf{H}_1 = \mathbf{S}_0^{-1}\mathbf{W}_1\mathbf{S}_0$, whose right-hand side is a diagonal matrix, in agreement with Eq. (141′).

[133] This follows of course, by inspecting Eq. (139), which remains unchanged in this case.

corresponding matrix and $\mathbf{W}_2$ on the right-hand side is supposed to be a diagonal matrix, yielding the second-order correction to the energy of the system under investigation. After that they solved the full second-order perturbation equation; that is, they calculated those elements $S_2(nm)$ of $\mathbf{S}_2$, the second-order correction to the transformation matrix $\mathbf{S}$, which lay *outside* the diagonal blocks (i.e., at positions where the matrix $\mathbf{S}_0$ has nonzero elements). Again the other elements of $\mathbf{S}_2$ could be obtained from the next higher-order equation, i.e., by time-averaging over the third-order perturbation equation. The calculation of still higher-order contributions to the transformation matrix could be continued by following the same procedure; i.e., by first evaluating a part of the elements of $\mathbf{S}_3$ by solving the third-order equation, and the rest by solving the time-averaged fourth-order equation, etc.[134]

Having obtained the perturbation scheme for treating degenerate systems, i.e., for systems whose energy matrix in zeroth-order approximation contains identical eigenvalues, Born, Heisenberg and Jordan still had to demonstrate that no contradiction would arise in applying it. 'The difficulty lies in the fact,' they acknowledged, 'that at each order of approximation equations have to be satisfied by matrices which are already fixed to a large extent [e.g., the second-order perturbation equation is used both to determine elements of the matrix $\mathbf{S}_1$ and of $\mathbf{S}_2$ with the same diagonal matrix $\mathbf{W}_2$], so that it is not possible to foresee whether or not these equations will really prove soluble' (BHJ, 1926, p. 580). However, they did not feel too bad about this situation, for the same difficulty existed in classical theory. As in the latter, they did not expect any particular trouble in quantum mechanics, at least not with those systems which became nondegenerate when higher-order perturbations were included because then the calculations could be continued with the perturbation scheme for nondegenerate systems in which no such complications arose.

With their perturbation scheme for degenerate systems the authors had achieved the goal of formulating the fundamental equations of matrix mechanics and the principal methods for solving them, in close parallel to the classical equations. At times they found that the quantum-mechanical equations possessed a simpler structure than the classical ones; the description of quantum systems also turned out to be less complicated. For example, in classical theory *different* perturbation schemes had to be applied to cases involving 'real' degeneracy (see Born and Pauli, 1922) and to those involving 'accidental' degeneracy (Born and Heisenberg, 1923a). In a case of 'real' degeneracy the frequency associated with the degenerate degree of freedom vanished for all values of the action variables, while 'accidental' degeneracy occurred when the associated frequencies were zero only for selected values of the action variables. In quantum theory such fine

[134]It need hardly be stressed how close the analogy is to the classical procedure (see, e.g., Born and Pauli, 1922, Section 2). For example, the fact that the elements outside of the above-mentioned blocks along the diagonal axis are zero in quantum mechanics corresponds to the fact that in classical theory one averages over the coordinates of the nondegenerate variables (see Eq. (141)). The arbitrariness of the elements $S_1(nm)$ in the diagonal blocks in first-order approximation is also reflected in the same step in classical dynamics.

distinctions disappeared, as one could not any longer speak directly about the motion of electrons in atoms but had to derive the existence of degeneracy from the identity of certain energy states. It might be argued that in quantum theory 'accidental' degeneracy should denote the fact that only a *few* energy states are identical, whereas 'real' degeneracy seems to imply that in the entire energy spectrum there are *very many*, *perhaps infinitely many*, sets of identical energy states. However, such a quantitative distinction makes little sense and, therefore, it was clear to Born, Heisenberg and Jordan that 'accidental' and 'real' degeneracies had to be treated by the same matrix perturbation scheme, in complete agreement with the spirit of quantum theory.

One must be careful when speaking about the analogy between the classical and the matrix mechanical perturbation schemes. In contrast to the dynamical variables in classical theory, *the quantum variables hitherto treated did not include what corresponded to action and angle variables*. In particular, it had not yet been possible to describe in matrix mechanics the coordinate corresponding to the classical angle variable w, which increases in time without limit. Thus far all the equations of the new mechanics employed exclusively only the so-called 'libration coordinates,' such as the position variable q ($= q(nm)\exp[2\pi i\nu(nm)t]$) of the harmonic or anharmonic oscillator, which vary in time within fixed limits. This restriction, which had been noted explicitly by Born and Jordan from the very beginning (Born and Jordan, 1925b, p. 859), could not be removed easily; it required nothing less than a genuine extension of matrix methods to include also the description of aperiodic systems. Although Born, Heisenberg and Jordan considered the possibility of introducing quantum-mechanical action and angle variables during their collaboration in fall 1925, they did not reach any definite results, and it was essentially this limitation which prevented them from attacking problems—other than the anharmonic oscillator, for which libration coordinates sufficed to make the Hamiltonian matrix diagonal—such as the calculation of the energy states of the hydrogen or the helium atom.[135] Neglecting this limitation, the three collaborators—among them especially Max Born—went ahead to generalize and complete the formal mathematical structure of quantum mechanics.

III.4 Born's *Idée Fixe* and a Letter to Niels Bohr

In establishing the perturbation scheme in quantum mechanics Born and Jordan had persuaded Heisenberg to adopt the mathematical point of view of their matrix transformation theory. The reason why they succeeded was that Heisen-

[135] In the course of the three-man collaboration, Jordan once proposed to discuss also the angular momentum and its conjugate angle coordinate in matrix mechanics. To this suggestion, Heisenberg replied by stating that he was not yet sure whether or not a new situation occurred in this case (Heisenberg to Jordan, 29 September 1925). In the published paper there was a detailed discussion of angular momentum in matrix mechanics (BHJ, 1926, Chapter 4, Section 1), but nothing about the canonically conjugate angle. Action and angle variables in quantum mechanics were first introduced by Paul Dirac in his first paper on the theory (Dirac, 1925d).

berg, who preferred a more intuitive, physical approach, recognized the superiority of the mathematical methods employed especially by Max Born. He expressed this change in his attitude in a letter to Wolfgang Pauli. After discussing the fundamental problem of quantum mechanics, i.e., the diagonalization of the Hamiltonian matrix, he remarked: 'The problem is (I shall not write this in greater detail) equivalent to a principal-axes transformation of infinitely many variables, i.e., to a system of infinitely many homogeneous linear equations. The energy values W_n are the eigenvalues of this problem; they may be discrete or continuous. The multiplicity of an energy value represents its statistical weight' (Heisenberg to Pauli, 12 October 1925). With these remarks Heisenberg pointed to a new formulation of matrix mechanics, which could be considered as being most general and complete, transcending even Born and Jordan's matrix mechanics in its sophistication. This theory formed part of the mathematical tradition of Göttingen and was now made use of by Max Born.

Just when Heisenberg wrote to Pauli from Copenhagen the lines quoted above, Niels Bohr received a long letter from Born. In this letter Born explained in some detail the fundamental mathematical ideas on quantum mechanics that had taken shape in his mind. 'I am so happy,' he wrote, 'that you like Heisenberg's idea. I am quite sure that it represents a great progress and that the formulation, which Jordan and I have given to it, is perhaps in a certain sense something final, so far as one can say that at all in physics' (Born to Bohr, 10 October 1925). Born now proceeded to give reasons for his conviction to Bohr, who was anything but a formalist and who did not readily accept mathematical arguments. 'For me the possibility of this formulation has a very personal attraction,' wrote Born. 'Since my student days I have have been haunted by an *idée fixe*, namely, that all essential laws of physics must find their adequate formulation as invariants of linear substitutions.' Born had acquired this *idée fixe* about the structure of physical laws under the influence of his teachers in mathematics, and he told the story to Bohr:

> When I began to study in Breslau in 1900, I learned—from the old, little known, but very ingenious and stimulating mathematician [Jakob] Rosanes—analytic geometry immediately in the form of a theory of invariants; and I was introduced by him also to the works of Cayley, Frobenius and others, in which matrix calculus constituted the most important formal tool. Then in 1905 I came to Göttingen under the influence of [Felix] Klein, who saw everything from the point of view of the theory of invariants, and of [David] Hilbert, who had then just seized Fredholm's theory of integral equations and was about to extend it to a theory of the quadratic forms of infinitely many variables. All this strengthened my *idée fixe*. Then, from the side of physics, came relativity theory, which completely fitted into the scheme. (Born to Bohr, 10 October 1925)[136]

[136] Actually, Born had gone to Göttingen in 1904. He had witnessed there not only the development of the theory of integral equations in Hilbert's hands, but also Hermann Minkowski's steps towards a relativistic electrodynamics along the lines which Felix Klein had conceived many years earlier in connection with the theory of invariants in geometry. Thus, Klein had set up the programme of formulating theories in mathematics and physics in the framework of the algebraic theory of invariants, while Hilbert and Minkowski—whose student Born was primarily—had carried out the programme in practice.

In looking back on his student days, Born reminded himself of Hilbert's work on the generalization of the theory of integral equations, in which Hilbert had employed what he called 'Hermitian forms of infinitely many variables' (Hilbert, 1906a). In this work, Hilbert had, in particular, extended to infinite forms a proof given by Charles Hermite that a finite quadratic form $A(x, x^*)$ $(= \sum a_{nm} x_n x_m^*$ with $a_{nm} = a_{mn}^*)$ of the complex variables, $x_1, \ldots, x_k$, can always be reduced by a transformation to new variables, $y_1, \ldots, y_k$, to a sum of squares, $\sum \lambda_n y_n y_n^*$, each term being weighted by a coefficient which turns out to be the root of the characteristic equation associated with the form (Hermite, 1855). Hilbert's theory of the principal-axes transformation, developed further by his students Ernst Hellinger (1907, 1909) and Hermann Weyl (1908), had made a deep impression on Born, and he mentioned the consequences in his letter to Bohr by saying: 'Thus it happened that everything [in physics], which I did myself, e.g., the contributions to the lattice theory of crystals, always ended up in my treating the matter as transformations of quadratic forms or something similar' (Born to Bohr, 10 October 1925). In his work on lattice theory, Born had already explicitly performed a principal-axes transformation of a quadratic form of infinitely many variables (Born and von Kármán, 1912). Thus, it was not surprising that now, in fall 1925, he reminded himself of this experience and tried to apply it to the problems he was treating in quantum mechanics.

In writing to Bohr, however, the thought crossed Born's mind that the knowledge of all this mathematical background alone would only help a little, for he said: 'I am sure that such a one-sided method of viewing things might appear to you as being quite ridiculous, because the formalism is rather secondary for the actual physical relationships. I am, however, aware of the fact that I lack that physical intuition which you and Heisenberg possess; and all that I can contribute to advance the thing consists in putting the physical regularities into a mathematical scheme.' In spite of this confession of his own shortcomings, Born thought it fortunate that he was well prepared to take up Heisenberg's ideas and to formulate them in a proper mathematical language. Thus he wrote to Bohr: 'Therefore I am now so happy that Heisenberg's rules fit exceedingly beautifully into my *idée fixe*, and that one can express quantum mechanics directly in the form of an eigenvalue problem; thereby one can, it seems to me, perceive all theoretical possibilities quite well. I am very curious how far the presently formulated fundamental theorems will hold out; it may well be that new principles will have to be added. But I am firmly convinced that the basic skeleton of the formulation is reasonable' (Born to Bohr, 10 October 1925).

During the next eighteen days, until he left for an extended visit to the United States, Born worked on the application of Hilbert's principal-axes transformation theory to the problems of matrix mechanics, especially the fundamental problem of determining the eigenstates of the energy matrix. Born's discussion of the connection between the theory of Hermitian forms and quantum theory revealed the great beauty of the intimate relationship between physics and mathematics. In it, he first restated the problem of diagonalizing the Hamiltonian matrix as an eigenvalue problem of an infinite quadratic Hermitian form. Then he used his

formulation and the concept of 'eigenforms' ('*Eigenformen*'), developed by Hilbert and Hellinger, to throw new light on the treatment of perturbed systems in matrix mechanics. Born's use of the principal-axes transformation theory for the purpose of solving the quantum-theoretical problem of motion led to an important physical conclusion: in general the infinite matrices, considered in the quantum-mechanical description of physical variables, should possess both discrete and continuous indices and the Hamiltonian matrices, especially, should then have discrete *and* continuous eigenvalues. This result was received with great enthusiasm by Heisenberg and Jordan. They took over without any change Born's contribution to the 'connection [of quantum mechanics] with the theory of the eigenvalues of Hermitian forms' as Chapter 3 of their joint paper.[137]

III.5 The Eigenvalue Problem and the Transformation to Principal Axes

In order to deal in matrix mechanics with the fundamental problem of calculating the energy states of a system of f degrees of freedom, described by the Hamiltonian matrix H, Born, Heisenberg and Jordan had found the following prescription: One may select any special system of matrices p_k^0 and q_k^0 ($k = 1, \ldots, f$), such that the commutation relations (121) are satisfied. The dynamical problem then consists in searching for a transformation to new variables, represented by the matrices p_k and q_k, which obey the same commutation relations and at the same time reduce the Hamiltonian $\mathsf{H}(\mathsf{p}_k, \mathsf{q}_k)$ to a diagonal matrix. It was this particular problem to which Born wanted to apply the mathematical theory of principal-axes transformation of bilinear forms. Thus he remarked: 'The transformation of the matrices can most easily be grasped if one regards them as a system of coefficients of linear transformations or bilinear forms' (BHJ, 1926, Chapter 3, p. 582). With these words he not only referred to the fact that the very origins of matrix theory lay in algebraic equations, but especially to the connection between matrix mechanics and Hilbert's theory of quadratic forms, which he now brought into focus.

The formal relation between the quantum-mechanical problem and the algebraic problem was easily established. With every infinite Hermitian matrix $\mathsf{a} = (a(nm))$ (where the indices n and m assume integral values, starting, say, from zero)—whose transpose, a^T, equals the matrix $\mathsf{a}^* = (a^*(nm))$, in which all the elements of the original matrix are replaced by their complex conjugates—Born associated a quadratic Hermitian form, $A(x, x^*)$,

$$A(x, x^*) = \sum_{n,m=0}^{\infty} a(nm) x_n x_m^*, \tag{144}$$

[137] Born remarked about the writing of Chapter 3 as follows: 'I remember darkly that Heisenberg and Jordan were not interested much in this chapter and left it to me. Heisenberg checked everything, but he did not have a deep interest in it' (Born, Conversations, p. 46).

of an infinite set of complex variables, $x_0, x_1, \ldots$.[138] Evidently, the quadratic Hermitian forms assumed real values. In his theory of integral equations, Hilbert had considered linear transformations of the variables, especially what he called 'orthogonal substitutions' (Hilbert, 1906a, p. 196). His procedure was now taken over by Born, who performed linear transformations of x_n to new variables y_l, defined by the equations

$$x_n = \sum_{l=0}^{\infty} v(ln) y_l, \qquad n = 0, 1, \ldots, \tag{145}$$

involving the transformation matrix $\mathsf{v} = (v(ln))$. With this transformation the quadratic form $A(x, x^*)$, Eq. (144), could be written as

$$A(x, x^*) = B(y, y^*) = \sum b(nm) y_n y_m^*,$$

with

$$b(nm) = \sum_{k,l=0}^{\infty} v(nk) a(kl) v^*(ml). \tag{146}$$

The coefficients of the transformed quadratic form constituted a matrix b, which was obtained from the matrix a through the matrix equation

$$\mathsf{b} = \mathsf{vav}^{*T}. \tag{147}$$

The matrix b again turned out to be Hermitian, as followed at once from its definition and the application of the rule concerning the transposition of a matrix product, that is, $(\mathsf{ab})^T = \mathsf{b}^T\mathsf{a}^T$. The matrix v was first called 'orthogonal' and later 'unitary,' provided the substitution, expressed by Eq. (146), did not change the 'unit Hermitian form' ('*Hermitesche Einheitsform*') $E(x, x^*)$. Evidently, v had then to satisfy the equation

$$\mathsf{vv}^{*T} = 1 \qquad \text{or} \qquad \mathsf{v}^{*T} = \mathsf{v}^{-1}. \tag{148}$$

As a result, Eq. (147) defined a special transformation, a unitary transformation of the matrix a.[139]

[138] For the moment the time-dependent exponentials, $\exp[2\pi i\nu(nm)t]$, are included in the coefficients of the quantum-mechanical matrices.

[139] Hilbert had first defined 'orthogonal transformations' only in connection with quadratic forms possessing real coefficients (Hilbert, 1906a, p. 180, *passim*). Later (in Section XII of the same paper, p. 216, *passim*) he generalized the concept of orthogonal transformations so as to apply also to Hermitian forms with complex coefficients. Later in 1906 he would speak about 'Hermitian orthogonality' in this context (Hilbert, von Neumann and Nordheim, 1928, p. 22). In 1925 Born used the same notation as Hilbert, i.e., he called the transformation with the matrix v, satisfying Eq. (147), an 'orthogonal' transformation.

The notion of a 'unitary substitution' or a 'unitary orthogonal matrix,' on the other hand, had already been used in mathematics two decades earlier (L. Autonne, 1902, p. 108; I. Schur, 1909, p. 489). John von Neumann returned to this concept in 1927 in his first paper on the mathematical

Born now referred to the well-known mathematical result that it is always possible to reduce a Hermitian form of finitely many complex variables by a unitary transformation to a sum of squares, of which each term is multiplied by a real coefficient (Hermite, 1854, 1855). In matrix language this result meant that for any matrix $\mathbf{a}$ there had to exist a 'unitary' (or 'complex orthogonal') transformation involving a matrix $\mathbf{v}$ satisfying Eq. (148), such that

$$\mathbf{vav}^{-1} = \mathbf{W}, \tag{149}$$

where $\mathbf{W}$ is a diagonal matrix with elements $\delta_{nm} W_n$. Hilbert had obtained an analogous result for compact (completely continuous) infinite matrices (Hilbert, 1906a). In agreement with Hilbert's notation, Born called the W_n—obtained from the transformation of an infinite Hermitian form $A(x, x^*)$ to principal axes (the W_n could also be interpreted as the inverse of the lengths of the axes of the associated geometrical figure)—the 'eigenvalues' ('*Eigenwerte*') and their ensemble the 'mathematical spectrum' ('*mathematisches Spektrum*') of $A(x, x^*)$. The spectrum, he noted further, is made up of discrete and continuous parts; he retained Hilbert's names, i.e., '*Punktspektrum*' ('point spectrum') for the discrete eigenvalues W_n and '*Streckenspektrum*' ('continuous spectrum') for those eigenvalues which are not separated but continuously connected (BHJ, 1926, pp. 583–584).[140] Hilbert had proved the existence of a continuous spectrum, in addition to a discrete spectrum, for the bounded quadratic forms, i.e., for forms whose coefficients satisfy the condition given by Eq. (42). (See Hilbert, 1906a, pp. 189–190, Theorem II.)

Hilbert had created his mathematical concept of 'spectrum' in close analogy to the notion of 'term spectrum' in physical spectroscopy. That now proved to

---

foundation of quantum mechanics (von Neumann, 1927a). There he stated: 'What is meant is the complex orthogonality of the transformation matrix,

$$\{\alpha_{\mu\nu}\}: \sum_{\rho=1}^{\kappa} a_{\mu\rho} a_{\nu\rho}^* = \begin{cases} 1 & \text{for } \mu = \nu \\ 0 & \text{for } \mu \neq \nu \end{cases},$$

which leaves the "Hermitian unit form" $\sum_{\rho=1}^{\kappa} x_\rho y_\rho^*$ invariant. In the mathematical literature one uses in general for this [orthogonality property] the name "unitary"' (von Neumann, 1927a, p. 3, footnote 8). Finally, Hermann Weyl, in his *Gruppentheorie und Quantenmechanik*, called those transformations unitary that leave the 'unit Hermitian form' invariant, and added: 'The name "orthogonal" has been used in the physical literature to denote these transformations, but in mathematics it is necessary to have different names for these two different concepts [i.e. orthogonal with respect to real and complex quadratic forms, respectively]' (Weyl, 1928, p. 15; English translation, Dover reprint, p. 16, footnote). This nomenclature was accepted by the physicists; as, for example, by Born and Jordan in their book *Elementare Quantenmechanik* (1930). We shall, therefore, use the term 'unitary transformations' for those transformations that are generated by a matrix $\mathbf{v}$ satisfying Eq. (148).

[140] In comparing Hilbert's use of the term 'eigenvalues' to Born's one should remember that Hilbert had called not W_n, but its inverse, W_n^{-1}, an eigenvalue of the quadratic form $A(x, x^*)$. The reason for the difference had to do with the fact that Hilbert's investigations of infinite matrices originated from his study of linear integral equations in which the eigenvalue parameter λ multiplied the integral term. (See Eq. (35) from which Fredholm and Hilbert started.)

have been a genuinely prophetic insight for the purpose of applying the mathematical theory of infinite quadratic forms to matrix mechanics, especially in connection with the fundamental problem of diagonalizing Hermitian matrices describing quantum systems. Born was immediately able to translate Hilbert's result into the appropriate physical language and conclude that the matrix **S** of Eq. (133), which generates canonical transformations in quantum mechanics, should be a unitary matrix and satisfy the relation (148).

At first sight it appeared that in this manner of transforming **H**, the Hamiltonian matrix of a given quantum system, into a diagonal matrix **W**, little had been gained towards the solution of the dynamical problem beyond what had been achieved already. For example, the uniqueness of the spectrum of eigenvalues for an arbitrary Hamiltonian matrix could be proved also if the transformation matrix **S** was an arbitrary nonsingular matrix having just an inverse.[141] However, the new mathematical representation soon revealed its advantages, especially when one reformulated the problem of diagonalizing the Hamiltonian matrix **H** by the equivalent algebraic problem of solving the infinite set of linear equations,

$$W v_k - \sum_{l=0}^{\infty} H(kl) v_l = 0, \tag{150}$$

for the unknown variables, v_k, $k = 0, 1, 2, \ldots,$ and the unknown parameter W. The $H(kl)$ are to be identified with the elements of the Hamiltonian matrix (including the time-dependent exponentials). If **H** is a finite matrix, the associated set of Eqs. (150) is finite and yields nontrivial solutions (i.e., $v_k \neq 0$) only when the parameter W assumes certain discrete values, W_n, $n = 0, 1, 2, \ldots$; these W_n are given by the roots of the characteristic equation, $\det[H(kl) - W\delta_{kl}] = 0$. Furthermore, it turns out that these roots are identical with the elements of the matrix **W**, obtained by diagonalizing the matrix **H**. Hence the equivalence of the dynamical problem in quantum mechanics and the problem of finding a nontrivial solution of the set of linear equations is established in this case. However, Hilbert had extended the equivalence proof to include a large class of infinite matrices as well.[142] This could be verified most easily by assuming that *all eigenvalues, W_n, of the Hamiltonian matrix under consideration were nondegenerate*. Then, for each W_n, there existed a specific set of variables, $v_k = v_{kn}$

[141] The reason for the identity of the resulting diagonal matrices **W** ($= \mathbf{SHS}^{-1}$) and **W′** ($= \mathbf{THT}^{-1}$) for different unitary transformation matrices **S** and **T** lies in the fact that also the matrix product $\mathbf{M} = \mathbf{TS}^{-1}$ is a transformation matrix. Because of $\mathbf{W}' = \mathbf{MWM}^{-1}$, there follow the equations, $M(nm)(W'_n - W_m) = 0$, where not all elements $M(nm)$ are zero. Thus, for a given eigenvalue W'_n there exists at least one eigenvalue W_m and vice versa. Moreover, the multiplicity of a given eigenvalue, which determines the weight of the energy state of the system, is uniquely defined, independently of how the Hamiltonian matrix is transformed into the diagonal matrix.

[142] Hilbert (1906a) had especially dealt with what he called 'completely continuous' (compact) and 'bounded' matrices (see our discussion in Section I.2). His results were slightly generalized and later reformulated by Ernst Hellinger later (Hellinger, 1909), from where Born took them over in quantum mechanics. (However, the reference to Hellinger's paper in BHJ, 1926, p. 583, footnote 2, is not quite correct.)

($k = 0, 1, 2, \ldots$), which satisfied the equations

$$W_n v_{kn} - \sum_{l=0}^{\infty} H(kl) v_{ln} = 0. \tag{150a}$$

Evidently, the v_{kn} were determined only up to constant factors by Eqs. (150a), and not all of them should be zero. Two different sets, v_{kn} and v_{km}, associated with different energies, W_n and W_m, could be shown to be orthogonal; this followed by applying Eqs. (150) for the state W_n and the complex conjugate equations for the state W_m. By normalizing the nonzero expression $\sum_{k=0}^{\infty} v_{kn} v_{kn}^*$ to unity, Born obtained the orthogonality relations for the sets of v_{kn} ($k, n = 0, 1, 2, \ldots$), that is,

$$\sum_{k=0}^{\infty} v_{kn} v_{km}^* = \begin{cases} 1 & \text{for } n = m \\ 0 & \text{for } n \neq m. \end{cases} \tag{151}$$

Hence he identified the v_{kn} with the elements of the matrix **S**, with the help of which the infinite Hamiltonian matrix **H** (whose elements coincide with the coefficients $H(kl)$ in Eq. (150a)) had to be transformed to a diagonal matrix **W**. (See Eq. (133).[143])

To demonstrate the correctness of this assertion, Born turned to the Hermitian form, $H(x, x^*)$ ($= \sum_{k,l=0}^{\infty} H(kl) x_k x_l^*$), and transformed the variables x_k in it to new variables y_m with the help of the v_{kn}, i.e., $x_k = \sum_{m=0}^{\infty} v_{km} y_m$; thus he obtained the result

$$H(x, x^*) = \sum_{k,l=0}^{\infty} \sum_{m,n=0}^{\infty} H(kl) v_{km} y_m v_{ln}^* y_n = \sum_{m=0}^{\infty} W_m y_m y_m^*. \tag{152}$$

Though the right-hand side of this equation yielded the desired proof, its derivation included an intermediate step which needed justification. Specifically, Born had exchanged the infinite sums over m and n with the infinite sums over k and l. Such a procedure was permitted only if the Hermitian form was 'completely continuous' or 'compact' (see the discussion in Section I.2), which was guaranteed, for instance, when the sum of squares of all coefficients, $\sum_{kl} |H(kl)|^2$, existed. However, the quadratic Hermitian forms arising in matrix mechanics were not compact at all. Hence the result, expressed in Eq. (152), appeared to be very questionable. Still Born, who knew about the difficulty, expected that his result would remain valid for those mechanical systems in which the existence of a discrete spectrum was known empirically; he showed further that if it was so, the initially made assumption of nondegeneracy of the energy states could be dropped.

[143] In Hilbert's theory of quadratic forms (Hilbert, 1906a), the v_{kn} ($k = 0, 1, 2, \ldots$) were the coefficients of the so-called 'linear eigenforms' $L_n(x)$ ($= \sum_{k=0}^{\infty} v_{kn} x_k$). An orthogonal (and complete) system of the latter, denoted by the running index n ($= 0, 1, 2, \ldots$), served to reduce a Hermitian quadratic form to principal axes. (See, e.g., Hellinger, 1909, p. 222, *passim*.) This result was taken over by Born in matrix mechanics.

Forgetting about this admitted lack of mathematical rigour in establishing the relation between the algebraic methods employed in the theory of infinite quadratic Hermitian forms and the fundamental dynamical problem of matrix mechanics, one could profitably use Eqs. (150) or (150a) to understand more clearly the already formulated perturbation scheme of quantum mechanics, especially to clarify certain aspects which had been left unanswered before. For this purpose, Born considered a system described by a Hamiltonian matrix H, which could be expanded in terms of a small parameter λ, and substituted its matrix elements $H(kl)$ $(= \delta_{kl} W_l^{(0)} + \lambda H_1(kl) + \lambda^2 H_2(kl) + \dots)$ into the algebraic equations (150). He then also expanded the parameter W $(= W^{(0)} + \lambda W^{(1)} + \lambda^2 W^{(2)} + \dots)$ and the variables v_k $(= v_k^{(0)} + \lambda v_k^{(1)} + \lambda^2 v_k^{(2)} + \dots)$ in terms of λ and, by considering the terms of zeroth order, first order, second order, respectively, he arrived at the linear, homogeneous equations

$$v_k^{(0)}\left(W^{(0)} - W_k^{(0)}\right) = 0, \tag{153a}$$

$$v_k^{(1)}\left(W^{(0)} - W_k^{(0)}\right) = -v_k^{(0)} W^{(1)} + \sum_{l=0}^{\infty} H_1(kl) v_l^{(0)}, \tag{153b}$$

$$v_k^{(2)}\left(W^{(0)} - W_k^{(0)}\right) = -\left(v_k^{(1)} W^{(1)} + v_k^{(0)} W^{(2)}\right) + \sum_{l=0}^{\infty} \left[H_1(kl) v_l^{(1)} + H_2(kl) v_l^{(0)}\right]. \tag{153c}$$

Evidently, Eqs. (153) were completely equivalent to the first three equations of the matrix perturbation scheme, Eqs. (135), for nondegenerate systems. By solving Eqs. (153) and the following equations expressing higher-order approximations, solutions were obtained for the unknown $v_k^{(0)}$, $v_k^{(1)}$, $v_k^{(2)}$, etc. (namely, the values $v_{kn}^{(0)}$, $v_{kn}^{(1)}$, $v_{kn}^{(2)}$, etc.) and the parameters $W^{(0)}$, $W^{(1)}$, $W^{(2)}$, etc. (namely, the values $W_n^{(0)}$, $W_n^{(1)}$, $W_n^{(2)}$, etc.), which turned out to be identical with the solutions found from Eqs. (135) for the elements of the transition matrix S $(= \mathsf{S}_0 + \lambda \mathsf{S}_1 + \lambda^2 \mathsf{S}_2 + \dots)$ and the elements of the diagonal energy matrix W $(= \mathsf{W}^0 + \lambda \mathsf{W}^{(1)} + \lambda^2 \mathsf{W}^{(2)} + \dots)$.

Thus far, the algebraic perturbation scheme had achieved nothing new; however, its advantage showed itself in the fact that it allowed one to deal with both nondegenerate and degenerate systems, which the scheme of Eqs. (135) could not do. This was explicitly demonstrated by Born in Chapter 3, Section 2, of the three-man paper in the following way. For systems, whose unperturbed Hamiltonian had r-fold degenerate eigenvalues, say $W_n^{(0)} = W_{n+1}^{(0)} = \dots = W_{n+r-1}^{(0)}$, Eq. (153a) yielded the solution $W^{(0)} = W_n^{(0)}$ for the zeroth-order energy. However, for this energy value, r^2 nonzero solutions existed in the same (zeroth) approximation for the coefficients $v_k^{(0)}$, these being $v_{nn}^{(0)}, v_{n,n+1}^{(0)}, \dots, v_{n,n+r-1}^{(0)}; v_{n+1,n}^{(0)}, v_{n+1,n+1}^{(0)}, \dots, v_{n+1,n+r-1}^{(0)}; \dots; v_{n+r-1,n}^{(0)}, v_{n+r-1,n+1}^{(0)}, \dots, v_{n+r-1,n+r-1}^{(0)}$. All the

other solutions, namely, $v_{k,n}^{(0)}, v_{k,n+1}^{(0)}, \ldots, v_{k,n+r-1}^{(0)}$, where $k \neq n, n+1, \ldots, n+r-1$, were zero. This result reflected a general property of Eqs. (150): If an eigenvalue is r-fold degenerate, then r linearly independent solutions exist for the set of unknown variables v_k; this implies that *the statistical weight of an eigenvalue is determined by the algebraic structure of Eqs.* (150), which may be called the eigenvalue equations.

In order to arrive at the explicit solutions, $v_{kn}^{(0)}$, for the above-mentioned variables, one had to go—in agreement with the observations made in Section III.3 concerning the perturbation scheme for degenerate systems—to the first-order equation (153b). After substituting in it for the unperturbed energy, $W^{(0)} = W_n^{(0)}$, the right-hand side became zero; then arose r homogeneous linear equations for r unknown variables $v_k^{(0)}$, which had nontrivial solutions only if the matrix of the coefficients was singular. As a consequence, the determinant of this matrix had to be singular; that is, the condition

$$\det\begin{pmatrix} H_1(nn) - W^{(1)}, & H_1(n, n+1), & \ldots H_1(n, n+r-1) \\ H_1(n+1, n), & H_1(n+1, n+1) - W^{(1)}, & \ldots H_1(n+1, n+r-1) \\ \vdots & & \\ H_1(n+r-1, n), & \cdot & \ldots H_1(n+r-1, n+r-1) - W^{(1)} \end{pmatrix} = 0 \tag{154}$$

had to be satisfied. Evidently, Eq. (154) allowed one to calculate the r characteristic roots or eigenvalues, $W_n^{(1)}, W_{n+1}^{(1)}, \ldots, W_{n+r-1}^{(1)}$, and these values split the r-fold degenerate state with energy $W_n^{(0)}$ into states with energies $W_n^{(0)} + \lambda W_n^{(1)}$, $W_n^{(0)} + \lambda W_{n+1}^{(0)}, \ldots, W_n^{(0)} + \lambda W_{n+r-1}^{(1)}$.

Equation (154) possesses the structure of the 'secular' equation in classical astronomy and leads to similar consequences: for instance, slow ('secular') oscillations appear in the perturbed quantum system with periods—e.g., $(h/\lambda) \cdot (W_n^{(1)} - W_{n+1}^{(1)})^{-1}$—which are large compared to the periods of the unperturbed motion because of the smallness of the perturbation parameter λ.[144] Two situations are then possible: either the r roots of Eq. (154) are different or they are not. In the latter case, one must repeat the whole procedure for obtaining the values of $v_k^{(0)}$ in the next higher approximation, as we have discussed already in Section III.3. Thus, again Born's new algebraic perturbation scheme, Eqs. (153), confirmed the conclusions of the matrix scheme for degenerate systems; in

[144] In classical astronomy the motions of certain planets, which have periods that are approximately multiples of a ground period, experience slow oscillations ('secular' motions). In atomic physics the same phenomena occur, and again the 'secular' frequencies turn out to be many orders of magnitude smaller than those of normal transitions.

Let us recall that Born had come across a similar secular problem earlier in connection with lattice dynamics (Born and von Kármán, 1913b). There he had studied the modes occurring in a three-dimensional lattice; the degeneracy in question had originated because no distinction had been made in the three space directions in the zeroth-order approximation. The degeneracy was removed as soon as anisotropic crystals were considered.

addition it provided elegant and well-established methods for the actual calculation.[145]

The orthogonality relations for the coefficients v_{kn}, Eqs. (151), or the corresponding unitarity condition for the matrix $\mathbf{S}$ establishing a canonical transformation, led to an important consequence concerning degenerate systems. Born considered, in particular, a system having an r-fold degenerate value $W_n^{(0)}$ in zeroth-order approximation and concluded from the unitarity of the $(r \times r)$-submatrix of $\mathbf{S}_0$, discussed earlier, that the sum of absolute squares of r transition amplitudes, $\sum_{m=n}^{n+r-1} |q^{(0)}(mk)|^2$, with $k \neq n, n+1, \ldots, n+r-1$, had to remain invariant, if the canonical transformation generated by the matrix $\mathbf{S}_0$ was applied to the dynamical variables of the system. Since the sum of absolute squares of the transition amplitudes had been found to be closely related to the total intensities of line multiplets—in the approximation in which the emitting atom may be considered as a degenerate system—one could agree with Born's conclusion: 'This result provides a mathematical representation of the so-called spectroscopic stability, which has played an important part in the more recent theories of fine-structure intensities' (BHJ, 1926, p. 590). Heisenberg, who had contributed extensively to these 'more recent theories,' was willing to consider Born's conclusion as confirming the validity of the application of the theory of principal-axes transformation to atomic theory.

Born had another reason to be pleased with *his* reformulation of the quantum-mechanical transformation theory. 'The calculations with matrices,' he said later, 'seemed to me rather artificial and complicated, while the calculations in the algebraic formalism went on in a straightforward manner' (Born, Conversations, p. 48). Born was also glad to have taken further steps in emancipating quantum mechanics from the use of semiclassical concepts; the setting up of the new perturbation scheme indeed confirmed the opinion which he had expressed in the introduction to Chapter 3 of the three-man paper. There he had remarked: 'The treatment of the preceding sections has aimed at solving the basic quantum-theoretical equations in a manner as closely parallel to classical theory as possible. But behind the formalism of this perturbation theory there lurks a very simple, purely algebraic connection and it is well worth bringing it into the limelight' (BHJ, 1926, p. 581). The algebraic connection thus discovered allowed him, among other things, to obtain just *one* set of perturbation equations, Eqs. (153), which was equally applicable to systems having either nondegenerate or degenerate unperturbed Hamiltonians. However, for this reason alone one might not consider it necessary to employ the advanced mathematical tools of Hilbert's theory of the principal-axes transformation of infinite quadratic forms. There was another expectation to which Born addressed himself when he said: 'We shall thus arrive at a new definition of the energy constants ("terms") which also

[145] For example, when all roots, $W_n^{(1)}, \ldots, W_{n+r-1}^{(1)}$, are different, one may conclude that the r sets of solutions, $v_{kn}^{(0)}, v_{k,n+1}^{(0)}, \ldots, v_{k,n+r-1}^{(0)}$, with $k = n, n+1, \ldots, n+r-1$, are linearly independent. Hence they may be arranged in a $(r \times r)$-matrix $\mathbf{S}_0^r$, which appears as a finite submatrix of the zeroth-order infinite transformation matrix $\mathbf{S}_0$ of Eq. (141′).

remains valid in the case of aperiodic motion, i.e., of continuously varying indices' (BHJ, 1926, p. 581). Thus, the tools of higher algebra, developed by David Hilbert and his school, seemed to offer for the first time the possibility of dealing with quantum systems, such as the hydrogen atom, having both discrete and continuous energy terms. Born was convinced that this possibility would lead to fundamental progress in the efforts to formulate the quantum theory of atoms and he proceeded to discuss the necessary mathematical methods and their application to quantum mechanics.[146]

III.6 Continuous Spectra and the Significance of the Transformation Matrix

Around mid-September 1925, on his return trip to Göttingen after the summer vacation in Silvaplana, Born passed through Zurich and reported to Hermann Weyl, then a professor of mathematics at the Swiss Federal Institute of Technology (E.T.H.), about the new developments in quantum theory, especially the success achieved by the introduction of infinite matrices. Weyl's imagination caught fire; soon afterwards he informed Born that 'your *Ansatz* in quantum theory has made an enormous impression on me,' and wished him godspeed in the pursuit of his 'fabulous discovery' ('*fabelhaften Entdeckung*') (Weyl to Born, 27 September 1925).

Weyl's enthusiasm about the use of matrix methods in quantum theory, which led him at once to derive a few results himself, had good reasons.[147] After all, in Göttingen he had participated in the development of the theory of linear integral

[146] Heisenberg explained the reason why the three-man paper contained a repeated discussion of perturbation theory, with matrices *and* with purely algebraic methods, as follows:

> In some way we did not feel that we should do it too systematically; but we also thought that it might be of some help for other physicists to show in one part of the paper how similar it [i.e., the matrix perturbation scheme] is to those older things which one had tried. Then we could show that there is a new technique [the algebraic one] which seems to be more powerful and actually deals with the same problem in another way. So we felt, "Why not give everything in the paper?" We did realize that it was a kind of duplication of things, but . . . we did not know what would be the most convenient. We had the impression that there may be other still more convenient mathematical schemes behind it. [We] felt that there [was] a great transformability in the whole thing, but we were not so far ready to write down a general scheme. The canonical transformations were actually in the paper in some way—in the Born part of it—but not in this very special general way in which one would later use it to go over to the Schrödinger picture. (Heisenberg, Conversations, p. 283)

Born saw the development as follows:

> We were [initially] entirely in Heisenberg's footsteps. He had the idea that one should take matrices, although he did not know that his dynamical quantities were matrices . . . And when one had such a programme of formulating everything in matrix language, it takes some effort to get rid of matrices. Though it seemed quite natural for me to represent perturbation theory in the algebraic way, this was not a particularly new way. (Born, Conversations, p. 48)

[147] For example, in his letter to Born mentioned above, Weyl considered some of the mathematical implications of matrix mechanics. In particular, he proposed the commutation relations, Eqs. (121), for systems having several degrees of freedom. Weyl continued to be interested in the development of matrix mechanics and kept contact with Göttingen by exchanging letters with Pascual Jordan (Weyl to Jordan, 13, 23, and 25 November 1925). He also received the proof-sheets of the papers on matrix mechanics.

equations based on infinite matrices. In his doctoral thesis under Hilbert, Weyl had extended his teacher's methods to deal with integral equations whose kernels can neither be represented by compact nor by bounded infinite matrices (Weyl, 1908). In particular, he had shown that if the difference of two matrices is a compact matrix, then the original matrices have the same 'continuous spectrum.'[148]

The main outcome of the investigations of Hilbert (1906a), Hellinger (1907) and Weyl's own on the spectrum of infinite quadratic matrices or forms dealt with what Weyl had called the 'principal theorem concerning the orthogonal transformations of bounded quadratic forms' (Weyl, 1909, p. 373). This theorem stated that a bounded Hermitian quadratic form, $K(x, x^*)$, can be reduced to the following 'normal' form,

$$K(x, x^*) = \lambda_0 y_0 y_0^* + \lambda_1 y_1 y_1^* + \ldots + \int \lambda d\sigma(\lambda; y\, y^*), \tag{155}$$

by a unique unitary linear transformation of the original (infinitely many) variables $x_0, x_1, x_2, \ldots$ into new variables $y_0, y_1, y_2, \ldots$ (see Hilbert, 1906a, p. 198, Theorem III). The 'normal' form or the principal-axes representation, Eq. (155), contains two terms: first, a discrete sum whose coefficients are called the 'discrete' eigenvalues of the spectrum of $K(x, x^*)$; second, an integral over a continuous interval yielding the 'continuous' spectrum.[149] The absolute value of the discrete spectrum of eigenvalues, λ_α $(\alpha = 0, 1, 2, \ldots)$, is bounded by M, which is the bound of the quadratic form $K(x, x^*)$, defined by an equation of the type of Eq. (42) in Section I.2. The integral representing the continuous spectrum has boundaries lying between $-M$ and $+M$; it involves a 'spectral form,' $\sigma(\lambda; yy^*)$, which increases monotonically from zero, for $\lambda = -M$, to the sum of all $y_n y_n^*$ $(n = 0, 1, 2, \ldots)$ for $\lambda = +M$. In the case of a compact infinite quadratic form all eigenvalues can be shown to belong to the discrete spectrum, while in the most general case of an unbounded infinite quadratic form—that is, one in which the bound M in Eq. (42) is infinitely large—the spectrum of eigenvalues has discrete and continuous parts, both of which may extend from $-\infty$ to $+\infty$.[150]

[148] Weyl had continued to work on problems related to his thesis and proved, for instance, the theorem that the spectrum of a bounded quadratic form remained essentially unaltered if an arbitrary compact form was added to it (Weyl, 1909).

[149] In Hilbert's original paper the eigenvalues λ were the inverse of those defined in Eq. (155).

[150] Hilbert had illustrated the important point as to how a continuous spectrum of eigenvalues can arise in infinite quadratic forms by discussing the special real unbounded quadratic form, $K(x, x) = x_1 x_2 + x_2 x_3 + x_3 x_4 + \ldots$ (Hilbert, 1906a, pp. 207–208). For any finite section of this form—i.e., if one considers, say, only N variables and breaks off the rest—all eigenvalues are discrete. However, in the infinite case, one has to solve a recursive set of homogeneous linear equations of the type of Eqs. (150) in order to perform the transformation to principal axes. This set of equations yields a continuous spectrum for λ, extending from $-\infty$ to -1 and from $+1$ to $+\infty$.

It should be added that Hilbert had proved the fact that a nonbounded quadratic form can be reduced to the 'normal' form, Eq. (155), provided there exists a real number λ' which is not an 'accumulation' point of the eigenvalue spectrum; that is, λ' is neither contained in the continuous spectrum nor does any sequence of discrete eigenvalues converge towards it (Hilbert, 1906a, p. 175, Theorem I). The latter condition is obviously satisfied for the above-mentioned nonbounded form; it would be different for a form exhibiting a continuous spectrum from $-\infty$ to $+\infty$.

Born knew all these results perfectly well and he decided to exploit them in matrix mechanics to treat the fundamental problem of diagonalizing the Hermitian Hamiltonian matrices by means of the principal-axes transformation. In making use of these results, he occasionally assumed that certain theorems, which had been proved by the mathematicians for bounded matrices only, would also hold in the case of the unbounded forms occurring in matrix mechanics.[151] He had no proof for this essential assumption since even the expert mathematicians, such as David Hilbert himself or Richard Courant, did not know. He had just to rely on his own intuition and on the physical facts; these implied that atomic systems possessed discrete *and* continuous energy states; hence he concluded the same to be true in matrix mechanics.[152] He remarked: 'The simultaneous appearance of both continuous and line spectra as solutions of the same equations of motion and the same commutation relations seemed to us to represent a particularly significant feature of the new theory' (BHJ, 1926, p. 590).[153] The question now was how to describe and determine the continuous spectrum of a given Hamiltonian matrix or of the associated Hermitian form.

In classical physics, the continuous spectra of atomic systems had been assumed to be connected with aperiodic motions, and the latter could be treated sometimes by means of Fourier integrals. In the Fourier integral, say for $x(t)$, the

[151] Born referred explicitly to the problem of extending the results of the mathematical theory of bounded forms to nonbounded forms in Chapter 3 of the three-man paper. There he remarked: 'Up till now, the theory of quadratic (or Hermitian) forms of infinitely many variables has been developed mainly for a special class ("bounded" forms) (D. Hilbert, [1912c]; E. Hellinger, [1909]). But here we are concerned just with nonbounded forms. We may nevertheless assume that in the main the rules run likewise' (BHJ, 1926, p. 583, footnote 2).

It is easy to see, for example, that the matrices describing the quantum-mechanical harmonic oscillator are not bounded matrices because the discrete spectrum, $W_n = (n + \frac{1}{2})\, h\omega_0/2\pi$, $n = 0, 1, 2, \ldots,$ is not bounded by any number M.

[152] Pascual Jordan recalled as follows about the mathematical problems involved in the extension of the theory of infinite forms:

> What we needed from mathematics was partially what one really knew about the fundamentals of quadratic forms, etc. What the mathematicians did not know we had to find anyway by ourselves, especially what we wanted for the purposes of quantum mechanics. Born had a reprint of a paper by [Ernst] Hellinger, who, in connection with Hilbert's work, had extended the mathematical tools; he [Born] tried to extract something for us from it. But he did not find much that we could use. What we could use from the paper were results which we knew already anyway, and what went beyond it with respect to fine mathematical details [in Hellinger's paper] did not suit us very much. (Jordan, AHQP Interview, 18 June 1963)

The mathematicians themselves had not yet entered into the problems posed by the physicists in the domain of the new quantum mechanics. However, in winter 1925 Heisenberg gave lectures on the new theory in the mathematical circle of Göttingen, and soon thereafter Hilbert began to take interest in the mathematical foundations of quantum mechanics. Later, John von Neumann developed the general theory of nonbounded matrices (von Neumann, 1927a, b).

[153] The fact that the fundamental equations of quantum mechanics, in particular the quantum condition, did not exclude the treatment of aperiodic systems had occurred to Heisenberg before he learned about the mathematical theory of quadratic forms. (See Heisenberg to Pauli, 18 September 1925.)

Let us also remind ourselves here that David Hilbert, when developing his theory of the eigenvalues of infinite quadratic forms, had had in mind the spectra of chemical atoms, which include both discrete and continuous states. Therefore, bounded quadratic forms had seemed to him especially interesting objects of study.

position coordinate of an aperiodic system,

$$x(t) = \int \phi(\nu) \exp(2\pi i \nu t)\, d\nu, \tag{156}$$

the amplitude $\phi(\nu)$ had to be interpreted as the *amplitude density* in the frequency interval ν and $\nu + d\nu$. Also all the other physical quantities had then to be described by Fourier integrals of the type of Eq. (156) involving amplitude densities in the same sense. Born concluded that:

> We shall have to expect quite similar situations in quantum mechanics. Instead of the quantities $q(kl)$ [i.e., the elements of Hermitian matrices describing the properties of periodic quantum systems] we shall have quantities of the form $q(k, W)\,dW$ or $q(W, W')\,dW\,dW'$, depending on whether one [index W, denoting the energy in the continuous spectrum] or both of the two indices [denoting the energies W and W' in the continuous spectrum] lie in the continuous region. (BHJ, 1926, p. 591)

Born further assumed that, similar to the situation in the classical theory, the energy W in the continuous region will have to be a 'total energy' per interval, dW, 'since the probability for an atom to have an absolutely definite energy W in the continuous region is zero' (BHJ, 1926, p. 591).

Such an extension of the description of aperiodic motion to the quantum case was natural in the light of earlier considerations.[154] Even more important for Born, as he reminded himself, was the fact that 'the mathematical theory of continuous spectra which occur in bounded quadratic forms has, starting from the fundamental investigations of Hilbert, been developed in detail by Hellinger' (BHJ, 1926, p. 590). Ernst Hellinger had been Born's friend since their student days together in Breslau, and he had gone to Göttingen with Born to work on a doctoral thesis under Hilbert. Its results now came in handy for Born in trying to extend the matrix methods to include the treatment of continuous spectra.

In his thesis, Hellinger had discussed the problem of obtaining a complete set of 'orthogonal invariants' of bounded quadratic forms of infinitely many variables (Hellinger, 1907). As his main task, he had turned to the analysis of the continuous eigenvalues given by the integral term on the right-hand side of Eq. (155). For this purpose he had introduced what he called the 'linear differential forms' ('*lineare Differentialformen*'), $d\phi(\lambda, y)$ $(= \sum_n d\phi_n(\lambda) y_n)$, of infinitely many variables y_n (Hellinger, 1907, pp. 21, 74); the coefficients of these differential forms, $d\phi_n(\lambda)$, are supposed to be differentials of functions, $\phi_n(\lambda)$, such that the sum of the squares, $\phi(\lambda) = \sum \phi_n^2(\lambda)$, exists and represents a continuous, monotonic function of λ, the energy eigenvalue in the continuous spectrum. The principal-axes transformation, which gives the continuous part of the eigenvalue spectrum of a quadratic form, $K(x, x)$ $(= \sum_{n,m} K_{nm} x_n x_m)$, is then achieved with the help of

[154] In the case of aperiodic motion, the Fourier integral representation replaced the Fourier series representation, which applied to periodic systems. In their paper, submitted in June 1925, Born and Jordan had discussed this representation in connection with the correspondence principle approach (Born and Jordan, 1925a).

those linear 'differential eigenforms' whose coefficients solve the set of (eigenvalue) equations

$$\lambda d\phi_n(\lambda) - \sum_m K_{nm} d\phi_m(\lambda) = 0 \tag{157}$$

with the eigenvalues λ.[155] Since λ varies continuously within the interval which lies between $-M$ and $+M$, where M is the bound of the quadratic form, continuously many 'differential eigenforms' *appear* to solve Eq. (157). This, however, is really not so, since only those functions, $\phi_n(\lambda)$, which are continuous in λ and the sum of whose squares is a continuous function, $\phi(\lambda)$, are admitted; and of the latter only denumerably, though possibly infinitely, many exist.[156] Hence only denumerably many differential forms, $d\phi^{(\alpha)}(\lambda, y)$ $(\alpha = 1, 2, 3, \ldots)$ satisfy Eqs. (157). Different differential eigenforms of a quadratic form are orthogonal to each other and can be normalized. If one generalizes Hellinger's results to Hermitian forms, the continuous terms on the right-hand side of Eq. (155) assume the form

$$\int d\sigma(\lambda; yy^*) = \sum_{\alpha=1}^{\infty} \int \lambda \sum_n \frac{y_n y_n^* \, d\phi_n^{(\alpha)}(\lambda) \, d\phi_n^{(\alpha)*}(\lambda)}{d\phi^{(\alpha)}(\lambda)}, \tag{158}$$

where $\sum_n d\phi_n^{(\alpha)*}(\lambda) y_n^*$ represents the complex conjugate of the linear differential eigenform $d\phi^{(\alpha)}(\lambda, y)$, and $d\phi^{(\alpha)}(\lambda)$, the differential of the function $\phi^{(\alpha)}(\lambda)$ $(= \sum [\phi_n^{(\alpha)}(\lambda) \cdot \phi_n^{(\alpha)*}(\lambda)])$, expressing the normalization. Evidently, this result states the transformation to principal axes for the continuous part of the eigenvalue spectrum.

In summer 1925 Max Born became fascinated by the possibilities offered by Hellinger's complete orthogonal transformation of infinite bounded quadratic forms. He noticed that, in order to deal with the continuous eigenvalues of a Hermitian matrix, he had only to adopt the procedure of the mathematicians. Thus he assumed that, 'in analogy with the results for bounded forms' (BHJ, 1926, p. 591), one can reduce the Hermitian form $H(x, x^*)$ associated with the Hermitian matrix to the following 'normal' form:

$$\sum_{k,l} H(k,l) x_k x_l^* = \sum_n W_n y_n y_n^* + \int W(\phi) y(\phi) y^*(\phi) \, d\phi. \tag{159}$$

[155] For the case of real quadratic forms, $K(x, x)$, see Hellinger, 1909, p. 242, §5, Eq. (1′).

Evidently Eq. (157) replaces the corresponding equation (see Eq. (150)) for discrete eigenvalues. In order to see the connection one might just consider a finite interval from λ_1 to λ_2 in the continuous spectrum. Then, from Eqs. (157), one obtains the equations $\int_{\lambda_1}^{\lambda_2} \lambda d\phi_n(\lambda) - \sum_m K_{nm} \Delta\phi_m(\lambda) = 0$, where $\Delta\phi_n(\lambda) = \phi_n(\lambda_2) - \phi_n(\lambda_1)$; these have indeed the structure of the discrete Eqs. (150).

[156] If the function $\phi_n(\lambda)$ is analytic in λ, then it may be characterized by the coefficients of the (in general infinite) power series in λ. If the function ϕ_n is not analytic but still continuous, then one of its derivatives (or a denumerable set of its derivatives) possesses a jump or discontinuity; within a finite interval of λ, however, only denumerably many jumps of finite height will occur, hence again only denumerably many differential eigenforms exist.

All he had done so far was to replace the eigenvalue (i.e., spectrum parameter) λ (in Eqs. (155) and (158)) by the energy W (the eigenvalue of $H(x, x^*)$) and to rewrite the expression on the right-hand side of Eq. (158) appropriately. In order to determine the structure of the generalized transformation to principal axes in physics, Born discussed in detail the solution of the eigenvalue equations (150) for continuous values of the energy W, referring at each step to the corresponding mathematical investigation of Hellinger.

Born started out with the observation that two solutions, $v_k(W')$ and $v_k(W'')$, whose energy eigenvalues W' and W'' lie in the continuous spectrum, cannot be normalized similar to the solutions in the discrete case (see Eq. (151)) because the function of two variables $\sum_k v_k(W')v_k(W'')$, having the property

$$\sum_k v_k(W')v_k(W'') = \begin{cases} 1 & \text{for } W' = W'' \\ 0 & \text{otherwise} \end{cases} \tag{160}$$

would be badly irregular, if it exists at all' (BHJ, 1926, p. 592).[157] Hellinger had avoided dealing with this function by introducing the concept of differential forms, and Born now followed his lead. Instead of taking $v_k(W)$ as the solutions of Eqs. (150), he took the differential solutions $v_k(W)\,dW$; from the coefficients $v_k(W)$ he constructed the function $\phi(W)$ according to the equation

$$\sum_k \left| \int v_k(W)\,dW \right|^2 = \phi(W), \tag{161}$$

the so-called 'basis function' ('*Basisfunktion*,' Hellinger, 1909, p. 246), which allows one to normalize the differential solutions. That is, in matrix mechanics the $v_k(W)\,dW$, $k = 0, 1, 2, \ldots,$ with different values of the continuous energy parameter W, act as orthogonal differential solutions of Eqs. (150), normalized by the basis function $\phi(W)$ which corresponds to Hellinger's $\phi(\lambda)$ in Eq. (158). Two solutions, $v_k(W')\,dW'$ and $v_k(W'')\,dW''$, are orthogonal in the sense that for two (finite) intervals, Δ_1 and Δ_2, of the continuous spectrum with Δ_{12} as the common interval (overlap) one finds

$$\int_{\Delta_1} v_k(W')\,dW' \cdot \int_{\Delta_2} v_k^*(W'')\,dW'' = \int_{\Delta_{12}} d\phi(W) = \phi(W^{(2)}) - \phi(W^{(1)}), \tag{162}$$

where $W^{(1)}$ and $W^{(2)}$ are the end points of Δ_{12}.

Born considered the fact that it had been possible to construct the transformation matrix generating the principal-axes transformation as the most important consequence of the mathematical theory of Hermitian forms. He wrote this orthogonal matrix as

$$\mathsf{S} = (v_{kn}, v_k(W)\,dW) \tag{163}$$

[157] In agreement with the mathematical consensus at the time, Born did not consider the object $\sum_k v_k(W')v_k(W'')$ as a function at all. The concept and theory of so-called 'generalized functions' was developed by the mathematicians later, especially after Paul Dirac introduced his delta-function into physics (Dirac, 1927á).

because its elements comprise the totality of discrete values v_{kn} and of values $v_k(W)$ which are discrete in one index (k) and continuous in the other (W). For quantum systems he represented S schematically as

$$\mathsf{S} = \begin{pmatrix} & k\rightarrow & \\ n\downarrow & \cdots & \\ & \cdots & \\ & \cdots & \\ W\downarrow & \| & \end{pmatrix} \tag{163'}$$

because the ordering of the columns and rows in the matrices had been established in such a way that the energy values in the diagonalized Hamiltonian increased with increasing indices, and the continuous spectrum occurred for higher values of the energy than the discrete spectrum. The transformation matrix, therefore, contained two types of rows: first, those with a discrete index, corresponding to eigenforms, $\sum_k v_{kn}x_k$ $(n = 0, 1, 2, \ldots)$, of the discrete spectrum; second, those with a continuous 'index' and labelled by the (continuous) energy eigenvalue W, corresponding to the differential eigenforms, $\sum_k v_k(W)\,dW\,x_k$, of Hellinger.[158] Four sets of equations determine the orthogonality and normalization of the rows of S, namely,

$$\sum_k v_{km}\cdot v_{kn}^* = \delta_{nm}, \tag{164a}$$

$$\sum_k v_{km}\cdot v_k^*(W)\,dW = 0, \tag{164b}$$

$$\sum_k v_k(W)\,dW\cdot v_{kn}^* = 0, \tag{164c}$$

$$\sum_k v_k(W')\,dW'\cdot v_k^*(W'')\,dW'' = d\phi, \tag{164d}$$

while in the normalization equations for the columns (which have always discrete indices) there occurs a sum arising from the discrete terms and an integral from the continuous terms.[159] With the help of this unitary transformation matrix one can indeed introduce new variables into the Hermitian form $H(x, x^*)$, those having discrete indices (y_n) and others having continuous indices ($y(\phi)d\phi$); in

[158] Note that, in the above definition of the transformation matrix, Eq. (163'), Born did not follow the usual definition, for he used the first index k to denote the columns rather than the rows.

[159] The discrete contribution is $\sum v_{kn}v_{ln}^*$, and the continuous contribution is $\int v_k(W)\,dW\cdot v_l^*(W)/d\phi$; both add up to δ_{kl}.

these new variables the Hermitian form then assumes the normal form, given by the right-hand side of Eq. (159).[160]

As Born found in the example of the momentum variable of a system having one degree of freedom, the elements of the matrices describing the variables of a quantum system having discrete and continuous eigenvalues fell into four classes. Even if one started with a matrix $\mathsf{p}^{(0)}$ $(= p^0(kl))$, which possesses only discrete indices $(k, l = 0, 1, 2, \ldots)$, after the transformation by the matrix S—which diagonalizes the Hamiltonian matrix of the system under consideration—the matrix p would contain the following elements: first, amplitudes $p(mn)$ denoting the transitions between discrete states with energies W_m and W_n; second, 'amplitude densities' $p(m, W)\,dW$ denoting the transitions from a discrete state, W_m, to a state in the continuous spectrum whose energy lies between W and $W + dW$; third, the amplitude densities $p(W, n)\,dW$ for transitions from states in the continuous spectrum (energy between W and $W + dW$) to the discrete state W_n; finally, the amplitude densities $p(W', W'')\,dW'\,dW''$ for the transitions from an energy state between W' and $W' + dW'$ in the continuum to another continuum energy state lying between W'' and $W'' + dW''$. Therefore, in matrix mechanics, all matrices corresponding to dynamical variables, $\mathsf{f}(\mathsf{p}, \mathsf{q})$, have the general structure indicated by the following scheme:

$$\mathsf{f} = \left(\begin{array}{ccc} & n \rightarrow & W'' \rightarrow \\ m \downarrow & \cdots & \text{———} \\ & \cdots & \text{———} \\ & \cdots & \text{———} \\ & \cdots & \text{———} \\ W' \downarrow & \| \; \| \; \| & \text{▨} \end{array} \right) , \tag{165}$$

where the dots stand for discrete indices, the lines for one discrete and one continuous index, and the dark submatrix in the lower right corner contains elements with both indices continuous. 'The physical meaning of this scheme is self-evident,' Born concluded. 'There are four types of "transitions" which somehow furnish a simple analogue to the "transitions" postulated hitherto in the theory of the hydrogen atom, viz. (1) from ellipse to ellipse; (2) from ellipse to hyperbola; (3) from hyperbola to ellipse; (4) from hyperbola to hyperbola' (BHJ, 1926, p. 595).

This general result was extremely important for the conceptual structure of quantum theory. While in the earlier atomic theory, in addition to the discrete energy states calculated from quantum conditions, *the existence of the continuous spectrum had to be assumed*, in quantum mechanics the discrete and continuous energy states and the transitions between them could be calculated from the

[160] Evidently, S transforms the variables x_k into discrete variables, y_n $(= \sum_k v_{kn}\, x_k)$, and continuous variables, $y(\phi)d\phi$ $(= \sum_k v_k(W)\,dW\, x_k)$. By replacing the x_k in $H(x, x^*)$ through the new variables and applying the orthogonality relations, Eqs. (164), the result of Eq. (159) is obtained.

same fundamental dynamical equations and commutation relations. By Born's application of the principal-axes transformation, therefore, the formal mathematical structure of matrix mechanics was completed. It was capable of handling all those problems in atomic physics in which *only the libration coordinates appeared*, that is, where the dynamical variables which increased linearly for large times could be avoided.[161] About this, Heisenberg wrote to Pauli: 'We have now also worked out the problem of continuous spectra from the point of view of this mathematical theory. At first it looks very chaotic; later one realizes that the calculations are not worse than in the case of Fourier integrals' (Heisenberg to Pauli, 23 October 1925).

By introducing the mathematical tool of principal-axes transformation into quantum mechanics Born had to answer a question of principle. In spite of the fact that the principal-axes transformation acted only as a means of integrating the equations of motion of a given quantum system, he still had to understand what, after all, did the elements of the transformation matrix **S** represent, for they did occur explicitly in the equations of perturbation theory, as in Eqs. (135) or (153), and had to be calculated. Following the guiding principle applied hitherto in the formulation of the new quantum theory, namely, that only observable quantities should enter into the equations describing physical systems, one of two conclusions had to be adopted: either the principal-axes transformation itself had no physical meaning, hence one ought to eliminate it from the quantum-theoretical equations; or the matrix elements of **S** were observable quantities and their physical significance had to be determined.

Born was inclined to believe that the transformation matrix effecting the principal-axes transformation of the Hermitian form $H(x, x^*)$ possessed a genuine physical meaning, and he set out to ascertain it by considering the formal structure of the eigenvalue equations (150a). In each of these equations, for fixed n, there appear only the elements v_{kn} (with $k = 0, 1, 2, \ldots$), i.e., the elements forming a row of the matrix **S** (see Eq. (163), specialized to the situation that only a discrete spectrum is associated with $H(x, x^*)$). Born now assumed that a column or a row of the transformation matrix **S** must have a well-defined physical meaning, and he mentioned it to Heisenberg, who did not respond positively to this assumption. 'He disliked,' Born recalled later, 'having to take a column or row [of a matrix], that is to consider the quantities v_k ($k = 0, 1, 2, \ldots$) without adding a second index as in v_{kn}. He stuck to the argument that [in quantum mechanics] the absolute square of one matrix element $|v_{kn}|^2$, represents the probability of the transition from a state k to another state n, hence only the matrix elements, v_{kn}, have a physical significance' (Born, Conversations, p. 47). Born, on the other hand, believed that it made sense to talk meaningfully about

[161] The extension of the scheme to include nonlibrating coordinates was not treated by Born, Heisenberg and Jordan in their joint paper, for they explicitly excluded those variables whose analogues in classical theory could not be described by Fourier integrals. They remarked, however, that 'the observable effects of an atom (such as radiation, the force upon another atom, etc.) do not in general belong to this type of function' (BHJ, 1926, p. 595).

the columns of the transformation matrix, but he could not for the moment justify his conviction; therefore he avoided a discussion of this fundamental question in the three-man paper.

However, a trace of the debate between Born and Heisenberg may still be found in the joint paper. In discussing the meaning of the differential $d\phi$ entering Eq. (159), the authors stated: 'Thence $d\phi$ essentially represents the number of atoms having an energy which lies between W and $W + dW$, or the *a priori* probability that the energy of the atom lies between W and $W + dW$' (BHJ, 1926, p. 594). Since $d\phi$, according to Eq. (164d), denotes the absolute square of a 'continuous' row of the transformation matrix **S**, Eq. (163), the physical significance of this quantity had been found. Similary, the sum of the absolute squares of the discrete quantities, $\sum_k v_{kn} v_{kn}^*$, was interpreted as the probability that the discrete state, denoted by n, has the energy W_n. In October 1925 Born did not have the time to investigate the problem beyond these statements. But he continued to pursue his interest in finding the physical meaning of the transformation matrix, especially of its rows and columns. Many years later he recalled: 'This thought occupied my mind all the time, and that led me much later to the statistical interpretation [of Schrödinger's wave function]. At that time [in fall 1925] I could do nothing with the discrete method [i.e., matrix mechanics], but when Schrödinger came I saw that the same question could be asked in the wave theory' (Born, Conversations, p. 44). Thus, Born's early involvement in this question of principle prepared him to make a fundamental discovery: the statistical interpretation of the wave function.[162]

[162] Born had a further reason to study the meaning of the transformation matrix. According to Heisenberg's investigation, the squares of the transition elements—such as the square of the elements of the position matrix, $|q(nm)|^2$—yielded the transition probability. Born, however, argued that this was not really so because in dispersion theory the important quantity was rather the dipole moment; hence he asked the question: 'Why not take the coefficients of the momentum matrix or those of any matrix function $\mathbf{f}(\mathbf{p}, \mathbf{q})$?' (Born, Conversations, p. 45). In fall 1925 he was already looking for some general principle that would allow one to calculate transition probabilities.

Incidentally, many years later Heisenberg used another transformation matrix, denoted by the same letter **S**, in the theory of elementary particles (Heisenberg, 1943a, b; 1944). By definition, this matrix involved only observable quantities. Perhaps Heisenberg wanted to make up for his old mistake of 1925 in having misjudged the importance of the quantum-mechanical transformation matrix.

Chapter IV
The Success of Matrix Mechanics

After the theoretical scheme of quantum mechanics, including the matrix and the equivalent algebraic formulations, had been established by Born, Heisenberg and Jordan, it became necessary to deal immediately with two tasks: first, with tests to confirm the consistency of the new theory, especially by comparing its results with those which were considered reliable—such as the results of dispersion theory and radiation theory; second, to extend the application of the new theory beyond the simple examples which it had been possible to treat thus far. In both tasks the Göttingen theoreticians received the help of other physicists who had had close relations with them and had become familiar with their matrix methods and early results prior to publication. It had thus been possible for these physicists to work on quantum mechanics already, say, from October 1925, several weeks before the three-man paper was submitted or the first paper of Born and Jordan (1925b) appeared in print. In checking the consistency of the new theoretical scheme, Hendrik Kramers made an interesting interpretation of the quantum conditions or commutation relations, while Wolfgang Pauli calculated the energy states of the hydrogen atom by employing exclusively matrix methods.

Pauli's work on the hydrogen atom must be regarded as a real triumph of the matrix methods, especially because he bypassed a major difficulty of the available quantum-mechanical formalism: the obvious impossibility of describing the action and angle variables. Ever since the days of Hamilton and Jacobi, these variables had provided the key to the solution of multiply periodic systems; and to avoid their use seemed to imply that one could not transfer the methods of classical dynamics, including the perturbation scheme, to calculate the energy states of real atoms. Pauli, however, discovered a clever method to solve the hydrogen problem without invoking angle variables. He was thus able to diagonalize the Hamiltonian matrix for one-electron atoms and, in addition, to determine their behaviour in constant external electric and magnetic fields in perfect agreement with the experimental information. 'It was stupid,' Born confessed with some envy about Pauli's success, 'that we could not solve the hydrogen atom; it was straightforward' (Born, Conversations, p. 50). However, the procedure, which worked in the case of the hydrogen atom, could not be extended to two- and more-electron atoms. These problems had to wait until an action-angle formalism became available in quantum mechanics, that is, until the matrix scheme was extended to deal with nonlibration variables as well. Before

that generalization it was only possible to treat a very selected class of atomic problems, such as the rotational motion of diatomic molecules.

IV.1 The Treatment of Dispersion Phenomena

The first 'outside' contribution to matrix mechanics came from Hendrik Kramers. He had, however, been familiar with the results of this theory for some time. On one hand, he had visited Göttingen in June 1925 and had learned from Heisenberg about his efforts to revise the fundamental concepts of quantum theory; on the other, he had met Heisenberg again in Copenhagen the following September, while the latter was working on matrix theory in collaboration with Born and Jordan. Finally, Kramers spent a few days in November 1925 with Pauli in Hamburg (see Heisenberg to Pauli, 3 November 1925). Shortly afterwards he submitted a note to *Physica*, the Dutch journal, in which he made 'a few comments on the new quantum mechanics of Heisenberg' ('*Eenige Opmerkingen over de Quantummechanica van Heisenberg*,' Kramers, 1925c).[163] In this note he discussed in particular a specific connection between the dispersion formulae (Kramers, 1924a; Kramers and Heisenberg, 1925) and the quantum conditions, Eqs. (121).[164]

After introducing the matrix notation, Kramers started by rewriting the dispersion formula for the electric moment of an atomic system induced by an incident electromagnetic radiation having a frequency ν and electric field strengths $E^{(1)}$, $E^{(2)}$, $E^{(3)}$, in the x-, y-, and z-directions, respectively. He denoted the Cartesian coordinates, describing 'symbolically the motion of an electron in an atomic system' (Kramers, 1925c, p. 370), as $\mathbf{q}^{(k)} = (q^{(k)}(mn) = (1/2e)A^k_{m,n}\exp[2\pi i\nu(mn)t])$, where $A^k_{m,n}$ is the amplitude in the k (= 1, 2, 3 or x, y, z)-direction of the radiation of frequency $\nu(nm)$, emitted by the atom during

[163] It is of interest to note that Kramers, in the beginning at least, had been skeptical about the prospects of Heisenberg's endeavours. (In a postcard to Pauli, dated 29 June 1925, Heisenberg mentioned that he was increasingly becoming more convinced of his new ideas, and Kramers therefore accused him of being too optimistic.) Later, after Heisenberg had completed his paper (1925c), Pauli tried to convince Kramers of the importance of the ideas contained in it. He wrote: 'The "community of true believers" [in Bohr's faith—meaning, principally, Kramers] would neither gain much honour nor much success by attempting to fight the trend of the present development in quantum theory, which is directed towards an analysis of the concepts of motion and force. For it seems to me that there is great hope of moving forward *positively* along this path. In particular, I have greeted with jubilation Heisenberg's bold proposals (about which you have certainly heard in Göttingen)' (Pauli to Kramers, 27 July 1925). Then, while Heisenberg stayed in Copenhagen in late summer 1925 before returning to Göttingen, Kramers learned from him about the matrix formulation of quantum mechanics. He acknowledged this information in a footnote in his paper by stating: 'I want to express my sincere thanks to Mr. Heisenberg for his kind information about the recent development of his theory by Born, Jordan and himself' (Kramers, 1925c, p. 369, footnote 1).

[164] Kramers' paper (1925c) did not constitute his earliest positive response to Heisenberg's work. Earlier, in his presentation of dispersion theory to the sixth Scandinavian Mathematical Congress in Copenhagen in August 1925, he had mentioned the 'interesting and promising results' which had 'already been obtained in the papers [of Born (1924b) and Heisenberg (1925c)] . . ., which contain an audacious attempt of developing a new "*quantum mechanics*"' (Kramers, 1925b, p. 153).

the transition from the mth to the nth state. And for the matrix element, $e\Delta q^{(k)}(mn)$, of the induced electric moment he wrote the equation (see Kramers and Heisenberg, 1925, p. 697, Eq. (37), or our Volume 2, Eq. (56))

$$e\Delta q^{(k)}(mn) = \frac{e^2}{h}\,\mathrm{Re}\left\{\sum_{l=1}^{3} E^{(l)} \sum_{r}\left[\frac{q^{(k)}(mr)q^{(l)}(rn)}{\nu(rn)+\nu} - \frac{q^{(l)}(mr)q^{(k)}(rn)}{\nu(mr)+\nu}\right]\exp(2\pi i\nu t)\right\}, \tag{166}$$

where $-e$ is the charge of the electron and Re{ } stands for the real part of the expression within the curly brackets.[165] This formula accounts for the behaviour of an electron bound by forces within an atom in such a way that discrete energy states, W_m, W_n and W_r, exist with the associated transition frequencies, $\nu(mr) = (1/h)(W_m - W_r)$ and $\nu(rn) = (1/h)(W_r - W_n)$. Kramers considered the limit of very large frequencies ν of the incident radiation, i.e., the cases in which $\nu \gg |\nu(rn)|, |\nu(mr)|$; then Eq. (166) changed into the equation

$$e\Delta q^{(k)}(mn) = \frac{1}{\nu}\,\frac{e^2}{h}\,\mathrm{Re}\left\{\sum_{l=1}^{3} E^{(l)} \sum_{r}\left[q^{(k)}(mr)q^{(l)}(rn) - q^{(l)}(mr)q^{(k)}(rn)\right]\exp(2\pi i\nu t)\right\} \tag{166a}$$

$$- \frac{1}{\nu^2}\,\frac{e^2}{2\pi i h m_e}\,\mathrm{Re}\left\{\sum_{l=1}^{3} E^{(l)} \sum_{r}\left[p^{(l)}(mr)q^{(k)}(rn) - q^{(k)}(mr)p^{(l)}(rn)\right] \times \exp(2\pi\nu t)\right\} + \ldots$$

where m_e is the mass of the electron and the quantities $\nu(mr)q^{(k)}(mr)$ have been replaced by $(2\pi i m_e)^{-1}p^{(k)}(mr)$, where $p^{(k)}(mr)$ are the matrix elements of the momentum matrix $\mathsf{p}^{(k)}$, and the dots denote terms of order ν^{-3}, ν^{-4}, etc.

Having obtained Eq. (166a), Kramers argued as follows. In the limit of large ν the quantum dispersion formula should pass, due to correspondence considerations, into the classical formula; that is, the right-hand side of Eq. (166a) should become equal to $[e^2/4\pi^2 m_e \nu^2]\mathrm{Re}\{E^{(k)}\exp(2\pi i\nu t)\}$.[166] As a consequence, he

[165] We have kept the time-dependent factor $\exp[2\pi i\nu(mn)t]$ in the definition of the matrix elements of $\mathsf{q}^{(k)}$ in agreement with Kramers. In other respects we do not follow him in order to retain closer agreement with the notation of Born, Heisenberg and Jordan. For example, Kramers wrote $q^k_{m,n}$ instead of $q^{(k)}(mn)$. His notation for the matrix product was also a bit unusual; thus, he expressed the first term in the square bracket on the right-hand of Eq. (166) as $\{q^k_{m,n}\}\{q^l_{m,n}/(\nu_{m,n} + \nu)\}$ including the sum over the intermediate states (denoted by the index r in Eq. (166)). (See Kramers, 1925c, p. 373, Eq. (14).)

[166] This classical formula has been discussed in Volume 1, Section VI.1. In case the frequency ν is not very large, an additional term proportional to ν^{-3} has to be considered, whose magnitude depends on the strength of the forces which bind the electron within the atom.

obtained for the system under consideration (having three degrees of freedom) the following quantum conditions:

$$\sum_r q^{(k)}(mr)q^{(l)}(rn) - q^{(l)}(mr)q^{(k)}(rn) = 0 \qquad \text{for all} \quad k, l, \text{ and } m, n, \tag{167a}$$

and

$$\sum_r p^{(k)}(mr)q^{(l)}(rn) - q^{(l)}(mr)p^{(k)}(rn) = \begin{cases} 0 & \text{for } k \neq l \text{ and all } m, n, \\ \dfrac{h}{2\pi i}\delta_{mn} & \text{for } k = l. \end{cases} \tag{167b}$$

(In Eq. (167b), δ_{mn} denotes the Kronecker symbol, which is 1 for $m = n$ and zero otherwise.) These equations agreed with Eqs. (121), or with the quantum condition of Born and Jordan in the case of systems having only one degree of freedom (see Born and Jordan, 1925b, p. 871, Eq. (38), or our Eq. (67) in this volume); hence Kramers concluded: 'The so-called "quantum conditions" of Born and Jordan [Eq. (67), or Eqs. (121) for systems having several degrees of freedom] signify that, in Heisenberg's quantum mechanics, a bound electron in an atom behaves with respect to oscillating fields of high frequencies just as it does in classical theory' (Kramers, 1925c, p. 376).[167]

For one degree of freedom, Heisenberg had already noted the identity of his quantum condition (see Volume 2, Eq. (90)) with Kuhn's sum-rule (Kuhn, 1925a, p. 409) at an early stage (Heisenberg, 1925c, Section 2). Kramers had now generalized this result and thus provided a physical justification of the commutation rules, Eqs. (121), for systems of three degrees of freedom, and his contribution was welcome indeed. However, it is surprising that, at the same time, he did not develop or apply the formalism of matrix mechanics, although he had sufficient command of the mathematics involved.[168] A possible point of departure for that might have been the remark which Kramers made at the end of his note: namely, that *for large quantum numbers the commutation relations, Eqs.*

[167] The commutation relations for the momentum matrices and their elements, Eqs. (121b), could be obtained from Kramers' discussion of dispersion. For this purpose one had simply to consider the momentum changes $\Delta p^{(k)}$ induced by the incident radiation.

[168] At the end of his 'derivation' of the quantum conditions for three degrees of freedom, Kramers proposed also to extend his arguments to atomic systems having several electrons and consequently more than three degrees of freedom (Kramers, 1925c, p. 376). This suggestion somehow indicates the methods by which he wanted to approach atomic theory in the future: namely, by a cautious extension of the previous dispersion approach. In this context he might have thought of the earlier investigations of Werner Kuhn on the dispersion rules for two-electron atoms (Kuhn, 1925a); Kuhn had found that in those cases the dispersion sum assumed the value 2 instead of 1, and this result seemed to imply a possible change in the quantum conditions (121). In any case, Kramers advocated that future developments in quantum mechanics should be strongly subject to comparison with experimental results; thus, he proposed to continue the same approach which had been applied especially by Niels Bohr and his associates for several years. He was not yet convinced that meanwhile a breakthrough had been achieved by Heisenberg, Born and Jordan, which allowed one to develop the theory without explicit reference to detailed experimental results.

(167) *or* (121), *pass into the classical Poisson brackets involving the corresponding variables* (i.e., the position and momentum variables of the corresponding classical system). Although Pauli had directed Kramers' attention to this fact, neither he nor Kramers investigated the consequences of the connection between the quantum conditions and the classical Poisson brackets.[169] On the other hand, after reading Heisenberg's paper on the quantum-theoretical re-interpretation of kinematics (Heisenberg, 1925c), Paul Dirac independently discovered this connection and made it the basis of a complete scheme of quantum mechanics applicable to atomic mechanics in general, far beyond the range of matrix mechanics (Dirac, 1925d).

The Kramers–Heisenberg theory of dispersion had provided the immediate point of departure for the discovery of quantum mechanics. It seemed natural, therefore, to derive the well-known dispersion formulae. This had been done already, as we have mentioned, in September 1925 by Heisenberg. He had succeeded fully and had considered this success of the matrix theory as confirming its validity, for, as he wrote to Pauli: 'I believe that this [i.e., the derivation of the dispersion formulae] is all that one can demand for the moment' (Heisenberg to Pauli, 18 September 1925). Although he was satisfied for the moment, Heisenberg soon thought of testing whether the matrix methods could also be applied successfully to describe other 'dispersion' phenomena, such as the scattering of fast α-particles by atoms, or, more generally, all those situations which could be treated as the action of time-dependent external forces on atoms. These phenomena had been discussed in detail during the previous months both in Copenhagen and Göttingen (Bohr, 1925a; Born and Jordan, 1925a), and now people at those places, especially Niels Bohr, were interested in their renewed discussion in terms of quantum mechanics. On 21 October 1925, a few days after returning to Göttingen from Copenhagen, Heisenberg reported to Bohr about the first little progress he had made. He wrote:

> If, in analogy to the classical theory, one investigates the conditions under which it makes sense to talk about "external forces," that is, to neglect the reactions, then

[169] In his paper Kramers remarked:
In the limit of large quantum numbers these conditions [Eqs. (121)] pass over into the following expressions of classical mechanics:

$$(q^{(k)}q^{(l)}) = 0, \qquad (p^{(k)}q^{(l)}) = \begin{cases} 0 & (k \neq l) \\ 1 & (k = l) \end{cases}, \qquad (p^{(k)}p^{(l)}) = 0.$$

With the expressions (ab) one denotes the known symbol of Poisson:

$$(\mathrm{ab}) = \sum_k \left(\frac{\partial a}{\partial I_k} \frac{\partial b}{\partial w_k} - \frac{\partial a}{\partial w_k} \frac{\partial b}{\partial I_k} \right),$$

where the canonical variables, I_k, w_k, are nothing else but the uniformizing variables (the w's are Stäckel's "angle variables") which play such a great role in the quantum theory of periodic systems. The quantities I_k in the theory are put equal to integer multiples of Planck's constant. However, such "quantum conditions" have to be considered out of date from the point of view of Heisenberg's theory. The methods developed in the paper of Kramers and Heisenberg [1925] yield, when adjusted to Poisson brackets, directly the result that the equations [(121) or (167)] are the fitting quantum-theoretical "re-interpretation" of (classical) bracket expressions. To this interpretation of [Eqs. (121)] Pauli has already drawn attention. (Kramers, 1925c, p. 376)

one finds that also in quantum mechanics the computational rules of the *classical* theory are always valid in the same approximation for the periodic functions occurring in the external forces (these forces may be either periodic in time or expandable in a Fourier integral), provided these conditions [i.e., the reactions can be neglected] are satisfied. This appears to me again to constitute a beautiful example of the correspondence-like connection of the new theory with the classical theory. (Heisenberg to Bohr, 21 October 1925).

Heisenberg's discussion of the dispersion phenomena became the basis of a full section of the three-man paper, dealing with these problems in a systematic way from the point of view of matrix theory (see BHJ, 1926, Chapter 1, Section 5, entitled 'System for which the Time Enters Explicitly in the "Energy Function"').

Born, Heisenberg and Jordan formulated the problem of the action of time-dependent forces on atoms in quantum mechanics ' . . . as the limiting case of interaction between two systems in which the influence of the interaction on one of the two systems (called system A) is so small that the action upon the other (system B) remains unaffected by this influence' (BHJ, 1926, p. 568). In classical mechanics the total Hamiltonian function describing the two systems is then given by a sum of three terms, say, $H = H_A + \lambda H_B + \lambda\epsilon H_{AB}$, where the first, H_A, denotes the energy of free system A, the second, λH_B, the energy of the free system B, and the third, $\lambda\epsilon H_{AB}$, the coupling or interaction energy. The third term is supposed to be a correction, which is small compared to the total energy and to the energies H_A and λH_B (i.e., $\epsilon \ll 1$). Now the system A is taken to be known. Then the motion of system B may be determined by the Hamiltonian function $\lambda H_B + \lambda\epsilon H_{AB}$, where the interaction energy term, $\lambda\epsilon H_{AB}$, depends, of course, on the coordinates of the system A. The latter coordinates may be time dependent; however, for a fixed time t, only given values of these coordinates (and their conjugate momenta) enter into the classical calculation. If the reaction of the system B on system A may be neglected—say, due to the fact that system B has a much smaller energy than A (or $\lambda \ll 1$)—then the motion of B is completely described by the Hamiltonian function $\lambda H_B + \lambda\epsilon H_{AB}(p_B, q_B; t)$, where the interaction term depends explicitly on time t (which takes into account the dependence on the variables of A). Thus, an adequate classical description of the action of an external time-dependent force on system B is provided.

Born, Heisenberg and Jordan then turned to the discussion of the same problem in quantum mechanics. They began by noting:

In the quantum-mechanical calculation the situation is just the same, provided we restrict ourselves to first-order perturbations (i.e., terms proportional to ϵ in the coordinates and momenta, etc., of the system B). It is altogether different, however, for higher-order perturbations, since in evaluating higher-order terms we encounter products of quantities in which more than one implicitly contains the coordinates of A. But this means, according to the quantum mechanical rule of building a product, that it by no means suffices to know the "external forces as functions of time" merely for the *given* values of the constants in A, but these external forces must be known for *all* values of these constants. Thereby, however, the concept of external forces appears in fact to become devoid of meaning. (BHJ, 1926, pp. 568–569)

As mentioned above, the difficulties with the concept of external forces did not show up in the lowest-order approximation; in that case one neglects the reaction of system B on A, which leads to a change of the variables of A, and this change, in turn, induces additional terms of order $\lambda\epsilon^2$ in the variables of B.[170] The latter result even allows one to make a more precise statement concerning the concept of external forces in quantum mechanics. Evidently, it makes sense to talk about external forces if the action of the system A on the system B can be described in a manner independent of the dynamical variables of B. This is especially so when the parameter λ, which characterizes the influence of the relative magnitude of the system B with respect to the system A, is very small. The authors noted, however, that 'in this approximation the quantum mechanical construction of products, and thereby the calculation of the perturbation to higher orders in ϵ can again be effected. In fact, the rules for this building of products reduce simply to those of classical multiplication, as in this approximation the coordinates, amplitudes and frequencies which enter into H_{AB} do not depend on the constants in A' (BHJ, 1926, p. 569).

A careful examination of the conditions, under which the concept of external forces had a well-defined meaning in their theory, enabled Born, Heisenberg and Jordan to suggest that certain important problems of atomic physics could perhaps be solved by applying a time-dependent quantum-mechanical perturbation theory. In particular, they argued that:

> In this sense one could, for example, treat the action of a strong alternating electromagnetic field on an atom indeed as the influence of an "external force" by neglecting the reaction of the atom [on the incident electromagnetic radiation], since the field energy can be regarded as infinitely large compared with that of the atom. The action of α-particles upon the electrons of an atom could also be viewed as an "external force," as in classical theory, because of the large energy of the α-particles; in this approximation, the Fourier expansion in time of the force thereby exerted upon the electrons would also be the same as in classical theory. (BHJ, 1926, p. 569)

On the other hand, the authors were very well aware of the fact that an approach of that kind was not applicable to some other phenomena: for instance, the action of forces due to one atom upon another could not be treated as an operation of external forces (except in the first-order terms) since neglecting the reaction in higher-order terms would lead to wrong results.[171]

[170] System B acts on the System A via the coupling term $\lambda\epsilon\mathsf{H}_{AB}$. Hence the original dynamical variables of the system A might be shifted by terms of order $\lambda\epsilon$. Since system A reacts on B again via the coupling term $\lambda\epsilon\mathsf{H}_{AB}$, a change of the variables of B, p_B and q_B, is induced, which has the order $\lambda\epsilon^2$, i.e., it is of the second order in the small perturbation parameter ϵ and of first order in λ.

[171] In choosing their examples, the authors wanted to draw attention to the fact that their new theory allowed them to deal, in principle, with all the phenomena that Bohr had discussed earlier (Bohr, 1925a). However, the phenomena, which Bohr had not been able to treat either with classical theory or by the then existing perturbation approach, were now identified with situations where one had to apply higher-order time-dependent perturbation theory involving possibly products of non-commuting matrices and difficulties with the concept of external forces.

In order to develop their time-dependent perturbation scheme, Born, Heisenberg and Jordan turned to the standard problem of dispersion theory: the scattering of radiation of frequency ν by an atom.[172] For simplicity one may assume that **E**, the electric field vector of the incident radiation, is parallel to the position vector of the atom; the problem may then be treated in one dimension (having one degree of freedom). In classical theory the interaction with the radiation changes the Hamiltonian of the free atom by a term $eq_0E\cos(2\pi\nu t)$, where E is the amplitude of the electric field and $-e$, the charge of the electron, plays the role of a small perturbation parameter; q_0 denotes the position of the electron, i.e., its distance from the positively charged nucleus.[173] In matrix theory one has just to substitute the Hermitian matrix q_0 for the position coordinate q_0, and the interaction term becomes accordingly

$$\lambda\mathsf{H}_1 = e\mathsf{q}_0E\cos(2\pi\nu t). \tag{168}$$

Hence, in case the frequency of the incident radiation is finite, i.e., $\nu \neq 0$, the perturbation term in the Hamiltonian matrix (describing the atom plus external radiation) shows a time dependence which is not included in the dynamical variables of the atom. It is evident that in this case a perturbation scheme different from the one given by Eqs. (135) has to be used in order to calculate the first-order correction to the position matrix or the induced electric moment of the atom because Eqs. (135) had been derived only for those Hamiltonian matrices which depend exclusively on the dynamical variables p and q of the atomic systems under investigation.[174]

The fact that a time-dependent perturbation scheme was necessary for dealing with time-dependent external forces was already known from classical dynamics.

[172] Evidently, the authors expected to deal later on with other 'dispersion' problems, such as the scattering of α-particles by atoms. However, the available matrix formulation prevented them from treating aperiodic motions; only later on would Born return to the scattering problem by using Schrödinger's wave mechanical formulation (Born, 1926a, b, c).

[173] Since the electron's charge e has a dimension, it must be normalized in order to reveal its magnitude. The dimensionless parameter corresponding to e is the square root of Sommerfeld's fine-structure constant α $(= 2\pi e^2/hc)$, which is smaller than 0.1.

[174] In the problems treated so far by perturbation theory, Born, Heisenberg and Jordan had assumed that the position and momentum variables depend periodically on time with the characteristic frequencies, $\nu(mn)$, of the atomic system. In the dispersion problem, however, the frequency of the incident radiation is, in general, different from any of the frequencies $\nu(nm)$; hence the time dependence via the exponential function $\cos(2\pi\nu t)$ must be considered explicitly in the Hamiltonian.

The situation is, of course, different when the external frequency ν is very small compared to all frequencies $\nu(mn)$; the time dependence in Eq. (168) may then be forgotten. After calculating the transformation matrix S_1 according to the second equation (135) and substituting it into the equation for q_1 $(= \mathsf{S}_1\mathsf{q}_0 - \mathsf{q}_0\mathsf{S}_1)$, one finds the matrix elements as

$$q_1(mn) = \frac{E}{h}\left[\sum_{k\neq m}\frac{q_0(mk)q_0(kn)}{\nu_0(mk)} - \sum_{k\neq n}\frac{q_0(mk)q_0(kn)}{\nu_0(kn)}\right].$$

This result was hinted at by Heisenberg in a letter to Jordan, dated 13 September 1925, as reproducing the results of dispersion theory. Evidently, this is true only in the limit of very low frequencies of the incident radiation.

A corresponding scheme could be formulated in quantum mechanics, yielding instead of Eqs. (135), the new perturbation equations

$$\begin{aligned} &\mathsf{H}_0(\mathsf{p}_0, \mathsf{q}_0) = \mathsf{W}_0, \\ &\mathsf{S}_1\mathsf{H}_0 - \mathsf{H}_0\mathsf{S}_1 - \frac{h}{2\pi i}\frac{\partial \mathsf{S}_1}{\partial t} + \mathsf{H}_1 = \mathsf{W}_1, \\ &\mathsf{S}_2\mathsf{H}_0 - \mathsf{H}_0\mathsf{S}_2 - \frac{h}{2\pi i}\frac{\partial \mathsf{S}_2}{\partial t} + \left(\mathsf{H}_0\mathsf{S}_1 - \mathsf{S}_1\mathsf{H}_0 + \frac{h}{2\pi i}\frac{\partial \mathsf{S}_1}{\partial t}\right)\mathsf{S}_1 \\ &\quad + \mathsf{S}_1\mathsf{H}_1 - \mathsf{H}_1\mathsf{S}_1 + \mathsf{H}_2 = \mathsf{W}_2, \end{aligned} \tag{169}$$

involving an explicitly time-dependent transformation matrix S $(= 1 + \lambda\mathsf{S}(t) + \lambda^2\mathsf{S}_2(t) + \ldots)$.[175]

Now the first two Eqs. (169) allowed one to calculate the induced electric moment of an atom created by the incident radiation of frequency ν. By decomposing the periodic function $E\cos(2\pi\nu t)$ in Eq. (168) into two terms, $\frac{1}{2}E\exp(2\pi i\nu t)$ and $\frac{1}{2}E\exp(-2\pi i\nu t)$, the second of Eqs. (169) becomes satisfied by two types of transformation matrix elements,

$$S_1(mn, +1) = \frac{E}{2h}\frac{q_0(mn)}{\nu_0(mn) + \nu}\exp(2\pi i\nu t) \tag{170a}$$

and

$$S_2(mn, -1) = \frac{E}{2h}\frac{q_0(mn)}{\nu_0(mn) - \nu}\exp(-2\pi i\nu t), \tag{170b}$$

where the additional plus or minus sign in the matrix elements refers explicitly to the positive or negative sign in the factors $\exp(\pm 2\pi i\nu t)$. Evidently, $S_1(mn, +1)$ and $S_1(mn, -1)$ are associated with the respective terms in the perturbation Hamiltonian, Eq. (168). (The $\nu_0(nm)$ refer to the frequencies of the unperturbed atom.) With the help of this result, Born, Heisenberg and Jordan immediately obtained the perturbation correction to the position coordinate—i.e., $\lambda\mathsf{q}_1 = e(\mathsf{S}_1\mathsf{q}_0 - \mathsf{q}_0\mathsf{S}_1)$—and consequently also the formulae expressing the matrix elements of the induced electric moment $e\Delta\mathsf{q}$ $(= e^2\mathsf{q}_1)$, namely,[176]

$$e\Delta q(mn, +1) = \frac{Ee^2}{2h}\sum_k \left[\frac{q_0(mk)q_0(kn)}{\nu_0(mk) + \nu} - \frac{q_0(mk)q_0(kn)}{\nu_0(kn) + \nu}\right]\exp(2\pi i\nu t) \tag{171a}$$

[175] Heisenberg, in his letter to Pauli, dated 18 September 1925, had already indicated the additional terms which appear in Eqs. (169) as compared to Eqs. (135). He had written: 'If H depends explicitly on the time, then there occur—as in classical theory—additional terms on the left-hand side: $-(h/2\pi i)(\partial\mathsf{S}_1/\partial t)$, $-(h/2\pi i)(\partial\mathsf{S}_2/\partial t)$, etc. This then enables one to carry out the integration [of the equations of perturbation theory]' (Heisenberg to Pauli, 18 September 1925).

[176] Note that in the three-man paper (BHJ (1926), pp. 571–572) Eqs. (39) and (40) are incomplete because the matrix elements of S_1 and of the magnetic moment are defined without the factors $\exp(\pm 2\pi i\nu t)$.

and

$$e\Delta q(mn,-1)=\frac{Ee^2}{2h}\sum_k\left[\frac{q_0(mk)q_0(kn)}{\nu_0(mk)-\nu}-\frac{q_0(mk)q_0(kn)}{\nu_0(kn)-\nu}\right]\exp(-2\pi i\nu t). \tag{171b}$$

If the state of the atom remains unchanged, that is, if only matrix elements with $m=n$ are considered, Eq. (171a) becomes identical with the Kramers dispersion formula for coherent scattering (Kramers, 1924a). On the other hand, the matrix elements with $m\neq n$ reproduce the dispersion formulae for incoherent scattering (Kramers and Heisenberg, 1925, §4). Finally, the authors had no difficulty in considering Eqs. (171) in the limit of very high frequency of the incident radiation, $\nu\gg|\nu_0(mn)|$, and in deriving the classical Thomson formula for the scattering of light by a free electron.[177] Matrix mechanics was thus able to provide a complete justification for all the dispersion formulae which had been obtained earlier on the basis of correspondence arguments.

Born, Heisenberg and Jordan dealt with a still different approach to problems described by time-dependent Hamiltonians (BHJ, 1926, Chapter 1, Section 5, pp. 570–571). This approach may be applied to all cases in which the Hamiltonian H depends, besides the dynamical variables p_k and q_k, on the time t via a periodic function with a frequency ν, i.e., $\mathsf{H}=\mathsf{H}(\mathsf{p}_k,\mathsf{q}_k,\cos(2\pi\nu t))$; moreover, the time dependence may already exist in the unperturbed problem. To treat such problems, the authors suggested the following procedure: One formally introduces the new dynamical variables q' and p', where q' is a diagonal matrix having all its elements identical to $\cos(2\pi i\nu t)$, and p' the canonically conjugate momentum obtained from the new Hamiltonian matrix H',

$$\mathsf{H}'=\mathsf{H}(\mathsf{p}_k,\mathsf{q}_k;\mathsf{q}')+2\pi\nu\sqrt{1-\mathsf{q}'^2}\,\mathsf{p}'. \tag{172}$$

The explicit time dependence was thus absorbed into a new, independent degree of freedom, and the previous problem turned into a formally time-independent problem described by the *Ersatz* Hamiltonian H'.[178]

The authors unfortunately did not illustrate the application of the above idea by a practical example, nor did they return to it later; the brief episode was forgotten completely. Still, the procedure involving the introduction of a new pair

[177] For large ν the authors obtained the matrix formula,

$$e^2\mathsf{q}_1=-\frac{Ee^2}{2\pi ih\nu^2m_e}(\mathsf{p}_0\mathsf{q}_0-\mathsf{q}_0\mathsf{p}_0)\cos(2\pi\nu t),$$

where m_e is the mass of the electron and p_0 the momentum variable. By substituting the commutation relation, this yields the induced moment as $Ee^2\cos(2\pi\nu t)/(4\pi^2\nu^2m_e)\cdot\mathsf{1}$.

[178] Evidently, the introduction of the new variables does not influence the equations of motion for the variables p_k and q_k.

of variables and the *Ersatz* Hamiltonian H′ was certainly interesting, at least in a formal sense: it constituted the first attempt at replacing the time coordinate by a quantum-theoretical variable, in this case a matrix. The attempt to give a quantum-theoretical meaning to time would be made again several times later on, especially by Wolfgang Pauli and Paul Dirac, and would lead to some important consequences.[179]

IV.2 Fluctuations in Cavity Radiation

In October 1925 Heisenberg reported to Pauli:

> Another thing, which Jordan has [especially] done for our paper is [to treat] the statistical behaviour of eigenvibrations, say of a membrane, in the new theory. Jordan claims that one obtains for the interference fluctuations the right result, i.e., both the classical and the Einstein terms, and he believes that he sees an analogy of our calculations with Bose's statistics. I am a little unhappy about it, for I do not understand statistics enough in order to decide whether there is something reasonable in it, but I cannot also criticize it, because I find the problem itself and the calculation significant. (Heisenberg to Pauli, 23 October 1925)[180]

The extension of matrix mechanics to systems of many degrees of freedom had provided the necessary tool for investigating the problem of cavity radiation from a fresh point of view. What Heisenberg, at the time, considered a 'significant problem and calculation,' Pascual Jordan, the author of this investigation, later on described as his 'most important contribution to quantum mechanics' (Jordan to van der Waerden, 10 April 1962).

In his quantum-mechanical treatment of radiation, Jordan developed a consistent 'wave' theory of cavity radiation, which was completely in the spirit of Bohr's ideas and at the same time reproduced the light-quantum effects predicted by Einstein (1909a). In order to understand the progress made by Jordan we must recall how the radiation problem stood in summer 1925.

The theory of radiation was characterized by two distinct approaches whose fundamental assumptions seemed to be in serious conflict with each other. On one hand, Einstein's quantum theory of an ideal gas (Einstein, 1924c, 1925a, b) made use of his concept of the light-quantum and the statistics introduced by

[179] We shall discuss Pauli's first attempts to treat time as the variable conjugate to the energy in Chapter V, Section 4, of this volume. Dirac's treatment of 'quantum time' will be discussed in Volume 4, Section V.4.

[180] Pascual Jordan was the expert on radiation theory in Born's Institute ever since he dealt with this subject in his thesis (Jordan, 1924). In 1925 he had treated the problem of radiation in two papers: first, in the joint work with Born on the quantum theory of aperiodic motion (Born and Jordan, 1925a), where the authors derived dispersion-theoretical expressions for Einstein's emission and absorption coefficients (§10, pp. 500–503); second, in a note on the thermal equilibrium between atoms and cavity radiation (Jordan, 1925d), where he showed that the results of Pauli (1923b, 1924b) and of Einstein and Ehrenfest (1923) remained unchanged when Bose statistics was used for the atoms rather than the classical (Boltzmann) statistics.

Bose (1924a). In Göttingen and Copenhagen, on the other hand, there was a tendency, as embodied in the theory of Bohr–Kramers–Slater (BKS, 1924), to stick to the wave description of electromagnetic radiation even in considering the phenomena of absorption and emission of light by atoms. One of the basic assumptions of the BKS theory of radiation, namely, that energy and momentum were not strictly conserved in individual interactions between radiation and matter, was proved to be wrong by the experiments of Bothe and Geiger (1925a, b). It was, therefore, a challenge for Bohr's followers to demonstrate that all phenomena which had been explained hitherto on the basis of light-quanta would follow from a suitable quantum-theoretical reformulation of the classical wave theory.[181]

In summer semester, 1925, Paul Ehrenfest gave a lecture on 'The Energy Fluctuations in a Radiation Field due to the Superposition of Quantized Eigen-oscillations' in the Göttingen Seminar on 'The Structure of Matter.' When Jordan took up the discussion of the statistics of electromagnetic radiation in matrix mechanics, the content of this lecture—which Ehrenfest also published in *Zeitschrift für Physik* (Ehrenfest, 1925a)—was of great help to him for two reasons: First, Ehrenfest had put his finger precisely on that feature of the classical radiation theory which could not be true in quantum theory, this being the concept of the superposition of statistically independent waves. Second, he had illustrated his arguments by using the simplest model for cavity radiation: a string, fixed at both ends, capable of undergoing elastic (transverse) vibrations. Jordan not only adopted Ehrenfest's model for the quantum-mechanical treatment; he also found that the latter completely confirmed the failure of the classical superposition property.[182]

[181] The opinion that the light-quantum description did not give the true nature of radiation had been expressed several times during the previous years by Niels Bohr and his collaborators, and they had tried to avoid the light-quantum description completely. Thus, for example, Kramers and Heisenberg, in their joint paper on dispersion theory, had emphasized that the phenomena of secondary radiation—which Adolf Smekal had treated on the basis of light-quantum considerations (Smekal, 1923c)—could be dealt with satisfactorily according to the wave description of radiation (Kramers and Heisenberg, 1925, pp. 685–686, 695–707).

[182] Let us note here that in summer 1925 the problem of the quantum-theoretical wave concept had received particular attention in Göttingen, especially in connection with the idea of matter waves that had been proposed by Louis de Broglie (1924e, 1925). Walter Elsasser, a young student of James Franck's, had proposed to explain the Ramsauer effect (Ramsauer, 1921b, 1922, 1923) on the basis of de Broglie's ideas (Elsasser, 1925). De Broglie's ideas had also been related by Einstein (1925a) to the new statistics governing ideal quantum gases (Bose, 1924a; Einstein, 1924c). During his visit to Göttingen in summer 1925 Paul Ehrenfest helped Born and his collaborators to understand the connection between the matter waves and Bose–Einstein statistics. (See Born's remark in his letter to Einstein, dated 15 July 1925: '*Glücklicherweise erschien Ehrenfest hier und hat uns ein Licht aufgesteckt.*' ['Fortunately Ehrenfest turned up here and enlightened us.'])

All the physicists in Göttingen had enjoyed Ehrenfest's visit, which took place in the second half of June 1925. While his lecture had enlightened Born, Heisenberg and Jordan, others did not find much in it. Thus, Frits Zernike, after seeing the published paper (Ehrenfest, 1925a), wrote: '[I have] full appreciation of your clear statements—but, apart from the special clarity of the presentation, does it contain anything that had not been calculated by Ornstein and myself?' (Zernike to Ehrenfest, 1925, undated; reported in Ehrenfest, 1925b, p. 316).

In classical theory the vibrations of a one-dimensional string of length L may be described by u $(= u(x,t))$, the lateral displacement, which satisfies the equation of motion

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}\,, \tag{173}$$

if the wave velocity is normalized to unity. With the boundary condition $u(0,t) = u(L,t)$, Eq. (173) admits a periodic solution that can be expressed by an infinite sum of terms, $\sum_{k=1}^{\infty} q_k(t)\sin(k\pi x/L)$, where $q_k(t)$ is a periodic function of time (i.e., $q_k(t) = \sin(\omega_k t + \phi_k)$, $\omega_k = k\pi/L$ being the frequency and ϕ_k $(k = 1, 2, \ldots)$ the phase constant). The energy of the vibrating string is then given by $H(L,t)$, where

$$H(L,t) = \frac{L}{4}\sum_{k=1}^{\infty}\left\{\dot{q}_k^2\{t\} + \omega_k^2 q_k^2(t)\right\}, \tag{174}$$

which is obviously independent of time.[183] The vibrating string thus appears as a collection of harmonic oscillators with frequencies ω_k and effective 'mass' $L/2$. Ehrenfest had also evaluated the energy of a segment of length l $(\ll L)$ of the string and found the result[184]

$$H(l,t) = \frac{1}{4}\sum_{j,k=1}^{\infty}\left\{\dot{q}_j(t)\dot{q}_k(t)K_{jk} + \omega_j\omega_k q_j(t)q_k(t)K'_{jk}\right\}, \tag{175}$$

where

$$K_{jk} = \frac{\sin\{(j-k)(\pi/L)l\}}{\omega_j - \omega_k} - \frac{\sin\{(j+k)(\pi/L)l\}}{\omega_j + \omega_k} \tag{175a}$$

and

$$K'_{jk} = \frac{\sin\{(j-k)(\pi/L)l\}}{\omega_j - \omega_k} + \frac{\sin\{(j+k)(\pi/L)l\}}{\omega_j + \omega_k}\,. \tag{175b}$$

Ehrenfest's aim had been to analyze the status of radiation theory and the consistency of its assumptions. In particular, he had asked the question whether the assumption of the statistical independence of waves, which is valid in classical theory, leads to any conflict with quantum-theoretical results. His

[183] Equation (174) is obtained by evaluating the integral, $\frac{1}{2}\int_0^L \{\dot{u}^2 + (\partial u/\partial x)^2\}\,dx$, under the assumption that the length of the string is large compared to each of the wavelengths associated with one of the frequencies, ω_k, of the eigenvibrations.

[184] For calculating $H(l,t)$, the square of the spatial sine functions cannot be replaced by their average values. In the case that l becomes large, i.e., of the order of L, the second terms on the right-hand sides of Eqs. (175a) and (175b) can be neglected, and the first terms in both equations become equal to unity; hence Eq. (175) agrees in this limit with Eq. (174).

calculation of the energy fluctuation in the string had shown that Einstein's formula for blackbody radiation (Einstein, 1909a, p. 189) could not be derived even if he used Planck's law for the distribution of thermal energy among the vibrational modes of the string.[185] Since the new statistical approach of Bose and Einstein obviously yielded the right fluctuation formula, Ehrenfest had concluded that a fundamental inconsistency existed in the wave theory of radiation (Ehrenfest, 1925a, p. 362).

Jordan, in fall 1925, had quite a different goal in mind. He started with the conviction that the wave picture *must* yield a consistent treatment of radiation, including the quantum effects. He decided against invoking any reminiscence of the light-quantum concept, such as had been used, for example, by Peter Debye in his ingenious derivation of Planck's radiation law (Debye, 1910). As he remarked:

> Such a mixture of theoretical wave and light-quantum considerations would seem to us hardly to accord with the real nature of the problem. Rather, we believe it to be consistent to separate the theoretical wave-aspect of the problem completely from the theory of light-quanta, that is, to treat the wave-statistics of black-body radiation throughout by the more general statistical rules applying, e.g., to quantum-theoretical atomic sytems. (BHJ, 1926, p. 607)

Jordan had available a more or less complete quantum-theoretical scheme, namely, matrix mechanics, which, he believed, should be able to determine the validity of the *assumptions* that Ehrenfest, for instance, had used in his analysis: in particular, how the available energy was distributed among the vibrational modes of the string and whether it was justified to talk about statistically independent waves. Moreover, thought Jordan, if matrix mechanics was a valid theory it should also lead to the correct 'light-quantum term' in Einstein's fluctuation formula. He, therefore, started by simply translating Ehrenfest's classical treatment into the new quantum-theoretical language.

In matrix mechanics, the dynamical variables of the one-dimensional string, such as $u(x,t)$, the lateral displacement entering Eq. (173), and q_k, its Fourier amplitudes, as well as $H(l,t)$, the energy of a segment of the string according to Eq. (175), had to be represented by matrices. On the other hand, the spatial coordinate had now to appear on the same footing as the time coordinate t, that is, as a number; the reason being that if a string were replaced by a sequence of mass points connected by elastic forces, then x can be represented by the *number* of a mass point multiplied by the lattice constant.[186] Since the summations over k

[185] In particular, Ehrenfest had not been able to derive the correct 'light-quantum' term (see Ehrenfest, 1925a, especially §§4 and 6).

[186] This treatment of the string was taken over later on into lattice dynamics. It is of interest to note, however, that in 1925 Born did not immediately consider lattice dynamics from this point of view, and he did not indicate any consequences of it in his lectures at M.I.T. (Born, 1926d). The reason probably was that thus far the quantum-mechanical treatment of radiation had only been Jordan's favourite subject.

in Eqs. (174) and (175) extend to infinity, that is, the elastic string has infinitely many eigenvibrations, Jordan had to generalize the matrix scheme also to be able to handle systems with infinitely many degrees of freedom. Thus the quantities q_k, describing time-dependent amplitudes of the lateral displacements in the matrix re-interpretation of the classical problem, had to be regarded as infinite matrices, that is, $\mathsf{q}_k = (q_k(m_1, \ldots, m_k \ldots ; n_1, \ldots, n_k \ldots))$, $k = 1, 2, \ldots, \infty$. Evidently, all matrix elements of q_k, except those for which $m_j = n_j$ for $j \neq k$ and $n_k = m_k \pm 1$, are zero. The Hamiltonian matrix, $\mathsf{H}(l, t)$ describing the behaviour of a segment of the string in quantum mechanics, can be obtained by substituting the matrices q_j and q_k and their time derivatives in Eq. (175).

Jordan was especially interested in calculating the energy fluctuation term, that is, the time average, $\overline{\boldsymbol{\Delta}^2}$, of the difference $\boldsymbol{\Delta}$, where

$$\boldsymbol{\Delta} = \mathsf{H}(l, t) - \bar{\mathsf{H}}(l, t). \tag{176}$$

Here $\bar{\mathsf{H}}(l, t)$ denotes the phase or time average of the Hamiltonian matrix, a matrix containing only the diagonal elements of $\mathsf{H}(l, t)$. In evaluating the expressions entering Eq. (176) one is allowed, provided the length of the string under investigation is much larger than the wavelengths of the eigenvibrations, to replace j- and k-sums (appearing, e.g., in the quantum-theoretical reformulation of Eq. (175), whose right-hand side involves terms such as $\boldsymbol{\Delta}_1 = \frac{1}{4}\sum_{j,k=1}^{\infty} \dot{\mathsf{q}}_j(t) \cdot \dot{\mathsf{q}}_k(t) K_{jk}$ and $\boldsymbol{\Delta}_2 = \frac{1}{4}\sum_{j,k=1}^{\infty} \omega_j \omega_k \mathsf{q}_j(t) \mathsf{q}_k(t) K'_{jk}$) by integrals over a continuous path of frequencies: hence the sums $\sum_j \ldots$ go over into integral $(L/\pi) \cdot \int_0^{\infty} d\omega_j \ldots$. Assuming further that the length l of the string segment under consideration is large enough, so that the quantity $(j + k)(l/L) \gg 1$ for most values of j and k, then the second terms in K_{jk} and $K'_{jk'}$, Eqs. (175a) and (175b), can be neglected and both factors become equal. As a result, any double integral over the frequencies ω_j and ω_k will turn into a single integral; for example, $\mathsf{H}(l, t)$ is then given by the expression

$$\mathsf{H}(l, t) = \frac{lL}{4\pi} \int_o^{\infty} \left[\dot{\mathsf{q}}_\omega^2 + \omega^2 \mathsf{q}_\omega^2 \right] d\omega, \tag{177}$$

where $\dot{\mathsf{q}}_\omega$ and q_ω denote the former quantities, $\dot{\mathsf{q}}_j(t)$ and $\mathsf{q}_j(t)$, respectively; consequently, its time average becomes[187]

$$\bar{\mathsf{H}}(l, t) = \frac{lL}{2\pi} \int_o^{\infty} \overline{\mathsf{q}_\omega^2}\, d\omega. \tag{177a}$$

If one wants to obtain the integral expression for the matrix $\boldsymbol{\Delta}$, one has to be careful, for it is composed of two double sums, i.e., $\boldsymbol{\Delta}_1 = \frac{1}{4}\sum_{j,k=1}^{\infty} \dot{\mathsf{q}}_j(t) \dot{\mathsf{q}}_k(t) K_{jk}$

[187] The assumption of large length l simplifies the calculation considerably because then the second terms in K_{jk} and K'_{jk} become zero. However, it is also possible to evaluate the case in which l is small, hence $(j + k)l/L \sim 1$. (See Born and Jordan, 1930, §73.)

and $\Delta_2 = \frac{1}{4}\sum_{j,k=1}^{\infty}\omega_j\omega_k \mathsf{q}_j(t)\mathsf{q}_k(t)K'_{jk}$, which do not include terms with $j = k$. The same situation arises when one considers the square, Δ^2, and its time average, $\overline{\Delta^2}$, the matrix describing energy fluctuations. The latter matrix may be decomposed into two contributions, $(\overline{\Delta_1^2}+\overline{\Delta_2^2})$ and $(\overline{\Delta_1\Delta_2}+\overline{\Delta_2\Delta_1})$. Jordan evaluated the first contribution and obtained the result[188]

$$\overline{\Delta_1^2} + \overline{\Delta_2^2} = \frac{lL^2}{4\pi}\int_{\pi}^{\infty}\left(\omega^2\overline{\mathsf{q}_\omega^2}\right)^2 d\omega. \tag{178a}$$

In classical theory the second contribution becomes zero when one makes the obvious assumption that the eigenvibrations are statistically independent. Ehrenfest had derived this result explicitly in his lecture at Göttingen in summer 1925.[189] However, *the assumption of statistical independence of the eigenvibrations cannot be taken over into quantum theory*. There the expression $\overline{\Delta_1\Delta_2}+\overline{\Delta_2\Delta_1}$ must be calculated explicitly, and Jordan obtained the result

$$\overline{\Delta_1\Delta_2} + \overline{\Delta_2\Delta_1} = \frac{lL^2}{8\pi}\int_0^{\infty}\left[(\overline{\dot{\mathsf{q}}_\omega\mathsf{q}_\omega})^2+(\overline{\mathsf{q}_\omega\dot{\mathsf{q}}_\omega})^2\right]\omega^2 d\omega. \tag{178b}$$

This expression involves products of noncommuting matrices q_ω and $\dot{\mathsf{q}}_\omega$, for the latter (if multiplied by the 'mass' of the string, i.e., the factor $L/2$) can be considered as the momentum p_ω. By making use of the commutation relations for q_ω and p_ω, i.e., Eqs. (121), Eq. (178b) can also be written as

$$\overline{\Delta_1\Delta_2} + \overline{\Delta_2\Delta_1} = -\frac{l}{4\pi}\int_{\pi}^{\infty}\left(\frac{h\omega}{2\pi}\right)^2 d\omega\cdot\mathsf{1}, \tag{178b'}$$

where $\mathsf{1}$ is the unit matrix.

[188]The derivation of Eq. (178a) presents no difficulties, because the matrices in $\overline{\Delta_1^2}$ and $\overline{\Delta_2^2}$ all commute. Hence one obtains for $(\overline{\Delta_1^2}+\overline{\Delta_2^2})$ the expression

$$\frac{L^2}{4\pi^2}\int_0^{\infty}d\omega\int_0^{\infty}d\omega'\,\frac{\sin^2[l(\omega-\omega')]}{(\omega-\omega')^2}\left(\omega^2\omega'^2\,\overline{\mathsf{q}_\omega^2}\;\overline{\mathsf{q}_{\omega'}^2}\right).$$

By changing the integration variables ω and ω' to $\omega_+ = \frac{1}{2}(\omega+\omega')$ and $\omega_- = \frac{1}{2}(\omega-\omega')$, and observing the relation

$$\lim_{l\to\infty}\frac{1}{l}\int_{-\Omega}^{+\Omega}\sin^2\frac{(l\omega_-)}{\omega_-}\,f(\omega_-)\,d\omega_- = \pi f(0),$$

one obtains indeed Eq. (178a).

[189]The main reason is that under the assumption of statistical independence the relation

$$\overline{q_j}\;\overline{q_k} = \overline{a_j}\;\overline{a_k}\;\overline{\cos(\omega_j t+\phi_j)}\cdot\overline{\cos(\omega_k t+\phi_k)} = 0,$$

holds; hence all terms contributing to $\Delta_1\Delta_2 + \Delta_2\Delta_1$ can be shown to be zero in classical theory. (See Ehrenfest, 1925a, p. 369, §9.)

In order to compare the quantum-mechanical fluctuation matrix $\overline{\mathbf{\Delta}^2}$ with the expression from Einstein's fluctuation formula (Einstein, 1909a, p. 189), Jordan considered the contribution from a specific frequency ν $(=\omega/2\pi)$, which he reformulated as

$$\overline{\mathbf{\Delta}_\nu^2} = \frac{1}{2l}\left[\overline{\mathsf{H}}_\nu(l,t)\right]^2 - \frac{l}{2}(h\nu)^2 1, \tag{179}$$

where $\overline{\mathsf{H}}_\nu(l,t)$ is the time average of the Hamiltonian matrix $\mathsf{H}_\nu(l,t)$,

$$\mathsf{H}_\nu(l,t) = lL(2\pi\nu)^2\mathsf{q}_\nu^2, \tag{180}$$

associated with the frequency ν.[190] Apart from a factor $2l$, $\mathsf{H}_\nu(l,t)$ may be regarded as the Hamiltonian matrix of a harmonic oscillator of mass $L/2$ and frequency ν, whose eigenvalues—given by $(n+\frac{1}{2})h\nu$, where $n=0, 1, 2, \ldots$—all contain $\frac{1}{2}h\nu$, the constant zero-point energy term. $\mathsf{H}_\nu(l,t)$ may, therefore, be expressed as the sum of two terms,

$$\mathsf{H}_\nu(l,t) = \mathsf{H}_\nu^{\text{th}}(l,t) + 2l\frac{h\nu}{2}1, \tag{180a}$$

where the first term on the right-hand side represents the 'thermal' part of the Hamiltonian, while the second term denotes the constant zero-point energy. After substituting this expression for $\mathsf{H}_\nu(l,t)$ in Eq. (179), Jordan obtained the following result for $\overline{\mathbf{\Delta}_\nu^2}$:

$$\overline{\mathbf{\Delta}_\nu^2} = h\nu\overline{\mathsf{H}_\nu^{\text{th}}}(l,t) + \frac{1}{2l}\left[\overline{\mathsf{H}_\nu^{\text{th}}}(l,t)\right]^2. \tag{179a}$$

Thus, the fluctuation matrix, $\overline{\mathbf{\Delta}_\nu^2}$, consisted of two terms, which had to be related to the terms in Einstein's formula (see Einstein, 1909a, p. 189). The latter is

$$\overline{\Delta_\nu^2} = h\nu\,\overline{E_\nu} + \frac{\overline{E_\nu^2}}{z_\nu V}, \tag{181}$$

giving the energy fluctuation, $\overline{\Delta_\nu^2}$, of blackbody radiation in the frequency ν contained in a volume V; $\overline{E_\nu}$ denotes the average energy distributed to the frequency ν at a given temperature T, and z_ν $(=8\pi\nu^2\,d\nu/c^3)$ is the number of modes. By identifying l with the volume V in the case of the one-dimensional

[190] Evidently, the time-averaged part of the Hamiltonian for the frequency interval, ν and $\nu+d\nu$, follows from Eq. (177) as

$$\overline{\mathsf{H}_\nu}(l,t)d\nu = lL(2\pi\nu)^2\overline{\mathsf{q}_\nu^2}\,d\nu,$$

where q_ν is identical with q_ω, except that the variable ω becomes replaced by ν.

string, and putting z_ν equal to 2, Jordan arrived at a complete agreement between the result of his quantum-mechanical calculation, Eq. (179a), and Eq. (181).[191]

Jordan's result pleased Born and Heisenberg, and the three collaborators concluded that, 'If one bears in mind that the question considered here is actually somewhat remote from the problems whose investigation led to the development of quantum mechanics, the result [Jordan's] . . . can be regarded as particularly encouraging for the further development of the theory' (BHJ, 1926, p. 615). Indeed, by considering waves in matrix mechanics, Jordan had been able to derive the same fundamental result: namely, a formula containing the light-quantum term, which followed from the application of Bose statistics to light-quanta (Einstein, 1925a, p. 8, Eq. (34)), thereby taking, in a sense, the first step towards the quantization of wave fields.

Although this line of investigation appeared to be very promising, it was not pursued further. Jordan recalled that one reason was: 'Later on Heisenberg explicitly doubted whether this application, which I had made of quantum mechanics to the case of a system of *infinitely many degrees of freedom*, was correct' (Jordan to van der Waerden, 10 April 1962). To be sure, Jordan's treatment of wave fields was not without difficulties, as it involved formally divergent expressions, such as the integral in Eq. (178b′). Though, in the three-man paper, the authors avoided any discussion of the complete frequency spectrum of the string extending, in principle, to infinity, they were conscious of the mathematical difficulties presented by the treatment of infinitely many degrees of freedom.

Heisenberg's doubts, as well as Born's, were perhaps enhanced by the objections raised by Albert Einstein against the treatment of the vibration of the string in the three-man paper. As Einstein informed Jordan: 'The thing with the fluctuations is wrong. It is true that one can compute with the zero-point term $\frac{1}{2}h\nu$ the average magnitude of the fluctuations, but not the probability for a very large fluctuation. For weak (i.e., Wien's) radiation, for example, the probability for finding the total radiation in a fraction V of the total volume V_0 is $W = (V/V_0)^{E/h\nu}$. This, evidently, cannot be explained with the zero-point term, although the expression is substantiated on the basis of thermodynamics' (Einstein to Jordan, 6 March 1926).

In spring 1926 Jordan did not have any good answer to Einstein's argument concerning the zero-point energy of radiation, and his satisfaction in having derived the fluctuation formula (179a) became muted at the time.[192] However, he continued his interest in the radiation problem and would later on make important contributions to the quantum theory of wave fields.

[191] In order to establish the agreement one has to assume that the quantum-theoretical expression $\overline{\mathsf{H}_\nu^{th}}(l, t)$ corresponds indeed to the average energy at a given temperature. Born and Jordan would discuss the implications of this assumption later on and work out the details of the comparison between Jordan's formula (179a) and Einstein's formula (181). (See Born and Jordan, 1930, §73.)

[192] Jordan referred to Einstein's objection in his later review of the light-quantum hypothesis, where he pointed out a possibility of removing the difficulty (Jordan, 1928, p. 194).

IV.3 The Conservation of Angular Momentum

Compared to the detailed formalism of matrix mechanics that had been developed, its applications thus far had been quite poor; they were confined to systems which consisted of harmonic or slightly anharmonic oscillators. For instance, the derivation of the dispersion formulae implied only the use of a harmonic oscillator model in which the transition frequencies of the scattering atoms had been introduced as empirically known constants. However, in a dynamical theory of atomic systems the frequencies and the intensities of the spectral lines should actually be obtained from the equations of motion; hence the appropriate general methods of integrating these equations had to be developed in matrix mechanics. The hope, of course, was that these general methods would then allow one, as in the classical case, to perform the integration for realistic atoms and molecules having one or more electrons. In spite of serious endeavours, Born, Heisenberg and Jordan, during their collaboration in fall 1925, did not succeed in solving any such atomic problem. They assembled all the preliminary results obtained in that direction in Section 4.1 of the three-man paper (BHJ, 1926, pp. 595–605). Besides some general remarks on the use of conserved quantities in quantum-mechanical problems, the authors dealt in detail with the properties of angular momentum and a preliminary application of their formalism to intensity problems of complex spectra. The latter parts were worked out entirely by Werner Heisenberg and Pascual Jordan.[193]

In the discussion of conserved properties in quantum mechanics, the authors could rely heavily on the analogous situation in classical mechanics. In the latter, the general properties of H, the Hamiltonian describing a given dynamical system, serve directly to integrate the equation of motion. For instance, the conservation of energy can be derived as a consequence of the property that H does not depend explicitly on time (except via the dynamical variables p and q). The conservation of momentum and angular momentum can be similarly derived from other properties of H, which are also called symmetry properties. If for a system of f degrees of freedom the values of f independent constants of motion are known, then—according to a fundamental theorem of classical mechanics—

[193] Jordan had begun to work on the matrix formalism of angular momentum very early, probably even before the completion of the first paper on matrix mechanics in September 1925. By the time of Heisenberg's return from Copenhagen to Göttingen in October, Jordan had written a first draft of the three-man paper, which included a section on the angular momentum. Heisenberg had worked independently on this problem in Copenhagen, and it was then decided that he and Jordan should compose the final draft of this section. Jordan later specified the respective contributions as follows:

> Up to about p. 600 [i.e., the end of the first part of the angular momentum treatment, including the commutation relations for the components of the angular momentum and the polarization rules] the content was from me. The important Eq. (18) [i.e., the formula describing the line intensities] and the consequences derived from it, which led to the selection rule for j [the total angular momentum] came from Heisenberg. However, the formula (24) [i.e., the formula describing the quantization of the total angular momentum], as also (25) [i.e., the formula yielding the relations between certain matrix elements of the angular momentum components], were already contained in the older manuscript of Born and myself. Page 604 [where the intensity formulae for the multiplets are presented] again came from Heisenberg. (Jordan to van der Waerden, 23 July 1962)

the motion of the system is completely determined. Born, Heisenberg and Jordan now asked whether to these classical results analogous ones corresponded in quantum mechanics. There were two aspects to this question: first, how to find the conserved dynamical variables for a given quantum system; second, how such properties could be used in order to integrate the equations of motion, which meant specifically how the Hamiltonian matrix H describing the system could be diagonalized.

To answer the second of these questions, the authors reminded themselves of the fact that in classical mechanics the knowledge of a conserved dynamical variable for a system of f degrees of freedom reduces the integration problem to one in which the system has $(f-1)$ degrees of freedom.[194] In matrix mechanics a similar situation existed; that is, there also the conserved quantities could be used to solve the dynamical equations of motion. To demonstrate this point, Born, Heisenberg and Jordan referred to the following argument. A conserved dynamical variable of a quantum system may be represented by a matrix, say A $(=\mathsf{A}(\mathsf{p},\mathsf{q}))$, which is constant in time. If the system is not degenerate, then A becomes diagonal. Let $\boldsymbol{\phi}$ $(=\boldsymbol{\phi}(\mathsf{p},\mathsf{q}))$ be any other dynamical variable of the same system. The difference,

$$\boldsymbol{\phi}\mathsf{A}-\mathsf{A}\boldsymbol{\phi}=\boldsymbol{\psi}, \tag{182}$$

can be calculated with the help of the commutation relations for the p's and q's, the pairs of canonically conjugate dynamical variables of the system under investigation. The matrix equation (182) leads to a set of infinitely many algebraic equations for the elements of the matrix $\boldsymbol{\psi}$. Since A is diagonal, i.e., $\mathsf{A}=(\delta_{nm}A_n)$, the equation for a particular matrix element, $\psi(nm)$, contains only one matrix element, $\phi(nm)$, of the matrix $\boldsymbol{\phi}$ and the two diagonal elements, A_n and A_m, of the matrix A, that is,

$$\phi(nm)(A_n-A_m)=\psi(nm). \tag{182a}$$

If the system under consideration is degenerate, then A, the matrix of the conserved variable, need not be diagonal. Even in this case, however, the algebraic equation for any particular element $\psi(nm)$, derived from Eq. (182), will contain only a finite number of elements of the matrix $\boldsymbol{\phi}$ (and of the matrix A).

The application of this result to the integration of the equations of motion of a quantum system is straightforward.[195] For this purpose, one has just to replace

[194] In classical mechanics a system of f degrees of freedom is treated by solving the Hamilton–Jacobi partial differential equation for the action function S. If one constant of motion is known, say the action variable J (for a periodic system), then S may be written as the sum of a term wJ (where w is the angle variable which is canonically conjugate to J) and of a function S^*, which does not depend on J anymore but only on the remaining $(f-1)$ integration constants of the system. Thus, the problem of solving the Hamilton–Jacobi equation of a system of f degrees of freedom is reduced to the problem of solving an equation of $(f-1)$ degrees of freedom.

[195] The considerations which follow were not stated in the three-man paper, but they follow immediately from the results obtained previously (in Chapter III).

the matrix Φ by the Hamiltonian matrix H of the system. Obviously, H commutes with any conserved quantity A; hence the matrix Ψ on the right-hand side of Eq. (182) must have only zero elements. In the case of a nondegenerate system of one degree of freedom the matrix A is diagonal; so must be, because of Eq. (182), the Hamiltonian matrix H. Then, however, the problem of integrating the equations of motion is solved automatically because the quantum-theoretical Hamiltonian can be expressed as a function of the variable A analogously to the classical case; in particular, the diagonal elements of H may be calculated from the matrix elements of A. In the case of systems with several, say f, degrees of freedom, the equation

$$\mathsf{HA} - \mathsf{AH} = 0 \tag{183}$$

does not necessarily imply that both matrices H and A are diagonal. Even in the case that the system is nondegenerate, different energy states may be associated with one value, A_n, of the conserved variable A. However, if one succeeds in obtaining a complete set of f independent conserved variables, $\mathsf{A}, \mathsf{A}', \mathsf{A}'', \ldots$, the integration of the equations of motion of the system would be achieved without difficulty. The reason is that the Hamiltonian H does not depend on the variables which are canonically conjugate to $\mathsf{A}, \mathsf{A}', \mathsf{A}'', \ldots$; hence it must be a function of the conserved variables alone, and its matrix elements will depend only on the elements of the matrices $\mathsf{A}, \mathsf{A}', \mathsf{A}'', \ldots$. In order to obtain the dependence of H on the A's one has to solve simultaneously f matrix equations of the type of Eq. (183). The eigenvalues of the conserved quantities, $A_n, A'_{n'}, A''_{n''}, \ldots$, i.e., the elements of the diagonalized matrices, $\mathsf{A} = (\delta_{nm} A_n)$, $\mathsf{A}' = (\delta_{n'm'} A'_{n'})$, $\mathsf{A}'' = (\delta_{n''m''} \mathsf{A}''_{n''})$, etc., determine the Hamiltonian matrix H completely; especially, H becomes diagonal in f pairs of indices $(nm), (n'm'), (n''m''), \ldots$. In that sense the knowledge of the conserved variables plays a similar role as in classical theory for the solution of quantum-theoretical problems.

The practical question now arose about finding suitable conserved variables of quantum systems. A possible guide for the search was provided by the conserved variables of the corresponding classical systems. In the previous models of atomic structure, angular momentum had played an important role as a conserved quantity; some atomic properties, such as the energy states of diatomic molecules or the intensities of components of complex spectra of many-electron atoms, had been treated by considering the atoms as rotating systems. Heisenberg and Jordan, therefore, discussed in detail the matrices describing the angular momentum of quantum systems. To begin with, they defined the components of the total angular momentum of a system consisting of N particles (electrons and nuclei) by re-interpreting the corresponding classical expressions in terms of matrices. That is, they replaced the Cartesian coordinates, q_{kx}, q_{ky}, q_{kz}, and the conjugate linear momentum coordinates, p_{kx}, p_{ky}, p_{kz}, of the kth particle ($k = 1, \ldots, N$), by the Hermitian matrices, $\mathsf{q}_{kx}, \mathsf{q}_{ky}, \mathsf{q}_{kz}$ and $\mathsf{p}_{kx}, \mathsf{p}_{ky}$, p_{kz}, describing the analogous variables in matrix mechanics. Thus, they arrived at the following definition of the Cartesian components of the total angular

momentum of the system[196]:

$$\mathsf{M}_x = -\sum_{k=1}^{N} (\mathsf{p}_{ky}\mathsf{q}_{kz} - \mathsf{q}_{ky}\mathsf{p}_{kz}), \tag{184}$$

$$\mathsf{M}_y = -\sum_{k=1}^{N} (\mathsf{p}_{kz}\mathsf{q}_{kx} - \mathsf{q}_{kz}\mathsf{p}_{kx}), \tag{184b}$$

$$\mathsf{M}_z = -\sum_{k=1}^{N} (\mathsf{p}_{kx}\mathsf{q}_{ky} - \mathsf{q}_{kx}\mathsf{p}_{ky}). \tag{184c}$$

The usefulness of the formal introduction of angular momentum in quantum mechanics became immediately evident when Heisenberg and Jordan considered systems, which can be described by separable Hamiltonian matrices. In that case the time derivatives of the angular momentum components, $\dot{\mathsf{M}}_x$, $\dot{\mathsf{M}}_y$ and $\dot{\mathsf{M}}_z$, could be shown to be separable functions of the position and momentum variables of the system; i.e., they were sums of functions, $\phi(\mathsf{p}_k)$ and $\psi(\mathsf{q}_k)$, each of which depended only on the momentum (p_k, $k = 1, \ldots, N$) *or* position variables (q_k, $k = 1, \ldots, N$).[197] Since all p_k commute with each other, as all q_k commute, an important consequence followed: if in a given classical system a Cartesian component of angular momentum is conserved, the same holds for the corresponding quantum system.[198]

The most important property of the matrices M_x, M_y and M_z, defined by Eqs. (184), is expressed by their commutation relations. The latter were obtained with the help of the commutation relations, Eqs. (121), for systems having several degrees of freedom. Thus Born, Heisenberg and Jordan found[199]

$$\mathsf{M}_x\mathsf{M}_y - \mathsf{M}_y\mathsf{M}_x = -\frac{h}{2\pi i}\mathsf{M}_z, \tag{185a}$$

$$\mathsf{M}_y\mathsf{M}_z - \mathsf{M}_z\mathsf{M}_y = -\frac{h}{2\pi i}\mathsf{M}_x, \tag{185b}$$

[196] The definition of the angular momentum coordinates by Eqs. (184) agrees with the definition in the three-man paper only up to a negative sign (see BHJ, 1926, Chapter 4, p. 596, Eqs. (1)). The above definition, however, is the same as given by Pauli in his treatment of the hydrogen atom in the case of one rotating particle (Pauli, 1926a, p. 345, Eqs. (26)), and by Born in his lectures on atomic dynamics (Born, 1926d, Lecture 16, Eqs. (3)). We have kept the original notation of BHJ, i.e., M_x, M_y and M_z, for the components of the angular momentum; Pauli used P_x, P_y and P_z instead.

[197] The time derivative of, say, M_x, contains terms such as $\dot{\mathsf{q}}_{ky}\mathsf{p}_{kz}$ and $\mathsf{q}_{ky}\dot{\mathsf{p}}_{kz}$, which may be expressed as $(\partial\mathsf{H}/\partial\mathsf{p}_{ky}) \cdot \mathsf{p}_{kz}$ and $-\mathsf{q}_{ky} \cdot \partial\mathsf{H}/\partial\mathsf{q}_{kz}$, respectively. Now, if H is a separable Hamiltonian, i.e., a sum of two matrix functions, $\phi(\mathsf{p}_{kx}, \mathsf{p}_{ky}, \mathsf{p}_{kz})$ and $\psi(\mathsf{q}_{kx}, \mathsf{q}_{ky}, \mathsf{q}_{kz})$, then $(\partial\mathsf{H}/\partial\mathsf{p}_{ky}) \cdot \mathsf{p}_{kz}$ is a function of the momentum coordinates only and $-\mathsf{q}_{ky} \cdot \partial\mathsf{H}/\partial\mathsf{q}_{kz}$ is a function of the position coordinates only.

[198] The same statement is true for the total linear momentum of the system, whose components are $\mathsf{p}_x = \sum_{k=1}^{N}\mathsf{p}_{kx}$, $\mathsf{p}_y = \sum_{k=1}^{N}\mathsf{p}_{ky}$, $\mathsf{p}_z = \sum_{k=1}^{N}\mathsf{p}_{kz}$.

[199] Born, Heisenberg and Jordan, in their paper, gave only the first of these equations (BHJ, 1926, Chapter 4, p. 597, Eq. (3)); the other two follow in the same way.

$$\mathsf{M}_z\mathsf{M}_x - \mathsf{M}_x\mathsf{M}_z = -\frac{h}{2\pi i}\mathsf{M}_y\,. \tag{185c}$$

These equations stated the noncommutativity of the Cartesian coordinates of angular momentum in quantum mechanics, and they led to an immediate consequence: either *all three* components M_x, M_y and M_z are conserved, or *none* of the three components is conserved, or just *one* component is conserved.[200] The same theorem holds for classical systems; however, while it is proved easily in matrix mechanics by a few algebraic steps, it takes a long computation to establish it in classical mechanics. The reason is that the dynamical equations, from which the conservation of a given classical variable has to be derived, are very complicated partial differential equations; in quantum theory, on the other hand, the corresponding dynamical equations are purely algebraic equations. The theorem on the conservation of the components of angular momentum in quantum theory, contained in Chapter 4 of the Born–Heisenberg–Jordan paper, thus represented the first example of a general situation: namely, that the conservation laws for dynamical systems can be established much more easily in quantum mechanics than in classical mechanics.[201]

Before applying the matrix formalism of angular momentum to specific systems, the authors studied the general question concerning the possible eigenvalues of the matrices M_x, M_y and M_z. Since it is sufficient to consider the problem just for one matrix, M_z, they proceeded in the following way. They first treated a system whose motion is symmetrical with respect to the z-axis. As a result of this symmetry property, M_z must be conserved at least for systems described by Hamiltonian matrices, which are separable in the Cartesian coordinates and momenta of their constituent particles, just as in the classical theory.[202] Moreover, the authors assumed that the energy states of the system under investigation are nondegenerate, i.e., $W_n \neq W_m$ for $n \neq m$. Then the matrix M_z becomes diagonal, or $\mathsf{M}_z = (\delta_{mn}\mathsf{M}_{zn})$, with the diagonal elements numbered by the same index, n, as the energy states. By commuting the position matrices of the constituent, q_{kx}, q_{ky}, q_{kz}, with the matrix M_z, Heisenberg and Jordan

[200] The last two statements of the above theorem are evident. To prove the first statement, one first assumes that two components, say M_x and M_y, have zero time derivatives. Then it follows from Eq. (185b) that the expression $\mathsf{M}_y(\mathsf{H}\mathsf{M}_z - \mathsf{M}_z\mathsf{H}) - (\mathsf{H}\mathsf{M}_z - \mathsf{M}_z\mathsf{H})\mathsf{M}_y$ must be zero. Since M_y and M_z do not commute, it follows that M_z must commute with H; hence M_z is also conserved if M_x and M_y are conserved quantities.

[201] Later, especially Hermann Weyl (1928) would discuss the symmetry properties in quantum mechanics in detail.

[202] These systems include most of the complex atoms, whether they are free or under the influence of a static electric or magnetic field in a given direction. (In the first case the preferred direction is provided by the direction of the core's angular momentum in the atomic core–series electron model of complex atoms.) In classical theory the time independence of the z-component of the angular momentum follows from the fact that the Hamiltonian of a system (having symmetry around the z-axis) does not depend on the angle variable ϕ_z (canonically conjugate to M_z). In matrix mechanics this argument could not be taken over because ϕ_z had not yet been included among the variables described by Hermitian matrices. By restricting themselves to separable Hamiltonians, however, Heisenberg and Jordan could take over the classical result without involving the existence of the angle variable in matrix mechanics.

obtained the three infinite sets of equations:

$$q_{lx}(nm)(M_{zn} - M_{zm}) = \frac{h}{2\pi i} q_{ly}(nm), \tag{186a}$$

$$q_{ly}(nm)(M_{zn} - M_{zm}) = -\frac{h}{2\pi i} q_{lx}(nm), \tag{186b}$$

$$q_{lz}(nm)(M_{zn} - M_{zm}) = 0, \tag{186c}$$

where M_{zn} and M_{zm} are two eigenvalues of M_z. These equations describe the polarization of the radiation emitted by the system; they state in particular that radiation with finite intensity (which is proportional to $|q_x(nm)|^2$, $|q_y(nm)|^2$ and $|q_z(nm)|^2$) occurs only if the value of the angular momentum component M_z changes by the amounts $-h/2\pi$ or zero.[203] As a consequence, all eigenvalues of M_z may be ordered in sequences of the form

$$M_{zn} = \frac{h}{2\pi}(n_1 + C), \quad n_1 = \ldots, -2, -1, 0, 1, 2, \ldots, \tag{187}$$

where C is a constant, which still has to be determined. Moreover, one is allowed to conclude that all eigenvalues of M_z of a system with axial symmetry can be separated into several sequences with different constants, C', C'', etc., such that C' is not equal to C'' plus an integer; no transitions will occur between the states belonging to different sequences because of the selection rule which states that the eigenvalues of M_z may change only by 0 or $h/2\pi$.

The question must be asked why Heisenberg and Jordan numbered the eigenvalues of M_z in Eq. (187) by the index n_1 rather than the index n, which characterizes the energy eigenvalues. The reason was that in a physical system many states of different energy may be associated with a given value, $(n_1 + C) \cdot h/2\pi$, of angular momentum component. Therefore, one should replace the index n by the sum of two indices, $n_1 + n_2$ (both n_1 and n_2 are integer numbers, of course), with n_2 denoting the different energy states which have the same M_z-value. This decomposition of the index n allowed the authors to write all matrices describing variables of the axially symmetric system as four-dimensional matrices with two index pairs, $(n_1 m_1)$ and $(n_2 m_2)$. Thus, for example, the dependence of the matrices q_{kx} (or q_{ky}) and q_{kz} on the first index pair may be expressed by the factors $\delta_{1,|n_1 - m_1|}$ and $\delta_{n_1 m_1}$, respectively, and the same holds for the angular momentum components M_x and M_z. Since, however, M_z is supposed to be a diagonal matrix, its elements may be simply written, in agreement with Eq. (187), as

$$M_z(n_1 n_2 ; m_1 m_2) = \delta_{n_1 m_1} \delta_{n_2 m_2} \frac{h}{2\pi}(n_1 + C). \tag{187'}$$

[203] The last result follows immediately from Eq. (186c), the former by combining Eqs. (186a) and (186b).

While M_x and M_y do not commute with M_z, the matrix M^2 describing the square of the total angular momentum, i.e., $\mathsf{M}^2 = \mathsf{M}_x^2 + \mathsf{M}_y^2 + \mathsf{M}_z^2$, does; hence it will also be diagonal in both index pairs.

Having exhausted the information obtainable from an axially symmetric, nondegenerate system, Heisenberg and Jordan considered systems described by a Hamiltonian H_0 plus a small perturbation $\lambda\mathsf{H}_1$. They assumed further that for a vanishing perturbation parameter ($\lambda = 0$) all three angular momentum components are conserved, i.e., $\dot{\mathsf{M}}_x^0 = \dot{\mathsf{M}}_y^0 = \dot{\mathsf{M}}_z^0 = 0$; hence the system described by H_0 is spherically symmetric, and the perturbation term $\lambda\mathsf{H}_1$ destroys the symmetry around the x-axis and y-axis. Then for $\lambda = 0$ the system must be degenerate; in spite of the fact that all matrices are time independent, one can choose only M_z^0 (and $(\mathsf{M}^0)^2$) as diagonal. However, it is still possible to assume the other components, M_x^0 and M_y^0, as being diagonal in the index pair (n_2m_2) discussed above. From the latter representation the authors concluded immediately that the diagonal elements of the matrix $(\mathsf{M}_x^0)^2 + (\mathsf{M}_y^0)^2$ were always positive; the same must be true, of course, of the diagonal elements of the matrix $(\mathsf{M}^0)^2 - (\mathsf{M}_z^0)^2$.[204] Now $(\mathsf{M}^0)^2$, the matrix of the degenerate system, does not depend on n_1, the index characterizing the eigenvalues of the matrix M_z^0; as a consequence the number of possible eigenvalues of M_z^0 for a given index n_2 will be finite. This means that for a fixed value of n_2 the two-dimensional submatrices of M_x^0, M_y^0 and M_z^0 in the index pair (n_1m_1) must be finite matrices; and these finite matrices still satisfy Eq. (185a). In taking the diagonal sum, $D(\mathsf{M}_x^0\mathsf{M}_y^0 - \mathsf{M}_y^0\mathsf{M}_x^0)$, involving the finite matrices, Heisenberg and Jordan noticed that the result was zero; hence they derived the equation

$$\sum_{n_1}(n_1 + C)\frac{h}{2\pi} = 0, \tag{188}$$

where the n_1-sum extends to all possible values n_1 for a fixed n_2. Equation (188) implies that the constant C in Eq. (187) can assume only two values, $C = 0$ or $C = \frac{1}{2}$. The authors therefore concluded: 'If for the angular momentum around the z-axis [i.e., M_z] we now introduce the notation usually used in the literature, namely, $m(h/2\pi)$ in place of $(n_1 + C)(h/2\pi)$, we have shown indeed that the selection rule $m \to (m + 1, m, m - 1)$ applies to m and that m is either "integer" or "half-integer"' (BHJ, 1926, p. 600).

The fact that the matrix formalism of angular momentum admitted only *integral* or *half-integral* eigenvalues for M_z settled an old argument, which had arisen in connection with the explanation of the anomalous Zeeman effects and the band spectra of molecules in the earlier quantum theory of multiply periodic systems. On occasion, it had seemed necessary to introduce half-integral quantum numbers (see, e.g., Heisenberg, 1922a; Kratzer, 1922). Now this earlier ad hoc assumption was endowed with a firm basis, and at the same time other

[204] This result follows from the expression for the elements of the matrix, i.e., $((M_x^0)^2 + (M_y^0)^2)(n_1n_2;m_1m_2) = \delta_{n_1m_1}\delta_{n_2m_2} \cdot \sum_{k_1}\{|M_x^0(n_1n_2;\, k_1n_2)|^2 + |M_y^0(n_1n_2;\, k_1n_2)|^2\} \geq 0$.

fractional quantum numbers were excluded from being ascribed to the components of angular momentum.[205]

The results obtained thus far were sufficient to discuss the properties of radiation emitted by an atom (having N electrons), on which an external (static electric or magnetic) field acts in a given direction, say in the direction of the z-axis. In this case the transitions are characterized by the transition in m, the eigenvalue (up to a factor $h/2\pi$) of the matrix M_z, and the polarization of the emitted radiation follows from Eqs. (186): i.e., if the value of m does not change at all, then the plane of radiation emitted by the kth electron is parallel to the z-axis; if, however, the value of m changes by one unit, the amplitude of the emitted radiation parallel to the z-axis becomes zero, hence the plane of the emitted radiation will be perpendicular to the z-axis (the radiation being circularly polarized in this case). In free atoms such an axis of symmetry, which is fixed in space, does not exist. In the Bohr–Sommerfeld theory of atomic structure, the practice then was to introduce the direction of the total angular momentum (provided the latter was different from zero) as a preferred direction. Evidently, this direction is not fixed in space, for the atoms collide with each other and in the absence of an external field the total angular momenta of different atoms have an arbitrary direction. For the purpose of calculating the properties of the radiation emitted by the kth electron in a given atom, it had been found convenient to consider the electron's position coordinate, X_k, Y_k and Z_k, in a coordinate system whose Z-axis was in the direction of the total angular momentum of that atom. Heisenberg and Jordan reformulated these position coordinates of the classical theory in a suitable way as the matrices X_k, Y_k and Z_k, and they succeeded in establishing commutation relations between these matrices and M^2, the matrix describing the square of the total angular momentum of the atom under consideration. From these relations they then derived the eigenvalues of the matrix M^2, that is,

$$M_j^2 = j(j+1)\left(\frac{h}{2\pi}\right)^2. \tag{189}$$

In transitions with finite intensity the value of j changes either by one unit or by zero.[206] Moreover, the authors were able to identify the value of j with the *maximum value* of the index m, which characterizes the eigenvalues of the

[205] Kramers and Pauli (1923) had argued previously that quantum numbers other than integral and half-integral ones might occur in the band spectra of molecules. Now Heisenberg was especially happy that besides the integral quantum numbers only half-integral ones were permitted in quantum mechanics (see Heisenberg to Pauli, 12 October 1925).

[206] In order to derive the commutation relations between X_k, Y_k and Z_k, and M^2, it was assumed that M^2 was a diagonal matrix with diagonal elements M_n^2. Then it was found that Z_k commuted with M^2, but X_k and Y_k did not. As a consequence of the commutation relations the eigenvalues of M^2 could be written in terms of an index j, Eq. (189); moreover, transitions associated with finite elements of the matrix Z_k involved no change of the index j, while transitions associated with finite elements of the matrices X_k and Y_k involved a change of j by one unit.

angular momentum in a given direction.[207] Hence j corresponded to what had been called the quantum number of the total angular momentum of an atom in the former theory of atomic structure. In agreement with the values of m, therefore, j could also assume either integral or half-integral values.

To some extent the above results confirmed those that had been obtained earlier. For example, the same selection rules, i.e., $j \to j, j \pm 1$ and $m \to m, m \pm 1$, came out as in the quantum theory of multiply periodic systems. On the other hand, important differences showed up. From the analysis of spectroscopic data it had been concluded that certain transformations in atoms, such as the transition from a state with $m = 0$ to another state with $m = 0$, was forbidden. In the new theory such exclusions could not be justified at all. As the authors remarked: 'Our result demonstrates further that the exclusion of individual states, such as was, for example, necessary in the earlier theory of the hydrogen atom in order to prevent collisions between the electron and the nucleus, has no place in the theory proposed here' (BHJ, 1926, p. 600).[208] Another strong point in favour of the new quantum theory was that the eigenvalues of $\mathbf{M}^2$, the square of the total angular momentum of an atom, were given by $j(j + 1)$ in units of $(h/2\pi)^2$, which explained the former ad hoc introduction of 'washed out' ('*verschwommen*') values for the squares of the angular momenta in the formulae describing the spectroscopic data. Thus, quantum mechanics offered a genuine chance of deriving, for example, the intensity rules for the components of complex spectra and the associated anomalous Zeeman effects.

The most important test case of the new angular momentum formalism seemed to be the problem of the anomalous Zeeman effects. Early in 1925, Samuel Goudsmit and Ralph Kronig (1925) and Helmut Hönl (1925) had derived formulae for the intensities of Zeeman components on the basis of specific mechanical models using correspondence arguments. A few months later, in June 1925, Heisenberg had indicated that these intensity formulae might be obtained by a quantum-theoretical re-interpretation of the kinematical relations describing the classical rotator (Heisenberg to Kronig, 5 June 1925; Heisenberg, 1925c). Now, in fall 1925, one had to give a consistent and complete derivation of the Goudsmit–Kronig–Hönl formulae. For this purpose Heisenberg and Jordan had just to treat a complex atom with a finite angular momentum and to express the position matrices, $\mathbf{q}_{kx}$, $\mathbf{q}_{ky}$, and $\mathbf{q}_{kz}$, of the kth electron in terms of the matrices $\mathbf{X}_k$, $\mathbf{Y}_k$, and $\mathbf{Z}_k$. Thus, for transitions in which the total angular

[207] This result was obtained by considering the expression, $(\mathbf{M}_x + i\mathbf{M}_y)(\mathbf{M}_x - i\mathbf{M}_y) = \mathbf{M}^2 - \mathbf{M}_z^2 + (h/2\pi)\mathbf{M}_z$, and then substituting the eigenvalues, $j(j + 1)(h/2\pi)^2$, $m^2(h/2\pi)^2$ and $m(h/2\pi)^2$, of the matrices $\mathbf{M}^2$, $\mathbf{M}_z^2$ and $(h/2\pi)\mathbf{M}_z$, respectively. Then followed the relation, $M_x(j, m - 1; j, m) + iM_y(j, m - 1; j, m) = (h/2\pi)\sqrt{j(j + 1) - m(m - 1)}$. Since the transitions from $m_{\max}$ to $m_{\max} + 1$ should lead to zero-matrix elements, j must be identified with $m_{\max}$.

[208] This point was first noted explicitly by Heisenberg in his letter to Pauli, dated 12 October 1925. It would play an important role in Pauli's subsequent analysis of the hydrogen problem in matrix mechanics.

momentum did not change—where the matrix $\mathbf{Z}_k$ had nonzero elements—the intensity ratios were given by the equation

$$|q_{lz}(j,m;j,m)|^2 : |(q_{lx}+iq_{ly})(j,m-1;j,m)|^2 : |(q_{lx}-iq_{ly})(j,m;j,m-1)|^2$$

$$= m^2 : \tfrac{1}{4}\left[j(j+1)-m(m-1)\right] : \tfrac{1}{4}\left[j(j+1)-m(m-1)\right]. \qquad (190)$$

A similar result was obtained for transitions with $j \to j \pm 1$, in which the matrices $\mathbf{X}_k$ and $\mathbf{Y}_k$ had nonzero elements.[209]

All these results, however, did not go beyond the conclusions of the formulae which had been accepted earlier as representing correctly the available empirical spectroscopic data. The virtue of the matrix derivation of the known results lay in the following: First, it became possible to discard specific models of atoms, which had been employed earlier to deal with the role of angular momentum in the analysis of spectroscopic data. All that seemed to be necessary in matrix mechanics was the basic concept of classically rotating electrons and re-interpretation of this motion in matrix language. Second, the empirically devised quantum numbers and their selection rules, by means of which the data had been organized and ordered, now received a well-defined meaning as being related to the total angular momentum (characterized by the index j) and to one of its components (characterized by the index m). In the matrix formulation of Born, Heisenberg and Jordan there was no room for arbitrariness, and the question to be answered with all urgency was whether *all* empirical spectroscopic data could be accounted for by their scheme. This question was related to the important dynamical problem: How could the angular momentum of an atomic system, having a well-defined Hamiltonian matrix, be calculated? How was it related to the energy eigenvalues, say in the case of the simplest atom—hydrogen—and what values would the indices j and m assume in this case? The answer to these, and other, questions had to be obtained soon, for then alone would it be possible to decide whether the matrix formulation of quantum theory could deal with realistic atoms.

IV.4 Wolfgang Pauli's Conversion

In their meeting in Hanover on Sunday, 19 July 1925, Wolfgang Pauli had rejected Max Born's invitation to collaborate with him on a consistent mathematical formulation of Heisenberg's newly discovered rules for the quantum-theoretical re-interpretation of kinematics and mechanics. Pauli had had several reasons in mind for doing so. Psychologically the most important was that in the

[209] In Eq. (190) we have corrected the factors $\frac{1}{4}$ in the second and third expressions on the right-hand side. (See, BHJ, 1926, p. 604, Eq. (30); compare with Born, 1926d, Lecture 17, Eqs. (4) and Born and Jordan, 1930, Chapter 4, p. 160, Eqs. (15).)

course of his investigations of previous years he had arrived at the conclusion that the *real* quantum theory, which would give a coherent account of atomic phenomena, would be born from fresh insights, from the insights of young men such as Heisenberg and himself. Pauli believed that the new quantum theory was again '*Knabenphysik,*' the young man's game, just as it had been for Einstein and Bohr in its earlier stages. He thought that an 'oldster' like Born, forty-three years of age in 1925, would kill the freshness of new ideas by casting them too soon into the framework of an erudite mathematical formalism. Born, Pauli believed, should better stick to his work on crystal lattices and leave quantum theory alone.[210]

Even during the following months of summer and fall 1925 Pauli did not change his negative attitude towards Born's involvement in the new quantum mechanics. On 9 October he wrote a long letter to Ralph Kronig, then at Copenhagen, in which he discussed again, as he had done often during the past year, a detailed mechanical model with the help of which he sought to explain the multiplet structure of spectral lines and their behaviour in magnetic fields. Pauli mentioned to Kronig that he would present this matter along similar lines in his article on quantum theory in *Handbuch der Physik*, which he was about to complete within two or three weeks (Pauli, 1926b). He continued:

> I have already a great longing for a reasonable scientific activity! Heisenberg's mechanics has again given me hope and joy in living. It has not yet brought the solution of the riddle, but I believe that it is again possible to go forward. One must first seek to liberate Heisenberg's mechanics from Göttingen's [i.e., Max Born's] deluge of formal learning (*formalen Gelehrsamkeitsschwall*) and better expose its

[210]The idea that he was not young enough to contribute really decisive ideas had been mentioned by Born himself in a letter to Pauli written as early as 1919. In this letter Born had talked about the necessity of introducing completely new concepts in order to find a way out of the difficulties of quantum theory, and had remarked: 'But how one might do that, I don't have any idea. Though I am not yet old, I am already too old and burdened to arrive at the solution. That is your task; according to what I have heard about you, to solve such problems is your calling.' And then he continued: 'However, if you wish to roam around for the moment in the lowlands of conventional physics, then I recommend to you a problem, which is connected with a just published paper of Stern and myself [Born and Stern, 1919], which I am sending you herewith' (Born to Pauli, 23 December 1919). The problem had been to prove that the surface energy of a cubic sodium chloride (NaCl) crystal is a minimum. Pauli, however, was not interested in the theory of crystal lattices, while Born continued to work in this field over the years.

In my conversations with Pauli during the years 1952–1958 he often referred to the fact that the discovery of and work on the new quantum mechanics and its applications had been '*Knabenphysik*' ('boys' physics'), for all the people involved—Heisenberg, Dirac, Jordan, Pauli himself, and the others, except Born and Schrödinger, had been young men. Unlike his teacher Arnold Sommerfeld, towards whom Pauli always showed great respect, he held Born in mild contempt, principally for the latter's overwhelming interest in mathematical formulations rather than the clarification of physical ideas themselves. In my last conversation with Heisenberg, on 25 February 1975 in Munich, I brought up the question of how sharply Pauli used to treat Born. Heisenberg agreed and said: 'Yes, he was quite disrespectful towards Born. It would not have occurred to me to treat Born like that; he was my professor. But Pauli was different. He could be very sharp and impatient' (Conversations with Pauli and Heisenberg, J. Mehra).

> physical essence. Then one must try, on one hand, by deepening the foundations [of Heisenberg's mechanics] also to get hold of the "stress" ("*Zwang*") and, on the other, to connect it [Heisenberg's mechanics] naturally with the collision phenomena. (Pauli to Kronig, 9 October 1925)

While Pauli was berating Göttingen's deluge of formal mathematical erudition Heisenberg and Jordan, in one of their contributions to the three-man paper, were developing the methods for treating angular momentum in matrix theory by using a trickle of that erudition. To them it seemed that the specific models that had been employed in the classification schemes of 'term zoology'—including Pauli's latest, about which he wrote to Kronig—could now be forgotten. Heisenberg, to whom Kronig showed Pauli's letter, therefore reacted strongly to the criticism of matrix mechanics and he wrote to Pauli:

> With respect to both of your last letters I must preach you a sermon, and beg your pardon for proceeding in Bavarian: It is really a pigsty that you cannot stop indulging in a slanging match. Your eternal reviling of Copenhagen and Göttingen is a shrieking scandal. You will have to allow that, in any case, we are not seeking to ruin physics out of malicious intent. When you reproach us that we are such big donkeys that we have never produced anything new in physics, it may well be true. But then, you are also an equally big jackass because you have not accomplished it either..... (The dots denote a curse of about two-minute duration!) Do not think badly [of me] and many greetings. (Heisenberg to Pauli, 12 October 1925)

Ten days later, the usually good-natured Heisenberg explained to Pauli the reasons for his sudden explosion. He wrote:

> The story of this letter which I composed in a fit of rage (on my own stupidity) is this: Concerning your letter to Kronig I wrote to you what I thought about your models [of atomic structure]; that was very little, for I do not wish to think about these questions for the moment. Further I wrote to you about some new results in quantum mechanics. Finally, it occurred to me, in connection with your letter [to Kronig], that I had been thinking for three days about continuous spectra and understood nothing, and that made me write the said [angry] conclusion [of my letter]. (Heisenberg to Pauli, 23 October 1925)

Indeed, Heisenberg had been struggling hard with the application of matrix methods to actual physical problems. He felt justified in wishing that, instead of playing 'variations on an old theme' ('*Variationen über eine alte Weise*') with such virtuosity on the 'organ of Zoology' ('*die Orgel der Zoologie*') and sending forth blasts of sarcasm against the work of people in Göttingen and Copenhagen, Pauli would 'accomplish something new in physics' ('*etwas physikalisch Neues fertig bringen,*' Heisenberg to Pauli, 12 October 1925).

In the early fall of 1925 Pauli's scientific attitude was still characterized by a deep liking for Bohr's correspondence principle as the guide for treating problems of atomic physics. During the writing of his *Handbuch* article on quantum theory (Pauli, 1926b) he had acquired a special mastery in handling the most

subtle problems of atomic spectroscopy, such as the combination lines of atoms appearing in the presence of an electronic field by means of the correspondence approach (Pauli, 1925d). With this experience Pauli found it even more compelling to adhere to the physical content of the correspondence principle in treating the problems of atomic mechanics. He particularly admired Heisenberg's success in sharpening the correspondence principle and giving it a more precise meaning by re-interpreting kinematical and mechanical relations. The mathematical formulation which Heisenberg's discovery received in Göttingen in the hands of Born, Jordan and Heisenberg himself could only mean, in Pauli's opinion, a fallback to the barrenness of (Born's) formal techniques.

To Heisenberg, Pauli's negative attitude towards the use of matrix methods seemed to be unreasonable, especially since all the pertinent features of the correspondence principle had already been successfully incorporated in matrix mechanics. Heisenberg believed that there was no longer room for any doubt about the consistency of the mathematical framework of the new quantum theory; one had only to apply the theory to solve actual problems.

It was not difficult to decide which problem was the most urgent and had to be treated first. For some time Heisenberg had shared Pauli's view that Bohr's theory of the hydrogen atom contradicted certain fundamental notions which, on the basis of the analysis of the fundamental problems of atomic physics that had taken place since Bohr's lectures in Göttingen in summer 1922, one expected *the real* quantum theory should satisfy. The fact that the Balmer spectrum had been obtained from Bohr's theory appeared to Pauli and Heisenberg, and to Bohr himself, as an 'accident.' One of the principal objections was the use, in the atomic domain, of such 'unobservable' classical pictures as electron orbits. Now that the notion of orbits had been abandoned in the purely quantum description of Born, Heisenberg and Jordan, Bohr's celebrated derivation of the Balmer formula seemed to have lost its justification entirely. From the beginning of his efforts at re-interpreting kinematical and mechanical relations in quantum theory, Heisenberg had aimed at the solution of the problem of determining the spectrum of the hydrogen atom and the intensities of the transition lines. Would the matrix methods now enable one to arrive at the desired goal?

Born, Heisenberg and Jordan, during their joint work on matrix mechanics, had always been aware of the necessity of treating realistic atomic systems. They hoped that the algebraic methods would indeed enable them to do so. Thus, for example, in introducing the theory of Hermitian quadratic forms they remarked: 'Thereby we attain the prospect of finding methods for direct calculation of the energy without explicitly solving the problem of motion: methods which correspond to Sommerfeld's method of complex integration' (BHJ, 1926, p. 581).[211] However, when it came to displaying the methods explicitly, the authors did not

[211] Heisenberg constantly pressed Born and Jordan to enter into the discussion of practical problems in atomic physics. Thus, in connection with the matrix theory of canonical transformations, he remarked: 'Born's form for the canonical transformation [see Section III.2, Eqs. (137)] is almost the high point of the theory; hopefully one would soon succeed in obtaining the methods of integration [of the equations of motion] from it' (Heisenberg to Jordan, 29 September 1925).

present any details, but referred simply in a footnote to the methods of Karl Heinrich Graeffe and Daniel Bernoulli, which one might use for determining the eigenvalues of finite and bounded infinite matrices.[212] Now both methods, cited in the three-man paper, were not applied at all to the solution of problems of atomic physics, for it was not expected that they would be successful in the case of unbounded Hermitian forms or matrices. Heisenberg noted this disparity between the elaborate formalism and the difficulty of using it to calculate the energy states of realistic atomic systems when he wrote to Pauli: 'The most important part of the paper seems to me to be, in spite of its mathematical character, the principal-axes transformation; it is thought to represent only a method of integration, and up to now it is only of little help, but formally it is still very beautiful' (Heisenberg to Pauli, 23 October 1925). To which Pauli rejoined that 'for this "in spite of" ("*trotz*") I could embrace you' ('*Für dieses "trotz" könnte ich Sie umarmen*'), for he just did not believe that the mathematical methods invoked by Born would be of any help in actual physical problems.[213] And Heisenberg, more than Born and Jordan, his collaborators in erecting the framework of matrix mechanics, felt the bitterness of the failure, in particular his own, in dealing with the problems of the hydrogen atom, the solution of which would supply the first crucial test of the new theory.[214]

[212]The problem of determining the eigenvalues of a finite quadratic matrix is equivalent to finding the roots of the associated characteristic equation. In the case of an $(n \times n)$-matrix the characteristic equation is an equation of order n. One can solve such equations rigorously only for the orders 2, 3, and 4. For higher, but finite orders, Daniel Bernoulli (1700–1782) had proposed the following method (Bernoulli, 1732). Suppose that all the roots of the characteristic equation of order n are different and can be arranged according to their decreasing absolute value as $|\lambda_1| > |\lambda_2| > \cdots$. Now consider the equations $(\lambda^h)^n - c_1^{(h)}(\lambda^h)^{n-1} + c_2^{(h)}(\lambda^h)^{n-2} - + \cdots = 0$ $(h = 2, 3, \ldots)$, having the roots, λ_1^h, $\lambda_2^h, \ldots, \lambda_n^h$, h being an integer greater than 1. The coefficients, $c_i^{(h)}$, may be obtained from the corresponding coefficients, c_i, in the characteristic equation (i.e., $\lambda^n - c_1\lambda^{n-1} + \cdots = 0$) by using the method of 'recurrent series.' Then the largest eigenvalue, λ_1, may be calculated from the equation $\lambda_1 = \lim_{h\to\infty}(c_1^{(h)}/c_1^{(h-1)})$, that is, from an approximation procedure involving the evaluation of the coefficients $c_1^{(h-1)}$ and $c_1^{(h)}$. Having obtained λ_1, one may find the second eigenvalue $(\lambda_2 = \lambda_1^{-1} \cdot \lim_{h\to\infty}(c_2^{(h)}/c_2^{(h-1)})$, the third eigenvalue, etc., from a similar approximation method. Bernoulli's method of calculating the eigenvalues of a finite quadratic matrix was extended by Erhard Schmidt to compact matrices (Schmidt, 1905). His method could also be applied to bounded matrices, but in the case of matrices with continuous eigenvalues it made no sense.

Another method for finding the roots of an algebraic equation of finite order (which includes the characteristic equations for $(n \times n)$-matrices) had been proposed by K. H. Graeffe (1799–1873) in a book on the solution of higher-order equations (Graeffe, 1837). His method exhibited some similarity with the one of Bernoulli: again one assumed that the eigenvalues were ordered in a sequence, such that $|\lambda_1| > |\lambda_2| > \cdots$, and one considered equations of the type, $(\lambda^h)^n - c_1^{(h)}(\lambda^h)^{n-1} + \cdots = 0$, but now h had to be an *even* number. Graeffe's method, which can be extended to obtain the eigenvalues of bounded infinite quadratic matrices, converges faster than Bernoulli's.

[213]The above remark of Pauli was reported by Pascual Jordan in a letter to van der Waerden, dated 10 April 1962.

[214]Later, Born described their efforts in Göttingen to solve the hydrogen problem as follows: 'We tried, of course, the hydrogen atom, but we could not work it out. It was stupid because it was straightforward. In my lectures at M.I.T. I presented it already, after learning about Pauli's work from letters. I remember that Heisenberg was also quite angry with himself that he had not done it' (Born, Conversations, p. 50). Born, of course, agreed with Heisenberg that the practical application of matrix methods was an important issue. Thus, he began his Lecture 18 at M.I.T., in which he treated

On the other hand, Pauli's pride was hurt when, on his provocation, Heisenberg had called him a 'jackass who had not accomplished anything physically new.' He will show them that such an epithet did not apply to him and decided to beat the *Göttinger Gelehrten* at their own game of matrix theory. He will solve the problem by means of methods, which he had only railed against thus far, thus picking the pearl of physical wisdom out of the chaff of mathematical formalism.

By mid-October 1925 Pauli knew enough about matrix mechanics; Heisenberg had written to him three times from Copenhagen, describing in detail the progress achieved by Born, Jordan and himself.[215] Pauli knew that in order to find the key to the solution of the hydrogen problem, which had escaped Heisenberg and his collaborators, and to carry the calculation through to completion, he would have to make good use of his own considerable formal power and mathematical virtuosity. And, maybe, if matrix mechanics did not yield the energy states of hydrogen correctly, it could still be pronounced wrong.

Less than three weeks after he had been accused of not having 'accomplished anything physically new,' Pauli reported to Heisenberg about his successful diagonalization of the Hamiltonian matrix of the hydrogen atom. Matrix mechanics had passed its first crucial test. Heisenberg, with great joy and a touch of envy, marvelled that Pauli had completed the theory so soon ('*daß Sie diese Theorie so schnell herausgebracht haben,*' Heisenberg to Pauli, 3 November 1925).

The reason for Pauli's quick success was that he had recognized at once a major weakness of the matrix theory and had found a way of curing it in the case of the hydrogen problem. In his paper he later on characterized this weakness as follows: 'The presently known version of the laws of the new quantum mechanics requires that we avoid the introduction of a polar angle ϕ. Since the latter is not confined within finite limits, it cannot formally be represented as a matrix in the same way as the above-mentioned coordinates [Cartesian coordinates of multiply periodic systems], which execute librations in classical mechanics' (Pauli, 1926a, p. 345).

The difficulty of re-interpreting classical coordinates, which do not describe *librations* and thus cannot be represented by Hermitian matrices, was certainly the principal obstacle in solving the problem of the hydrogen atom in matrix

---

the hydrogen problem, by stating: 'We now come to the crucial question for the whole new theory: Is it able to account for the properties of the hydrogen atom? Let us recall that the explanation of the hydrogen spectrum (Balmer's formula) was the first great success of Bohr's theory [of 1913] and has since remained its keynote. If the new theory failed here it would have to be abandoned in spite of its many conceptual advantages, but, as Pauli has shown, it stands the test successfully' (Born, 1926d; 1970 reprint, p. 113).

[215] Heisenberg had presented to Pauli a detailed introduction to matrix mechanics in a letter in September 1925; this letter included the complete outline of the theory for systems of one degree of freedom and a sketch of the perturbation scheme for that case (Heisenberg to Pauli, 18 September 1925). He wrote again to Pauli on 24 September and 12 October; in the last letter he discussed Born's general transformation theory, by means of which a quadratic Hermitian form can be reduced to a sum of squares.

mechanics. Surprisingly, Born, Heisenberg and Jordan had not encountered any difficulty in treating in matrix language the angular momentum, which is the dynamical variable conjugate to ϕ, the angle of rotation. In fact, however, they had not cared at all about the angle variable, but had considered the angular momentum by itself and drawn certain important consequences for atomic systems from its conservation. Heisenberg had brought one of these results to Pauli's attention, when he remarked in a letter: 'For systems with axial symmetry the angular momentum around the axis is an integral multiple of $h/2\pi$, such that the possible states lie symmetrically with respect to zero and constitute a sequence free of gaps (with no exclusion of zero being possible), whose members differ by $h/2\pi$' (Heisenberg to Pauli, 12 October 1925). This was important information because it implied two observations: first, the angular momentum could be quantized in matrix mechanics; second, the eigenvalues of the z-component of angular momentum, when described by integral quantum numbers, included the value $m = 0$. (See Eq. (188) with $C = 0$ and $n_1 = m$.) Pauli knew that in the Bohr–Sommerfeld theory of atoms the state $m = 0$ had to be excluded at times, and this always caused serious difficulty in the application of the 'principle of mechanical transformability.' From Heisenberg's report it appeared that such a situation did not occur in matrix mechanics, which aroused Pauli's interest to investigate this question in the case of the hydrogen atom.[216]

In the Bohr–Sommerfeld theory the hydrogen atom had been treated as having two degrees of freedom, with the electron orbiting along a Kepler ellipse in a plane. The integration of the equation of motion of the electron, expressed in polar coordinates, was made possible by determining two constants of motion: for instance, the angular and the radial momenta. This equation could not be genuinely re-interpreted in matrix mechanics because of the difficulty of representing the angle as a matrix. However, Pauli discovered a way out of this difficulty. He recalled that Wilhelm Lenz, his Hamburg colleague, had integrated the classical equation of motion of an electron of mass m_e, moving in the Coulomb field of a nucleus of charge Ze, with the help of a vector quantity, $\mathbf{A}$, defined as

$$\mathbf{A} = \frac{1}{Ze^2 m_e}(\mathbf{M} \times \mathbf{p}) + \frac{\mathbf{r}}{r}, \tag{191}$$

where $\mathbf{M}$ $(= M_x, M_y, M_z)$ is the angular momentum vector, $\mathbf{p}$ $(= p_x, p_y, p_z)$ the linear momentum vector, $\mathbf{r}$ $(= x, y, z)$ the position vector of the electron relative to the nucleus (with r its absolute value), and the product on the right-hand side denotes the 'cross' or 'vector' product (i.e., $(\mathbf{M} \times \mathbf{p})_x = M_y p_z - M_z p_y$, etc.). The

[216] It should be noted that, in the letters of October 1925, Pauli did not receive from Heisenberg a complete account of the matrix theory of angular momentum, such as was contained in Section 4.1 of the three-man paper. However, the information about the specific result was enough to stimulate Pauli to work out a theory of angular momentum in matrix mechanics for himself. He presented this theory in his published paper on the hydrogen atom (Pauli, 1926a, Section 3, pp. 348–349).

vector **A** is a conserved quantity in the classical Kepler problem and is related to W, the energy of the system, by the equation

$$1 - \mathbf{A}^2 = -\frac{2W}{Z^2 e^4 m_e}\mathbf{M}^2, \tag{192}$$

where $\mathbf{M}^2$ is the square of the total angular momentum of the electron. In the paper, in which Wilhelm Lenz presented his method, he had specifically remarked: 'The emphasis must be placed on the fact that the quantum states are obtained here without applying the method of separation and altogether without immediate assistance of the concepts of action and angle variables' (Lenz, 1924, p. 198). This was exactly what Pauli needed in quantum mechanics.

Several reasons may be thought of why Pauli considered using **A**, the 'Lenz vector,' in his treatment of the hydrogen atom.[217] First, he was familiar with Lenz' method and had just recently discussed it in his *Handbuch* article on quantum theory (Pauli, 1926b, pp. 133–137, 159–162). Second, since Lenz had successfully employed the vector **A** in his treatment of the problem of the hydrogen atom in crossed electric and magnetic fields within the framework of the Bohr–Sommerfeld theory (Lenz, 1924), Pauli expected to be able to do the same in matrix mechanics; then he could see whether the difficulties encountered previously would disappear or not. Probably a still deeper reason was the fact that, provided a proper re-interpretation of **A** existed in matrix language, the quantum-theoretical energy would be expressible in terms of observable quantities alone with the help of an equation which corresponds to the classical Eq. (192). In particular, no variables to which a physical significance could not be attached, such as the angle ϕ, would show up in intermediate steps of the calculation. Psychologically, quite another argument may have played an important role in Pauli's decision to make use of the vector **A**: he could thereby afford to ignore all the learned algebraic tools—including infinite quadratic forms and what not—which Born had assembled. Pauli had only to keep in mind that the dynamical variables in quantum theory had to be represented by Hermitian matrices; with that knowledge alone he could play the game of '*Knabenphysik*,'

[217] Wilhelm Lenz had used the vector **A**, which he called 'axial vector' ('*Achsenvektor*'), in his paper dealing with the simultaneous action of crossed electric and magnetic fields on the electron orbits of the hydrogen atom in the Bohr–Sommerfeld theory of atomic structure (Lenz, 1924, especially p. 198). There he had stated that for the free atom (i.e., without a perturbing electric or magnetic field) the vector **A** is constant in time, as follows 'from a little known, but easily sketched, method of integration of the equations of motion [of the electron in the hydrogen atom]' (Lenz, 1924, p. 198), referring at the same time to Carl Runge's book on *Vektoranalysis* (Runge, 1919, p. 68). Pauli, in his *Handbuch* article on quantum theory, referred to the use of the axial vector by Lenz ('*Zu diesem Zweck bilden wir mit Lenz den Achsenvektor* **A** . . . ,' Pauli, 1926b, p. 133).

The vector **A** had in fact been introduced a long time ago in connection with the integration of the motion of a celestial body in the gravitational field of the sun. While the earliest discussion may be found in the work of Jakob Hermann and Johann Bernoulli in 1710, Pierre Simon de Laplace presented in his *Traité de Mécanique Céleste*, Volume I (p. 165ff), the three constants of motion which make up the vector **A**. For a detailed account of the history of this vector quantity we refer to the articles of Herbert Goldstein (1975, 1976).

juggling suitably chosen classical dynamical relations until they fitted properly into matrix equations. In playing this young man's game he would, of course, abide by the rules concerning the quantum-theoretical re-interpretation of classical mechanical quantities, which his friend Heisenberg had discovered.[218] Under these conditions the application of matrix methods to the problem of the hydrogen atom was going to be intellectual fun, which Pauli was willing to enjoy.

It took Pauli less than three weeks to derive the Balmer formula purely by means of the methods of matrix mechanics.[219] His quick success converted him completely to the scheme of Born, Heisenberg and Jordan. Pauli now became convinced that this theory constituted a 'tremendous fundamental advance' ('*ungeheuren prinzipiellen Fortschritt,*' Pauli to Bohr, 17 November 1925). This new feeling of progress stimulated his scientific work greatly. Dispelled suddenly from his mind was the desperation of a few months ago when physics had seemed to him 'much too difficult' that he wished he had chosen the occupation of a 'film comedian' (Pauli to Kronig, 21 May 1925). In the wake of 'Heisenberg's mechanics' he had acquired 'hope and joy of living' ('*Die Heisenbergsche Mechanik hat mir wieder Lebensfreude und Hoffnung gegeben,*' Pauli to Kronig, 9 October 1925). The solution of the hydrogen problem had now proved the usefulness of 'Göttingen's deluge of [formal] learning' ('*Göttinger Gelehrsamkeitsschwall*'). For the time being, at least, Pauli's heart had been captured by matrix mechanics; it appeared to him that only by means of this theory could one hope to make further progress in the understanding of atomic phenomena. Quantum theory had a future again.

IV.5 The Solution of the Hydrogen Problem

In his lectures on problems of atomic dynamics at the Massachusetts Institute of Technology, Max Born described the key to the solution of the hydrogen problem in matrix mechanics as follows: 'In the classical theory of Keplerian motion it is customary to operate with polar coordinates. This process fails here because it does not seem possible to consider angular variables as matrices. Pauli avoided this difficulty by retaining rectangular coordinates and introducing an additional coordinate, the radius vector r' (Born, 1926d, Lecture 18; 1970 reprint, p. 113). Born implied with this statement that he and his collaborators in Göttingen had not thought of using the quantity corresponding to the magnitude r of the radius vector in classical theory, although it occurs explicitly in the expression for the Hamiltonian H describing the hydrogen atom. For Pauli the

[218] Pauli regarded Heisenberg as a congenial partner in the '*Knabenphysik,*' but not the much more senior Born. He was ultimately willing to accept Born's mathematical scheme, the matrix mechanics, only to the extent it was necessary in order to formulate Heisenberg's ideas about dynamical variables in a consistent mathematical language.

[219] Pauli had not begun to treat the hydrogen problem until he received Heisenberg's letter of 12 October. On 3 November Heisenberg acknowledged that he had received two letters from Pauli containing the solution (Heisenberg to Pauli, 3 November 1925).

necessity of introducing r into matrix mechanics was even greater because the expression for the vector **A** again involved it (see the right-hand side of Eq. (191)). Thus, the first step towards the solution of the hydrogen problem indeed consisted in defining the matrix r, which involved, especially, the derivation of the commutation relations of r with the dynamical variables of the system.[220]

Pauli knew that if two independent constants of motion were found, he could integrate the equation of motion of the hydrogen atom. For these, he decided to choose the angular momentum, $\mathbf{M} = (\mathsf{M}_x, \mathsf{M}_y, \mathsf{M}_z)$, and **A**, the quantum-theoretical reformulation of the vector given by Eq. (191) in classical theory. He defined the angular momentum by the expression, $m_e(\mathbf{r} \times \mathbf{v})$, which involves the vector product of the vector matrices $\mathbf{r} = (\mathsf{x}, \mathsf{y}, \mathsf{z})$ and $\mathbf{v} = (\dot{\mathsf{x}}, \dot{\mathsf{y}}, \dot{\mathsf{z}})$.[221] Pauli proved that the components of the angular momentum commute with H, the Hamiltonian matrix of the one-electron atom, hence they become constants of motion.[222] Then he wrote the classical vector **A** as the vector matrix **A**, where

$$\mathbf{A} = \frac{1}{Ze^2 m_e} \cdot \frac{1}{2}\left[(\mathbf{M} \times \mathbf{p}) - (\mathbf{p} \times \mathbf{M})\right] + \frac{\mathbf{r}}{\mathsf{r}}, \tag{191a}$$

and showed that its components, A_x, A_y and A_z, are constants of motion under the same condition as in classical theory, namely, that only central forces act on the electron.[223] However, the nonclassical properties of vector matrix **A** may be recognized from the fact that the vector product, $(\mathbf{A} \times \mathbf{r})$, is not the negative of the vector product, $(\mathbf{r} \times \mathbf{A})$, as in the case of the classical vectors **A** and **r**;

[220] Pauli defined r through the relation, $\mathsf{r}^2 = \mathsf{x}^2 + \mathsf{y}^2 + \mathsf{z}^2$, where x, y and z are the matrices describing the Cartesian coordinates of the electron. Then the matrix r commutes, of course, with all position coordinates, but not with the canonically conjugate momenta. Pauli wrote the commutation relations with the latter as (see Pauli, 1926a, p. 347, Eq. (34)): $\mathbf{p}\mathsf{r} - \mathsf{r}\mathbf{p} = (h/2\pi i)(\mathbf{r}/\mathsf{r})$, where $\mathbf{p}$ denotes the vector matrix of the momentum, i.e., $\mathbf{p} = (\mathsf{p}_x, \mathsf{p}_y, \mathsf{p}_z)$, and $\mathbf{r}$ denotes the vector matrix of the position, i.e., $\mathbf{r} = (\mathsf{x}, \mathsf{y}, \mathsf{z})$.

[221] As we have mentioned above, Pauli did not know about the angular momentum formalism which was developed at the same time by Heisenberg and Jordan, when he began to work on the hydrogen problem. Therefore, he had to derive independently the results contained in Chapter 4 of the three-man paper. From the presentation of the angular momentum formalism in Pauli's published paper (see Pauli, 1926a, Section 3), one finds that he occasionally emphasized different properties of the matrices M_x, M_y and M_z (which he denoted by P_x, P_y and P_z—see Pauli, 1926a, p. 348) than those mentioned in the paper of Born, Heisenberg and Jordan. In addition, Pauli reformulated in matrix language the two products of vector algebra, the scalar product and the vector product. He defined them in analogy with the classical definition: for instance, $\mathbf{r} \cdot \mathbf{M} = \mathsf{x}\mathsf{M}_x + \mathsf{y}\mathsf{M}_y + \mathsf{z}\mathsf{M}_z$, and $\mathbf{r} \times \mathbf{v} = (\mathsf{y}\dot{\mathsf{z}} - \mathsf{z}\dot{\mathsf{y}},\ \mathsf{z}\dot{\mathsf{x}} - \mathsf{x}\dot{\mathsf{z}},\ \mathsf{x}\dot{\mathsf{y}} - \mathsf{y}\dot{\mathsf{x}})$. The definition of the vector product $(\mathbf{r} \times \mathbf{v})$ is not ambiguous, because $(\mathbf{r} \times \mathbf{v})_x$ can be shown to commute with the matrix x, $(\mathbf{r} \times \mathbf{v})_y$ can be shown to commute with y, etc.

[222] The Hamiltonian matrix of the one-electron atom, whose nucleus bears the charge Ze, is obviously given by the expression $\mathbf{p}^2/2m_e - Ze^2/\mathsf{r}$. Note that the angular momentum components remain constants of motion if the second term in the Hamiltonian matrix is replaced by a central potential, i.e., by an arbitrary function of the matrix r.

[223] In proving the time independence of the vector matrix **A**, defined by Eq. (191a), Pauli made use of the fact that the time derivative of the matrix quotient $\mathbf{r}/\mathsf{r}$ is given by the expression $(2m_e)^{-1}[(\mathbf{M} \times \mathbf{r}/\mathsf{r}^3) - (\mathbf{r}/\mathsf{r}^3 \times \mathbf{M})]$. On the other hand, the time derivative of the first term on the right-hand side of Eq. (191a) becomes equal to $(2Ze^2m_e)^{-1}[(\mathbf{M} \times \dot{\mathbf{p}}) - (\dot{\mathbf{p}} \times \mathbf{M})]$ and, on substituting for $\dot{\mathbf{p}}$ the expression $-Ze^2\mathbf{r}/\mathsf{r}^3$, one finds that $\dot{\mathbf{A}}$ is indeed zero.

especially, the term in the square bracket can be rewritten as $2(\mathbf{M} \times \mathbf{p}) + 2(h/2\pi i)\mathbf{p}$, including an extra term, $(Ze^2m_e)^{-1} \cdot (h/2\pi i)\mathbf{p}$, which is not contained in the classical expression, Eq. (191).[224]

In spite of the differences between the classical and the quantum-theoretical variables, Pauli was convinced that the quantum-theoretical vector matrix, **A**, together with the angular momentum matrix, **M**, will provide the means for determining the energy eigenvalues of the hydrogen atom. He just went on to establish the commutation relations between these matrices, which, besides Eqs. (185), or

$$(\mathbf{M} \times \mathbf{M}) = -\frac{h}{2\pi i}\mathbf{M}, \tag{193}$$

included the equations

$$(\mathbf{A} \times \mathbf{M}) = -\frac{h}{2\pi i}\mathbf{A} \tag{194}$$

and

$$(\mathbf{A} \times \mathbf{A}) = \frac{h}{2\pi i}(Z^2e^4m_e)^{-1}2\mathsf{H} \cdot \mathbf{M}, \tag{195}$$

where H is the Hamiltonian of the system. With the help of Eqs. (193)–(195) Pauli immediately obtained the matrix re-interpretation of Eq. (192), relating the axial vector matrix **A** and the energy eigenvalues of the one-electron atom, as

$$1 - \mathbf{A}^2 = -(Z^2e^4m_e)^{-1}2\mathsf{H}\left(\mathbf{M}^2 + \frac{h^2}{4\pi^2}\mathbf{1}\right). \tag{192a}$$

(Here $\mathbf{A}^2 = (\mathbf{A} \cdot \mathbf{A}) = \mathsf{A}_x^2 + \mathsf{A}_y^2 + \mathsf{A}_z^2$ and $\mathbf{M}^2 = (\mathbf{M} \cdot \mathbf{M}) = \mathsf{M}_x^2 + \mathsf{M}_y^2 + \mathsf{M}_z^2$.) Again, an extra term arose on the right-hand side of Eq. (192a) as compared to the classical equation (192).

In the set of matrix equations, Eqs. (193)–(195) and (192a), Pauli had formulated the rules of the game he was about to play. He would need all his skill to obtain his desired goal: a formula expressing the eigenvalues of the Hamiltonian of the one-electron atom in terms of the eigenstates of the matrices **A** and **M**. The first difficulty arose already in the beginning of the calculation; it

[224] A difference from the classical situation also arises in the following equation: in classical theory, a vector **A** can be multiplied by the radius vector **r**, and the equation $(\mathbf{A} \cdot \mathbf{r}) = -(Ze^2m_e)^{-1}\mathbf{M}^2 + r$, represents a conic section. In quantum theory this equation must be reformulated as $\frac{1}{2}[(\mathbf{A} \cdot \mathbf{r}) + (\mathbf{r} \cdot \mathbf{A})] = -(Ze^2m_e)^{-1}(\mathbf{M}^2 + \frac{3}{2}(h^2/4\pi^2)\mathbf{1}) + \mathbf{r}$. Again there is an extra term, $-(Ze^2m_e)^{-1} \cdot \frac{3}{2}(h/2\pi)^2\mathbf{1}$, on the right-hand side of this equation; hence one recognizes immediately that the geometrical interpretation of the classical equation cannot be taken over into quantum theory.

had to do with the degeneracy of the system under investigation. Thus, Pauli remarked in his paper later on:

> From the existence of the vector matrix **A**, constant in time, we can infer that an atom with a single electron constitutes a degenerate system, even apart from the spatial orientation of the atom (a situation similar to the Kepler motion in classical mechanics). Namely, we can readily conclude from the relations derived above that in general $\mathbf{AM}^2 - \mathbf{M}^2\mathbf{A}$ cannot vanish. Since, on the other hand, $\mathbf{AH} - \mathbf{HA}$ vanishes, it is obvious that for a definite value of the energy H there is not just a single value of $\mathbf{M}^2$; the system is thus in fact degenerate. (Pauli, 1926a, p. 352)[225]

Now Pauli was, of course, familiar with the trick of how to deal with degenerate systems in the Bohr–Sommerfeld atomic theory. All one had to do in that case was to add a small perturbation term to the original (unperturbed) Hamiltonian and to claim that all results calculated from the perturbed Hamiltonian, in which the degeneracy did not occur, will pass over smoothly into those of the unperturbed case. Born and Pauli (1922) had based their perturbation theory on this device, and Heisenberg had used it again when he calculated the polarization of radiation emitted by a degenerate system (Heisenberg, 1925a). The same procedure could be followed in matrix mechanics; in fact, it was also applied in Göttingen by Born, Heisenberg and Jordan just when Pauli faced the problem of degeneracy in his treatment of the hydrogen atom. At the time, however, Pauli did not know the details of their algebraic methods—nor would he have cared to use them had he been fully aware of them, for he had railed enough already against the Göttingen formalism.[226] In any case, he chose to avoid the problem of degeneracy completely as had been done in the previous treatments. Thus, he added an appropriate perturbing term to the Hamiltonian of the free one-electron atom on purely physical grounds. In this way he removed the degeneracy of the problem artifically in order to be able to perform the calculation; however, he also made sure that the perturbing term altered the energy spectrum of the free atom only by an amount which could be neglected for practical purposes.

Pauli introduced two kinds of perturbing forces: a non-Coulomb central force and the directed force of a constant external magnetic field. These forces contributed an additional energy to the time average of λH_1, the first-order perturbation term in the Hamiltonian. Pauli explained the action of the two perturbing forces as follows: 'If, in particular, the perturbing field is due to an

[225] In the above quotation, we have replaced Pauli's notations **P** and **E** for the angular momentum and Hamiltonian matrices by **M** and **H**, respectively, in order to retain a consistent notation within the present chapters.

[226] In his published paper Pauli referred to the methods of Born, Heisenberg and Jordan concerning degenerate systems—i.e., to Chapters 2 and 3 of the three-man paper (Pauli, 1926a, p. 352). But he knew about them only after mid-November 1925 (when he received a copy of the manuscript from Heisenberg); at that time he had already performed his calculation of the hydrogen spectrum.

additional non-Coulomb central force, the above time average [i.e., $\lambda\bar{H}_1$] depends only on $\mathbf{M}^2$ (apart from its H-dependence), since there exists no preferred direction in space here. Further, the perturbation energy of a magnetic field in z-direction depends only on the momentum component $\mathbf{M}_z$, which is parallel to the field' (Pauli, 1926a, p. 353). These two perturbations thus completely remove the degeneracy of one-electron atoms and, as Pauli pointed out, they make it possible to discuss, at least in principle, certain more general questions which arise in the theory of these atoms: 'The requirement that $\mathbf{M}^2$ *and* $\mathbf{M}_z$ are to be *diagonal* matrices leads, therefore, to a special solution of Eqs. [(193)–(195), (192a)], adapted to the relativistic fine structure and to an additional weak magnetic field' (Pauli, 1926a, p. 353).[227] Although Pauli was not able to present the 'special solution' he had in mind in the published paper on the hydrogen atom, the fact, that the one-electron atoms can be perturbed in such a way that their degeneracy is removed and their energy eigenvalues are fully described by the eigenvalues of the matrices $\mathbf{M}^2$ and $\mathbf{M}_z$, played an important role in his calculation.[228]

In the particular case when the matrices $\mathbf{M}^2$ and $\mathbf{M}_z$ for the hydrogen atom were diagonal, Pauli assumed that the eigenvalues of $\mathbf{M}_z$ were given by[229]

$$M_z(j,m;\,j,m) = m\frac{h}{2\pi} \quad \text{where} - j \leq m \leq j. \tag{196}$$

It then followed from the commutation relation for the components of the angular momentum, contained in Eq. (193), that the matrices $\mathbf{M}_x$ and $\mathbf{M}_y$ could not be diagonal in the index m, the quantum number of $\mathbf{M}_z$. The only nonzero matrix elements of $\mathbf{M}_x$ and $\mathbf{M}_y$ were connected with transitions from the state m (of $\mathbf{M}_z$) to states $m \pm 1$. Pauli was able to calculate the squares of these matrix elements and, with them, the eigenvalues of the matrix $\mathbf{M}^2$, obtaining the result $j(j+1)(h/2\pi)^2$, in agreement with Born, Heisenberg and Jordan, Eq. (189).[230]

After he had made the fullest use of Eq. (193), Pauli set out to exploit Eqs. (194) and (195) for the purpose of determining the vector matrix **A**. This was a difficult task indeed and, in spite of his virtuosity in calculation, he might not

[227] In the above quotation we have again adapted Pauli's notation to the one used throughout the present chapter.

[228] Pauli submitted his paper for publication only in January 1926 because he had hoped all this time still to be able to include a calculation of the fine structure of the hydrogen atom.

It should be noted that Pauli, in dealing with the hydrogen atom in matrix mechanics, came closer to the spirit of Sommerfeld's discussion of the Balmer spectrum than to Bohr's original approach of 1913. He had just recently reviewed in detail Sommerfeld's theory, including the fine structure, in his *Handbuch* article (Pauli, 1926b, §§19–21). One may even say that his familiarity with *all* aspects of the earlier approaches—in short, Pauli's erudition—contributed importantly to the quick success of his hydrogen calculation in matrix mechanics.

[229] Pauli already knew this result from Heisenberg. (See Heisenberg to Pauli, 12 October 1925, quoted in Section IV.4.)

[230] It should be noted again that Heisenberg did not write Eq. (189) to Pauli earlier than his letter of 3 November 1925. Pauli denoted the quantum number of the total angular momentum by k rather than by j.

have gotten anywhere if he had not reminded himself of the (intensity) rules which had been obtained earlier by H. Hönl (1925) and S. Goudsmit and R. Kronig (1925) for the intensities of the Zeeman components. Since these intensities are characteristic of the squares of transition elements for an atomic system when a preferred direction, the z-axis, is given, their ratios should be identical to the ratios of the absolute squares of the corresponding matrix elements of **A**. Hence Pauli assumed that

$$\begin{aligned} |A_x(j+1,m;\, j,m\pm 1)|^2 &= |A_y(j+1,m;\, j,m\pm 1)|^2 \\ &= \tfrac{1}{4}C(j\mp m)(j\mp m+1), \\ |A_x(j,m;\, j+1,m\pm 1)|^2 &= |A_y(j,m;\, j+1,m\pm 1)|^2 \\ &= \tfrac{1}{4}C(j\pm m+1)(j\pm m+2), \\ |A_z(j+1,m;\, j,m)|^2 &= C\left[(j+1)^2-m^2\right] \end{aligned} \tag{197}$$

and discovered that these relations, together with the elements of the matrices, $\mathbf{M}_x$, $\mathbf{M}_y$ and $\mathbf{M}_z$, satisfied Eq. (194), provided C was a real, still undetermined coefficient which would depend on the index j (but not on m); thus $C = C_{j+1}$. The matrix elements of **A**, which do not occur in Eqs. (197), vanish; as, for instance, those in which the total angular momentum does not change (i.e., the transitions $j \to j$).[231]

To determine the numerical value of the coefficients C in Eqs. (197), Pauli referred to Eq. (195). This equation stated that the difference of matrix products, $\mathbf{A}_x\mathbf{A}_y - \mathbf{A}_y\mathbf{A}_x$, is a diagonal matrix in the indices or quantum numbers j and m, and it yielded a relation involving two adjacent coefficients, C_j and C_{j+1}, that is,

$$m\left[(2j-1)C_j - (2j+3)C_{j+1}\right] = m\,\frac{|W|h^2}{2\pi^2 Z^2 e^4 m_e}\,, \tag{198}$$

where $|W|$ is the absolute value of $H(j,m;\, j,m)$, the energy term.[232] Equation (198) was akin to a sum-rule, and Pauli made use of its contents by means of the following arguments. First of all, he looked for $j_{\min}$, the smallest value of the angular momentum quantum number, which occurs for a given *finite* energy W

[231] Equations (197) are identical with the equations expressing the transition intensities of the Zeeman components, given by Pauli in his *Handbuch* article (see Pauli, 1926b, p. 240, Eqs. (70_2) and (70_3)). In the case of the Zeeman effects, however, there also exist the transitions $j \to j$ with finite intensity (see Pauli, 1926b, p. 24, Eqs. (70_1)). In the case of the vector matrix **A** the corresponding elements are zero, as follows from the fact that **A** and $\mathbf{M}^2$ do not commute. (Hence the matrix $\mathbf{M}^2$ has only elements corresponding to the transition $j \to j$, but then the matrices **A** do not have any.)

[232] To obtain Eq. (198), Pauli substituted Eq. (196) for the eigenvalues of the angular momentum component in the z-direction and made use of the relation, $A_y(j,m;\, j',m\pm 1) = \pm iA_x(j,m;\, j',m\pm 1)$, between the elements of the matrix $\mathbf{A}_x$ and $\mathbf{A}_y$.

of the one-electron system. Evidently, the coefficient C associated with $j_{\min}$ must be zero because C_j is a factor connected with transitions $j \to j - 1$ and there is no state with $j_{\min} - 1$. By substituting $j = j_{\min}$ on the left-hand side of Eq. (198), one notices that it can be satisfied only by the specific value $m = 0$, else the two sides of the equation will assume different signs and will not match. Since, due to Eq. (196), m, the quantum number of $\mathbf{M}_z$, always ranges from $-j$ to $+j$, it follows that $j_{\min}$ must also be zero. However, by considering any energy value W above $H(0,0;\ 0,0)$, the ground state, one obtains from Eq. (198) the coefficients C_j, with $j = 1, 2, \ldots$, up to $j_{\max}$, the maximum value of the total angular momentum consistent with the energy W under consideration. From this argument Pauli derived the important conclusion that j and m are 'necessarily integral numbers' ('*Es sind demnach j und m notwendig ganzzahlig,*' Pauli, 1926a, p. 356), and that for a general energy state j runs from j to $j_{\max}$. He was thus able to exclude half-integral angular momenta in the case of one-electron atoms.[233]

In his calculation leading to Eqs. (196)–(198) Pauli had assumed that the degeneracy of the one-electron atom was removed by a small perturbation; on the other hand, in the discussion following Eq. (198), he associated different values of the total angular momentum, i.e., $j = 0, 1, 2, \ldots, j_{\max}$, with a single value W of the energy. Still the results did not contradict one another because the strength of the perturbing forces (which removed the degeneracy) did not enter anywhere in Eqs. (197) and (198); as a consequence the perturbation could be made arbitrarily small, and one was allowed to use Eqs. (196)–(198) in order to calculate the energy states of the one-electron atom from Eq. (192a). The left-hand side of the latter equation, taken for the matrix element denoted by $(j,m;\ j,m)$, reads[234]

$$1 - \mathbf{A}^2(j,m;\ j,m) = 1 - \frac{2|H(j,m;\ j,m)|}{Z^2 e^4 m_e}\left(\frac{h}{2\pi}\right)^2\left[j_{\max}^2 - j_{\max} - j(j+1)\right]. \tag{199}$$

On the right-hand side of Eq. (192a), Pauli had just to substitute the value $j(j+1)(h/2\pi)^2$ for the matrix $\mathbf{M}^2$; then it yielded the equation

$$H(j,m;\ j,m) = \frac{2\pi^2 Z^2 e^4 m_e}{h^2 (j_{\max} + 1)^2} \tag{200}$$

[233] In deriving consequences from the sum-rule, Eq. (198), Pauli followed Heisenberg's arguments, which the latter had used in evaluating the Thomas–Kuhn sum-rule in order to determine the constant of motion of the anharmonic oscillator (Heisenberg, 1925c, Section 3). As in Heisenberg's case, it was crucial for the solution of the hydrogen problem that an index, j, should assume a lowest value.

[234] The value of $\mathbf{A}^2(j,m;\ j,m)$ is obtained by adding the contributions, $A_x^2(j,m;\ j \pm 1, m \pm 1)$, $A_y^2(j,m;\ j \pm 1, m \pm 1)$ and $A_z^2(j,m;\ j \pm 1, m)$, which yields: $(j+1)(2j+3)C_{j+1} + j(2j-1)C_j$. On the other hand, the constant C_j has the value, $C(j_{\max} - j + 1)(j_{\max} + j + 1)(2j-1)^{-1}(2j+1)^{-1}$, with $C = 2|W|(Z^2 e^4 m_e)^{-1} h^2/4\pi^2$.

for the absolute value of the energy states of the one-electron atom. By identifying $(j_{\max} + 1)$ with n and denoting the energy value $H(j,m;j,m)$ appropriately by W_n, Pauli finally obtained Niels Bohr's formula of 1913, in which n is usually called the principal quantum number. And he concluded: 'Herewith it has been demonstrated that the Balmer terms come out correctly from the new quantum mechanics' (Pauli, 1926a, p. 357).

Besides Heisenberg, it was Niels Bohr who first learned about the results of Pauli's hydrogen calculation. Hendrik Kramers, who spent a few days in Hamburg in early November, had informed Bohr about it. Bohr immediately wrote to Pauli: 'To my great joy I heard from Kramers that you have succeeded in deriving the Balmer formula' (Bohr to Pauli, 13 November 1925). In the same letter Bohr asked for further details of the calculation and Pauli obliged without delay. Bohr and the theoreticians close to him found Pauli's success striking. Convinced that it was indeed true, Bohr wrote to Pauli: 'Kramers, Kronig and I, who have just gone once again with the greatest pleasure through your beautiful calculation on the hydrogen spectrum, send you many friendly greetings from Tisvilde' (Bohr to Pauli, 5 December 1925).

The great satisfaction which Heisenberg and Bohr expressed on Pauli's achievement arose for a reason which could be appreciated nowhere better than in Göttingen and Copenhagen. At these places the conclusion had been reached during the past three years that Bohr's apparently-so-successful theory of atomic systems, including his celebrated derivation of the Balmer formula (Bohr, 1913b), possessed no rational theoretical foundation whatsoever. However, Pauli's admirable calculation of the hydrogen spectrum, performed on the basis of the matrix mechanics, which could be regarded as the complete logical and mathematical formulation of the quantum concepts developed in Copenhagen and Göttingen since 1923, now provided the necessary justification of the fundamental formula of Bohr's theory. It was natural to ask whether matrix mechanics would be able to resolve the difficulties which had occurred earlier in the treatment of multiply periodic systems, and Pauli provided the answer in his matrix treatment of hydrogen-like atoms.

One of the great successes of the theory of multiply periodic systems had been the calculation of the Zeeman effect (Sommerfeld, 1916d; Debye, 1916a) and Stark effect (Schwarzschild, 1916; Epstein, 1916a, c) of the hydrogen atom. A difficulty had arisen later when one had discussed, especially since 1923, the simultaneous action of electric and magnetic fields on the one-electron atom (O. Klein, 1924a, b; Lenz, 1924). The term, which had to be added to the energy state W_n of the free atom in the case of the external fields, had been evaluated as

$$\Delta W_{n_1, n_2} = \left(\frac{n}{2} - n_1\right)\omega_1 h + \left(\frac{n}{2} - n_2\right)\omega_2 h, \tag{201}$$

where n_1 and n_2 assume integral values between 0 and n, the principal quantum number. In Eq. (201), ω_1 and ω_2 denote frequencies, composed of the Larmor frequency ν_L and the precession frequency ν_s of the Stark effect: in the case of

parallel magnetic and electric fields one has $\omega_1 = \nu_L + \nu_s$ and $\omega_2 = |\nu_L - \nu_s|$; in the case of crossed-fields one considers the vectors $\boldsymbol{\nu}_L$ and $\boldsymbol{\nu}_s$ in the direction of the magnetic field **H** and the electric field **E**, respectively, with magnitudes ν_L and ν_s, and one takes for ω_1 the absolute value of the vector sum, $|\boldsymbol{\nu}_L + \boldsymbol{\nu}_s|$, and for ω_2 the absolute value of the vector difference, $|\boldsymbol{\nu}_L - \boldsymbol{\nu}_s|$. The relation between the numbers n_1 and n_2, on one hand, and the quantum numbers m and s, associated with the Zeeman and Stark effects, on the other, had been found to be

$$m = n - (n_1 + n_2) \quad \text{and} \quad s = n_2 - n_1 \quad \text{for} \quad \nu_L > \nu_s$$

and (202)

$$m = n_2 - n_1 \quad \text{and} \quad s = n - (n_1 + n_2) \quad \text{for} \quad \nu_L < \nu_s.$$

In the Bohr–Sommerfeld theory of atomic structure the adiabatic hypothesis was valid; it claimed that through an adiabatic change one could always transform a possible state of an atomic system into another possible state. In particular, one could start from a situation, in which parallel magnetic and electric fields, of such strengths that $\nu_L > \nu_s$, act on a one-electron atom, and arrive after an adiabatic, i.e., infintely slow, change of strengths and directions of both fields (such that the frequencies ω_1 and ω_2 are always different from zero) at a situation, in which the fields are again parallel but $\nu_L < \nu_s$. According to Eqs. (202) the quantum numbers m and s exchange their values before and after the adiabatic transformation. Now, Sommerfeld and Bohr had argued that states of the atom with $m = 0$ and $s = m$ should not occur in their theory.[235] The difficulty connected with the one-electron atom in simultaneous magnetic and electric fields had then been the following: one may begin with an allowed electron orbit $s = 0$ and $m = n$ and transform it adiabatically into a forbidden orbit with $s = n$ and $m = 0$. This inconsistency of the theory of atomic structure had been considered to be a very serious one by the experts (see, e.g., Bohr, 1925b, pp. 850–851; Heisenberg, 1925c, p. 879).

When, in his letter to Pauli, dated 12 October 1925, Heisenberg reported that in matrix mechanics the state with the zero eigenvalue of the angular momentum M_z could not be excluded in principle, Pauli was very excited. The new theory, he thought, seemed to offer the opportunity of removing the difficulty with the crossed-field problem in the erstwhile atomic theory, which he had discussed himself at length just recently in his *Handbuch* article (Pauli, 1926b, Section 25, pp. 159–164). Hence he decided to repeat the calculations on that problem in matrix mechanics. The first part of it, the Zeeman effect of the one-electron atom, was quickly performed. The task was to find the eigenvalues of the matrix, $e/2m_e c(\mathbf{M} \cdot \mathbf{H})$, describing the perturbation term in the Hamiltonian. This matrix

[235] In his treatment of the Balmer spectrum of hydrogen, Sommerfeld had argued that electron orbits, in which the electron has zero angular momentum, must be excluded because they could come aribrarily close to the atomic nucleus (Sommerfeld, 1915b, pp. 445–446). From the condition, $j \neq 0$, Bohr had concluded that $m \neq 0$ (Bohr, 1918b, Section 5).

contains the scalar product of the vector matrix **M** and the magnetic field strength vector, **H**, a classical quantity (because **H** describes an external macroscopic field); it may also be written as $2\pi(\boldsymbol{\nu}_L \cdot \mathbf{M})$, where $\boldsymbol{\nu}_L$ is the vector quantity $(= e\mathbf{H}/4\pi m_e c)$, whose absolute value is the Larmor frequency ν_L. Evidently, the eigenvalues of the perturbation term could be expressed in terms of the eigenvalues of the matrix $\mathbf{M}_z$, i.e., the component of the angular momentum in the direction of the magnetic field **H**; if these are $m\ h/2\pi$, then the eigenvalues of the matrix $2\pi(\boldsymbol{\nu}_L \cdot \mathbf{M})$ or $2\pi\nu_L\mathbf{M}_z$ are just given by $h\nu_L m$, where the 'magnetic quantum number' m assumes the values $-(n-1), -(n-2), \ldots, +(n-1)$, n being the principal quantum number of the atom. Thus, Pauli obtained a normal Zeeman effect (because of the selection rule $m \to m \pm 1, m$), thereby reproducing quantum mechanically one of the results of the Bohr–Sommerfeld theory. The only difference from the older treatment was that, in agreement with Heisenberg's statement, the value $m = 0$ could not be excluded, and that m did not assume the values $-n$ and $+n$.

Similarly, Pauli discussed the Stark effect of the one-electron atom in an external electric field of strength **E** by examining—in the approximation in which terms proportional to $\mathbf{E}^2$ can be neglected—the eigenvalues of the matrix, $\frac{3}{2}n(\mathbf{A} \cdot \boldsymbol{\nu}_s)$, where **A** is the Lenz axial vector matrix, and $\boldsymbol{\nu}_s$ a quantity proportional to **E** (i.e., $\boldsymbol{\nu}_s = (3nh/8\pi^2 m_e eZ)\mathbf{E}$) whose absolute value is the precession frequency ν_s of the Stark effect.[236] In order to determine the perturbation term of the energy in the electric field **E**, Pauli had only to compute the eigenvalues of the component of the vector matrix **A**, say $\mathbf{A}_z$, which is parallel to the electric field **E**. He thus found the result[237]

$$A_z(s, m; s, m) = \frac{s}{n}\ , \tag{203}$$

where s assumes the integral values

$$s = -(j_{\max} - |m|), 2 - (j_{\max} - |m|), \ \ldots, (j_{\max} - |m|). \tag{203a}$$

[236] To arrive at the above result, Pauli reformulated the classical expression, $\Delta W_n = e(\mathbf{r}_{\mathrm{av}} \cdot \mathbf{E})$, for the perturbation term of the energy. The vector $\mathbf{r}_{\mathrm{av}}$ denotes the average position of the electron, i.e., the position **r** averaged over a complete period of the (nonperturbed) orbit. (Bohr called $\mathbf{r}_{\mathrm{av}}$ the position of the 'electric centre' of the atom: Bohr, 1918b, p. 70.) Lenz established the relation, $\mathbf{r}_{\mathrm{av}} = \frac{3}{2}a_n\mathbf{A}$, between $\mathbf{r}_{\mathrm{av}}$ and the axial vector **A**, where a_n is the semi-major axis of the Kepler ellipse having the quantum number n (Lenz, 1924, p. 199, Eq. (4)). (For the free atom a_n is given by the expression $\frac{1}{2}Ze^2|W_n|^{-1}$, where W_n is the energy of the atom in the state n.) Pauli demonstrated that in matrix mechanics the same relation holds between the corresponding matrices $\mathbf{r}_{\mathrm{av}}$ and **A** as in the classical theory by proving two facts: first, $\mathbf{r}_{\mathrm{av}}$ has the same commutation relations with the angular momentum matrices as the expression $\frac{3}{4}Ze^2|W_n|^{-1}\mathbf{A}$; second, the matrix $(\mathbf{r}_{\mathrm{av}} - \frac{3}{4}Ze^2|W_n|^{-1}\mathbf{A})$ commutes with **A**. (See Pauli, 1926a, pp. 253–254.) Hence he introduced the vector **A** instead of $\mathbf{r}_{\mathrm{av}}$ in the perturbation term $e(\mathbf{r}_{\mathrm{av}} \cdot \mathbf{E})$ and obtained $\frac{3}{4}Ze^3|W_n|^{-1}(\mathbf{A} \cdot \mathbf{E})$, which he finally wrote as $\frac{3}{2}n(\mathbf{A} \cdot \boldsymbol{\nu}_s) \cdot h$, on substituting the expression for ν_s.

[237] Equation (203) for the matrix element $A_z(s, m; s, m)$ was suggested by the previous atomic theory. Pauli had just to show that in quantum mechanics the same equation held provided one introduced a new index s which assumed the values given by Eq. (203a). In addition, he had also to calculate the elements of the matrices **M** and $\mathbf{A}_x$ and $\mathbf{A}_y$ in terms of the index pairs, (s, s') and (m, m') and to show that Eqs. (193)–(195) were satisfied. (See Pauli, 1926a, pp. 358–359.)

Again, there is a difference between the classical and the matrix calculations in the values admitted for the index s characterizing the Stark effect. In the former theory, s had assumed the values, $-(n-|m|), -(n(-|m|-2), \ldots, (n-|m|)$, with the subsidiary condition, $m \neq 0$; in quantum mechanics s is automatically bounded by $(n-1)$ because the maximum value of the angular momentum is not $nh/2\pi$ but $(n-1)h/2\pi$. In addition, contrary to the classical situation, the state with zero eigenvalue of $\mathbf{M}_z$ would not be excluded.

Finally, turning to the case of crossed magnetic and electric fields, Pauli followed closely the classical treatment of Wilhelm Lenz. In this way he obtained a formula for the eigenvalues of the perturbation term, $\Delta H(n_1 n_2;\ n_1 n_2)$ or $\Delta W_{n_1,n_2}$, which has a structure very similar to the classical formula, Eq. (201), that is,

$$\Delta W_{n_1,n_2} = \left(\frac{j_{\max}}{2} - n_1\right)\omega_1 h - \left(\frac{j_{\max}}{2} - n_2\right)\omega_2 h, \tag{204}$$

where n_1 and n_2 assume values between 0 and $j_{\max}(=n-1)$, and ω_1 and ω_2 are the same combinations of the Larmor frequency ν_L (or the vector $\boldsymbol{\nu}_L$) and the precession frequency ν_s (or $\boldsymbol{\nu}_s$) of the Stark effect as in the previous theory.[238] The only difference from the classical formula (201) is the occurrence of $j_{\max}$ rather than n, the principal quantum number of the one-electron atom. The relations between the indices m and s and n_1 and n_2 in Eq. (204) were now given by

$$\begin{aligned} &m = j_{\max} - (n_1 + n_2) \quad \text{and} \quad s = n_2 - n_1 \quad \text{for} \quad \nu_L > \nu_s, \\ &m = n_2 - n_1 \quad \text{and} \quad s = j_{\max} - (n_1 + n_2) \quad \text{for} \quad \nu_L > \nu_s. \end{aligned} \tag{204a}$$

Again, as in the case of the Zeeman or Stark effects, no specific value for s or n had to be excluded in the quantum-mechanical formula for the crossed-field problem.

In a long letter to Bohr, Pauli gave a detailed account of the status of atomic theory after the successful treatment of one-electron atoms in matrix mechanics. In it, he analyzed in particular the difficulty of the previous atomic theory, pointing out that

> the root of these difficulties is the arbitrary exclusion—from the point of view of the quantization rules—of pendulum orbits . . . This was the origin of all evil and the original sin of all considerations on the hydrogen atom, which rested on the

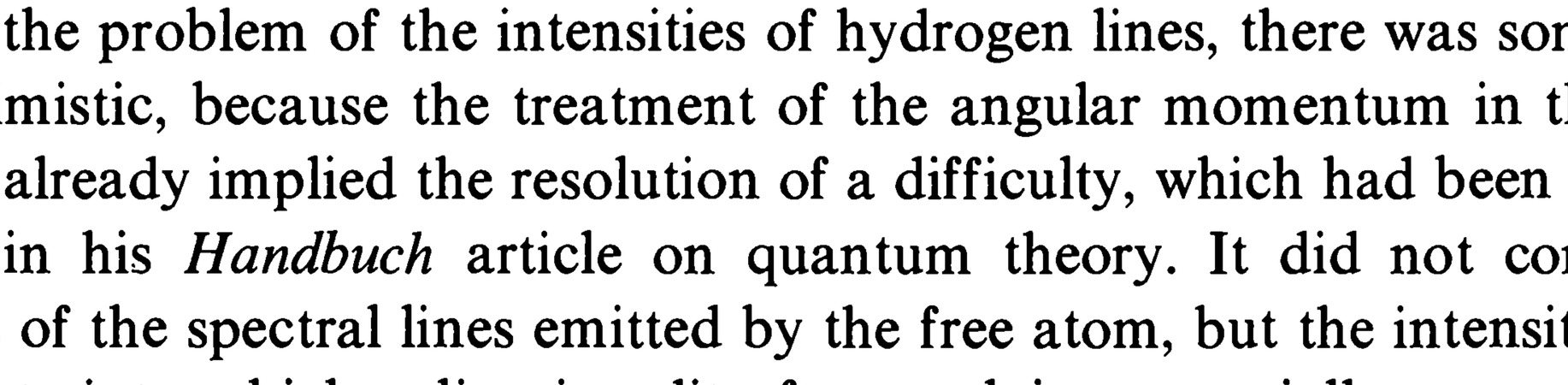
[238] To prove Eq. (204), Pauli introduced the vector matrices, $\mathbf{J}_1$ and $\mathbf{J}_2$, defined by $2\mathbf{J}_1 = (2\pi/h)\cdot \mathbf{M} + n\mathbf{A}$ and $2\mathbf{J}_2 = (2\pi/h)\mathbf{M} - n\mathbf{A}$ (following the analogous treatments by Klein and Lenz in the classical theory), and derived the commutation relations for them. He found that the components of $\mathbf{J}_1$ and $\mathbf{J}_2$ satisfy similar commutation relations as the components of $\mathbf{M}$, and that $\mathbf{J}_1^2$ and $\mathbf{J}_2^2$ have the eigenvalues $(j_{\max}/2)(j_{\max}/2+1)$. As a result, the eigenvalues of the component of $\mathbf{J}_1$, parallel to the vector $\boldsymbol{\omega}_1$, can be labelled by $(j_{\max}/2 - n_1)$, and the eigenvalues of the component of $\mathbf{J}_2$, parallel to the vector $\boldsymbol{\omega}_2$, can be labelled by $(j_{\max}/2 - n_2)$, where n_1 and n_2 run from 0 to $j_{\max}$. (See Pauli, 1926a, pp. 359–360.)

quantization rules for periodicity systems. The exclusion of $m = 0$ was only the logically necessary consequence of this first ban, and the adiabatic processes in crossed fields represent only the specially noticeable bad consequences of the exclusion of $j = 0$ in the case of the fine structure, which was necessary on one hand to describe the phenomenon, and was theoretically completely unfounded on the other . . . And yet, nobody at that time had been able to avoid the original sin of the exclusion of the pendulum orbit, in which I see a limitation of the applicability of the theory of periodic systems, since redemption from it can be achieved for no smaller price than by renouncing the mechanical, space-time pictures of the stationary states of the hydrogen atom. [And he concluded:] Since in the new quantum mechanics of Heisenberg such additional bans do not occur in the case of hydrogen, the difficulties in crossed fields also disappear. (Pauli to Bohr, 17 November 1925)

IV.6 The Problems of Intensities and the Diatomic Molecule

The calculation of the energy states of the hydrogen atom did more than convince Pauli about the correctness of matrix mechanics. He immediately began a serious search for new applications of the scheme. He knew that there were many questions concerning the hydrogen atom itself that still remained to be answered. For example, could one finally obtain intensities of the hydrogen lines? Or, could one calculate the relativistic corrections to the energy states given by Eq. (200)? On the other hand, as had happened before with Bohr's theory of 1913, the new results on the spectrum of hydrogen stimulated the hope that one might be able to extend the matrix methods so as to deal with systems containing several electrons, especially the helium atom. In November 1925 the prospects for making rapid progress with the help of a consistent quantum theory of atomic structure appeared to be very bright, and Pauli, without losing any time, immediately attacked the problems facing him.

As for the problem of the intensities of hydrogen lines, there was some reason to be optimistic, because the treatment of the angular momentum in the hydrogen atom already implied the resolution of a difficulty, which had been described by Pauli in his *Handbuch* article on quantum theory. It did not concern the intensities of the spectral lines emitted by the free atom, but the intensities of the components into which a line is split after applying an axially symmetric field (Pauli, 1926b, pp. 164–165). In the Bohr–Sommerfeld theory of multiply periodic systems, the transitions in one-electron atoms from states with the principal quantum number $n = 2$ to states with $n = 1$ occurred according to *Scheme* (a): If



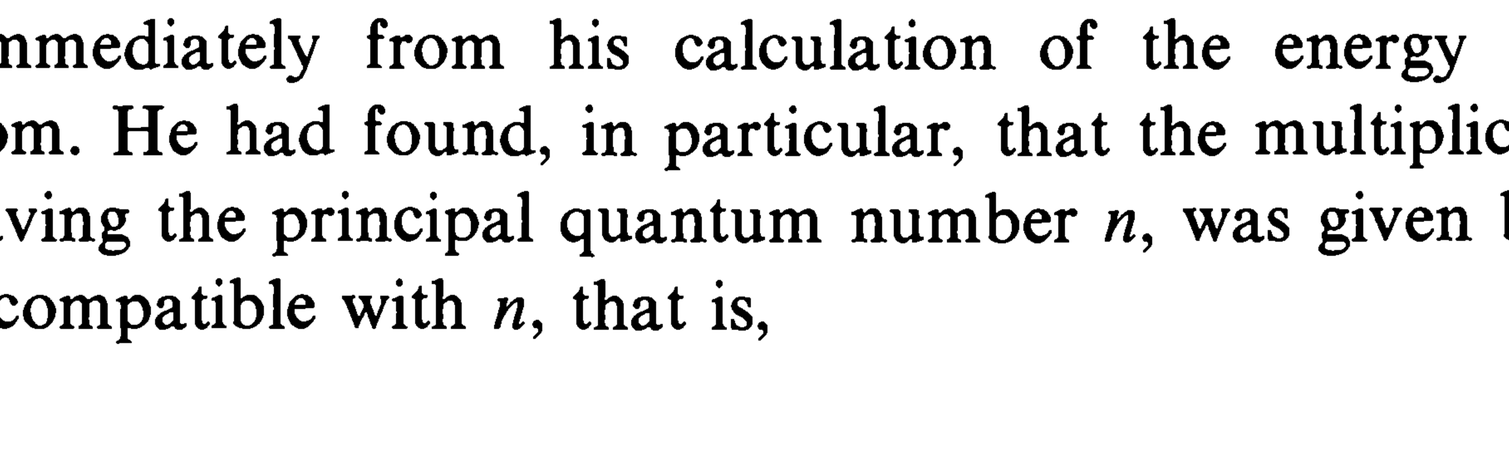
Scheme (a) Transition from the states with $n = 2$ to the states with $n = 1$.

no external field is present, all four transitions shown in *Scheme* (a) contribute to the same spectral line, which might appear as a resonance line. However, under the action of an axial field (say in z-direction), the transitions $(n, m) = (2, -2)$ to $(n, m) = (1, -1)$ and $(n, m) = (2, 2)$ to $(n, m) = (1, 1)$ give rise to two separated components, polarized circularly in opposite directions and shifted from the original line by the same distance, while the transitions $(n, m) = (2, -1)$ to $(n, m) = (1, -1)$ and $(n, m) = (2, 1)$ to $(n, m) = (1, 1)$ produce a linearly polarized, unshifted line component. The transition from the state $(n, m) = (2, 0)$ to the state $(n, m) = (1, 0)$ is, of course, excluded. Assuming that the emitted spectral line is unpolarized in the absence of the external field, Pauli had concluded from the above scheme that the intensity of each of the circularly polarized transition line components (with $\Delta m = \pm 1$) is as large as the sum of the intensities of the two linearly polarized transition line components (with $\Delta m = 0$). However, correspondence arguments demanded that the intensities of all transition lines should be the same. Since the two statements contradicted each other, Pauli had concluded that the transition *Scheme* (a) 'does not correspond to reality, and that also in atoms with a single electron combinations via the middle [i.e., transitions from and to states with $m = 0$] might occur in reality' (Pauli, 1926b, p. 165).

What Pauli had demanded earlier, now actually happened in matrix mechanics; i.e., the *Scheme* (a) had to be replaced by another scheme, *Scheme* (b), in which transitions occurred from the states $(n, m) = (2, \pm 1)$ and $(n, m) = (2, 0)$ to the ground state $(n, m) = (1, 0)$, and $m = 0$ to $m = 0$ transitions were especially not forbidden. In *Scheme* (b) all three components with $\Delta m = 0, -1$ and $+1$ have the same intensity, if the radiation was linearly polarized in the absence of the external field. And this was in full agreement with the correspondence arguments.

Much more involved than the intensity problem of the hydrogen lines appeared to be the problem of the structure of many-electron atoms. Pauli recognized this immediately from his calculation of the energy spectrum of the hydrogen atom. He had found, in particular, that the multiplicity of the energy state, W_n, having the principal quantum number n, was given by the sum of all states (j, m) compatible with n, that is,

$$\sum_{j=0}^{n-1} \sum_{m=-j}^{+j} m = \sum_{m=-j}^{n-1} (2j + 1) = n^2. \tag{205}$$



Scheme (b) Transitions from the states with $n = 2$ to the state $n = 1$.

Hence it seemed to follow from quantum mechanics that the state with $n = 1$ could be assumed by only one electron, the state with $n = 2$ by four electrons, the state with $n = 3$ by nine electrons, etc. This result could not be correct in the light of atomic structure and, while reporting this matter to Bohr, Pauli wrote:

> However beautiful and closed [i.e., internally consistent] the result of applying Heisenberg's quantum mechanics to the hydrogen atom—which especially leads to the consequence that the ground state of the hydrogen atom should exhibit no magnetic moment—may be, I still do not consider it to be final. To be sure, it would be logically possible that the features of atomic structure, which you have summarized under the notation of "stress" ("*Zwang*"), first occur in atoms having several electrons. Nevertheless, I have the definite feeling (which I cannot prove by logical arguments) that the difference between the atoms with one electron and those with several electrons with respect to the existence of the stress has often been enormously exaggerated. In particular, I believe that one would, if one performed the experiment of deflecting hydrogen atomic beams in inhomogeneous magnetic fields (unfortunately, so far I could not convince Stern to carry it out), see the same thing as in the case of the Ag-atom, namely ± 1 magneton and *not* non-magnetic atoms. In general I believe that reality deviates from the results, which can be derived from the present basis of Heisenberg's theory for the hydrogen atom, insofar as that for each orbital term $E_n = -Rh/n^2$ there exist instead of n^2 states in reality $2n^2$ states, such that for each energy value in the "stressless" ("*zwanglosen*") Heisenberg theory there arises in a magnetic field the additional term, ± 1 times the Larmor frequency times h, thereby doubling the number of terms. (Pauli to Bohr, 17 November 1925)

These statements gave the outline of a systematic programme for developing the theory of atomic systems within the framework of matrix mechanics. An immediate task in this programme was to complete the discussion of the hydrogen atom including the intensity problem and the relativistic corrections. The next most obvious and 'elementary' task seemed to be to apply the methods worked out by Born, Heisenberg and Jordan, and Pauli himself, to calculate the energy states of the helium atom. Pauli knew that for all these problems he needed an extension of the existing matrix scheme. For example, in the case of the helium atom there was no trick available—similar to the use of the axial (Lenz) vector in the hydrogen atom—which would help to avoid the use of action and angle variables. Moreover, he believed firmly that one had to bring into quantum mechanics the features described formerly by the notion of 'non-mechanical stress.' In November 1925, right after the calculation of the hydrogen states, he started to work hard on this programme. However, in spite of his efforts, in which he was soon joined by Heisenberg, progress could not be made as rapidly as he had hoped. In order to perform such tasks as the calculation of the relativistic hydrogen atom or even the helium atom, matrix mechanics had to be modified and extended considerably. Before that could be done, however, it was possible to treat only one other atomic problem: the diatomic molecule.

Just when Pauli became increasingly involved in treating the problems of atomic mechanics by using matrix formalism, Lucie Mensing, a former student

of Wilhelm Lenz' from Hamburg, learned about the new quantum mechanics in Göttingen and became interested in applying it to calculate certain properties of the spectrum of diatomic molecules.[239] The motion of diatomic molecules could be described by means of very simple mechanical models: the rotator and the anharmonic oscillator. In order to treat these models in quantum mechanics one did not have to invent any new clever technique, such as the one Pauli had used for diagonalizing the Hamiltonian matrix of the hydrogen atom; all the necessary tools were already available in the papers of the Göttingen theoreticians (anharmonic oscillator: Born and Jordan, 1925b; rotator: Born, Heisenberg and Jordan, 1926). Thus Mensing's work on the diatomic molecule (Mensing, 1926a), apart from Pauli's calculation of the spectrum of hydrogen, represented the earliest application of matrix mechanics to atomic problems.

The dynamical behaviour of a molecule consisting of two atoms (masses m_1 and m_2) may be described by a Hamiltonian matrix H, where

$$\mathsf{H} = \frac{1}{2m_0}\left(\mathsf{p}_x^2 + \mathsf{p}_y^2 + \mathsf{p}_z^2\right) + \mathsf{U}(\mathsf{r}), \tag{206}$$

which is the matrix re-interpretation of the corresponding classical Hamiltonian. In Eq. (206), m_0 is the so-called reduced mass of the system, that is, the two atoms bound together in a molecule have been replaced by a single object having the effective mass m_0 (with $m_0^{-1} = m_1^{-1} + m_2^{-1}$, where m_1 and m_2 denote the masses of the two atoms, respectively); p_x, p_y and p_z are the components of the momentum. The potential energy, $\mathsf{U}(\mathsf{r})$, analogous to its classical counterpart, U(r), is supposed to depend only on the matrix quantity r $(= \sqrt{\mathsf{x}^2 + \mathsf{y}^2 + \mathsf{z}^2}$, where x, y and z are the Cartesian coordinates of the effective mass). In the case of actual diatomic molecules, the potential energy had often to be expanded in a power series of the parameter $(r - a)$, where a denotes the equilibrium distance of the effective mass in a state when no rotation occurs.[240] Now the matrix $\mathbf{a}$,

[239] As Pascual Jordan recalled later: 'Miss Mensing was a physicist, who also worked in Göttingen. On our suggestion, she made an application, the first application of the new quantum mechanics, to—I think—a diatomic molecule' (Jordan AHQP Interview, Second Session, 18 June 1963). This implied that Mensing had begun to work on the diatomic molecule before Pauli's calculation of the hydrogen states became known in Göttingen. Her paper on this subject was received only much later, on 29 March 1926, by *Zeitschrift für Physik* (Mensing, 1926a); hence she could refer in it to certain results of Pauli's paper (1926a).

Lucie Mensing, born in Hamburg on 11 March 1901, studied physics at the University of Hamburg, where Wilhelm Lenz became her supervisor. She submitted her first paper on an application of perturbation theory to problems of molecular structure as late as September 1925 (Mensing, 1925a), although she had obtained the results already in winter 1923–1924. In 1925 she received her doctorate at Hamburg with a thesis on the influence of electric fields on the widths of spectral lines (Mensing, 1925b). After that she spent some time in Göttingen and became interested in applying matrix mechanics to the problems of molecules. Later, in 1926, she continued the work of Heisenberg and Jordan (1926) on the anomalous Zeeman effects (Mensing, 1926b). Lucie Mensing married the physicist Wilhelm Schütz.

[240] The classical theory of diatomic molecules had been treated, for example, by Max Born, in his *Vorlesungen über Atommechanik* to which Lucie Mensing referred (see Born, 1925, §20, pp. 140–148).

which replaces the classical equilibrium distance, is a diagonal matrix having all its elements equal to a constant a. Mensing then introduced new dynamical variables, the position matrix, $\mathsf{q} = (\mathsf{r} - \mathsf{a})$, and the canonically conjugate momentum, $\mathsf{p} = m_0\dot{\mathsf{r}}$, and rewrote H, the Hamiltonian of the diatomic molecule, Eq. (206), as

$$\mathsf{H} = \frac{1}{2m_0}\left[\mathsf{p}^2 + \frac{\mathsf{M}^2}{\mathsf{a}^2}\left(1 + \frac{\mathsf{q}}{\mathsf{a}}\right)^{-2}\right] + \mathsf{U}_0 + \alpha_2\mathsf{q}^2 + \alpha_3\mathsf{q}^3 + \ldots . \qquad (206a)$$

In Eq. (206a) M^2 denotes the square of the total angular momentum, U_0 the diagonal matrix $\mathsf{U}(\mathsf{a})$ whose elements are all equal, α_2 the (equal) elements of the diagonal matrix, $\frac{1}{2}(d^2\mathsf{U}(\mathsf{r})/d\mathsf{r}^2)_{\mathsf{r}=\mathsf{a}}$, and α_3 the (equal) elements of the diagonal matrix, $\frac{1}{6}(d^3\mathsf{U}(\mathsf{r})/d\mathsf{r}^3)_{\mathsf{r}=\mathsf{a}}$; the dots indicate terms multiplied by powers of q higher than the third.[241] By making this change in the dynamical variables, the system was reduced to a special one, which could be easily handled in matrix mechanics: on one hand, one could make use of the angular momentum methods of Born, Heisenberg and Jordan (1926) or Pauli (1926a); on the other, one could apply the perturbation methods of Born and Jordan (1925b) or Born, Heisenberg and Jordan (1926), if one assumed, as in classical theory, that q was a small quantity with elements much smaller than a, i.e., that the actual position of the centre of mass of the two atoms in the molecule never deviated appreciably from the equilibrium position.[242]

Since the total angular momentum had been proved to be a constant of motion, Mensing considered separately the specific submatrices of the Hamiltonian associated with a particular value (denoted by the index j) of the angular momentum, H_j. Then M^2 had to be replaced by $j(j+1)h^2/4\pi^2$ times the unit matrix $\mathsf{1}_j$, and H_j became equivalent to the Hamiltonian of a one-dimensional anharmonic oscillator.[243] This result, however, was not surprising at all; it simply indicated what in the classical model had been regarded as the oscillations of the

---

Mensing followed Born's treatment, though she did not take over the notation completely. Thus, she expanded the potential $\mathsf{U}(\mathsf{r})$ in a power series, $\mathsf{U}(\mathsf{r}) = \mathsf{U}_0 + \alpha_2(\mathsf{r}-\mathsf{a})^2 + \alpha_3(\mathsf{r}-\mathsf{a})^3 + \cdots$, where $\mathsf{U}_0 = \mathsf{U}(a)$, $\alpha_2\mathsf{1} = \frac{1}{2}(d^2\mathsf{U}(\mathsf{r})/d\mathsf{r}^2)_{\mathsf{r}=\mathsf{a}}$, and $\alpha_3\mathsf{1} = (1/3!)(d^3\mathsf{U}(\mathsf{r})/d\mathsf{r}^3)_{\mathsf{r}=\mathsf{a}}$. (Evidently, $(d\mathsf{U}(\mathsf{r})/d\mathsf{r})_{\mathsf{r}=\mathsf{a}}$ must have only zero elements because $\mathsf{r} = \mathsf{a}$ denotes the equilibrium position.) This expansion agrees with Born's (1925, p. 143, Eq. (8)) in the correspondence limit, if one relates Mensing's parameters α_2 and α_3 to Born's ω and a as follows: $\alpha_2 = (m_0/2)\omega^2$, $\alpha_3 = a$. (Note that in Mensing's expansion of $\mathsf{U}(\mathsf{r})$—Mensing, 1926a, p. 815, Eq. (1)—the matrices $(\mathsf{a}-\mathsf{r})$, $(\mathsf{a}-\mathsf{r})^2$ and $(\mathsf{a}-\mathsf{r})^3$ appeared, rather than $(\mathsf{r}-\mathsf{a})$, $(\mathsf{r}-\mathsf{a})^2$ and $(\mathsf{r}-\mathsf{a})^3$.)

[241] The matrix $\mathsf{M}^2/\mathsf{a}^2 \cdot (1 + \mathsf{q}/\mathsf{a})^{-2}$ enters Eq. (206a) because of rewriting $(\mathsf{p}_x^2 + \mathsf{p}_y^2 + \mathsf{p}_z^2)$ in Eq. (206) as $\mathsf{p}^2 + \mathsf{M}^2/\mathsf{r}^2$ and replacing r^2 by $\mathsf{a}^2(1 + \mathsf{q}/\mathsf{a})^2$. Further, the potential energy matrix $\mathsf{U}(\mathsf{r})$ is expanded in powers of the matrix q, as indicated in footnote 240. Note that q/a is equal to q/a, where a is the diagonal element of the matrix a; similarly, $\mathsf{M}^2/\mathsf{a}^2$ may be replaced by M^2/a^2.

[242] In this case one can expand the matrix $(1 + \mathsf{q}^2/\mathsf{a})^{-2}$ in powers of q/a and write H as $(1/2m_0)[\mathsf{p}^2 + \mathsf{M}^2/a - 2\mathsf{M}^2\mathsf{q}/a^3 + \mathsf{q}^2(2\alpha_2 m_0\mathsf{1} + 3\mathsf{M}^2/a^3) + \mathsf{q}^3(2\alpha_3 m_0\mathsf{1} - 4\mathsf{M}^2/a^5)] + \mathsf{U}_0 + \cdots$. (Note that M^2 commutes with q, q^2, etc.)

[243] Technically, Mensing looked at the matrices as four-dimensional objects in two index pairs (j, j') and (n, n'), where the second index pair described the nonrotational degrees of freedom.

two nuclei towards each other. By applying the matrix method of treating the anharmonic oscillator with the help of the reformulated perturbation scheme, Mensing found that $W_{n,j}$, the energy states of the diatomic molecule belonging to a fixed angular momentum (characterized by the index j), were given by the equation

$$W_{n,j} = U_0 + \frac{h^2}{8\pi^2 m_0 a^2} j(j+1) + h\left(n + \frac{1}{2}\right)\left[\nu_0 + \beta j(j+1)\right] + \alpha h^2\left[n(n+1) + \frac{11}{30}\right] + \ldots, \tag{207}$$

where U_0 is the diagonal element of the matrix U_0 in Eq. (206a), n the index characterizing the nuclear oscillations, and ν_0, β and α are constants for a given j.[244] The energy expression, Eq. (207), corresponded perfectly to the classical one. In the latter, the factor $j(j+1)$ in the second, the so-called rotational, term had been j^2; the third term had been $hn(\nu_0 + \beta j^2)$ and had been associated with the harmonic part of the oscillations of the nuclei and the mixed rotational–oscillational motion, respectively; and the fourth term, taking into account the anharmonicity of the oscillations, had been $h\alpha n^2$ (see, e.g., Born, 1925, p. 145, Eq. (12)). Thus, the result, Eq. (207), represented the quantum-theoretical formula describing all energy states of the diatomic molecule, except the excitations of the electrons.

Lucie Mensing did not stop at this point. She discovered further that only the transitions associated with the changes, $\Delta j = \pm 1$ of the rotational index j, and $\Delta n = 0, \pm 1, \pm 2, \ldots$ of the vibrational index n, had finite intensities, while transitions with $\Delta j = 0$ in the free molecule could not be observed.[245] On the other hand, the matrix treatment of angular momentum required integral 'quantum numbers' j and left no room for speculation about half-quanta. Both results, however, sufficed to remove a difficulty which had occurred in the previous theory of band spectra and had led to the introduction of half-integral values for the quantum number j.[246] Finally, Mensing noticed that in matrix theory the band lines of diatomic molecules should not exhibit a linear Stark effect—

[244] The constants ν_0, β and α appearing in Eq. (207) may be expressed in terms of the constants α_2, α_3, m_0, a and the index j (Mensing, 1926a, p. 819, Eqs. (22) and (24)).

[245] The reason for this lay in the fact that the diatomic molecule was treated as a one-body system; hence the matrix z of the coordinate perpendicular to the plane of motion had only zero elements, and consequently the intensities of the $\Delta j = 0$ transitions vanished.

[246] Mensing found, in particular, that the formula for the frequencies, $\nu(n, j; n', j-1)$ and $\nu(n, j-1; n', j)$, associated with finite intensities, can also be described by a classical formula involving half-integral rotational quantum numbers j (Mensing, 1926a, p. 820). Then she showed that the intensities of the band lines symemtric to the 'zero-intensity' line, i.e., the intensities of the transitions $(n, j; n', j-1)$ and $(n, j-1; n', j)$ are approximately equal (Mensing, 1926, pp. 821–823).

because it would be proportional to $\bar{z}$, the average of the z-coordinate of the unperturbed motion, whose elements vanish altogether—in agreement with observation (Mensing, 1926a, p. 823). These clear and straightforward conclusions pleased Pauli, who knew Lucie Mensing rather well—he had supervised her thesis during Wilhelm Lenz' illness—and he decided to collaborate with her on another problem of molecular physics: the calculation of the electric polarizability of diatomic molecules.

The reason why Pauli was interested in this particular question was that he had evaluated the electric polarizability and, with it, the temperature dependence of the dielectric constant of diatomic molecules a few years earlier by using the Bohr–Sommerfeld theory of multiply periodic systems. He had arrived at an expression (Pauli, 1921a, p. 327, Eq. (27)) which, in the limit of high temperatures, was larger than the classical expression of Debye (1912a) by a factor of five. Evidently, this situation contradicted the correspondence principle, in whose validity Pauli, in fall 1925, believed more than ever before.

The quantum-mechanical treatment of a diatomic molecule in the presence of an electric field of strength **E** (absolute value E), acting along the z-direction, was straightforward. The field perturbs the motion of the molecule in such a way that an average electric dipole moment, $\bar{\mu}_z$ $(= e\bar{z})$, which is proportional to E, may be induced, that is,

$$\bar{\mu}_z = \alpha E. \tag{208}$$

The matrix $\boldsymbol{\alpha}$ in Eq. (208) denotes the electric polarizability of the molecule.[247] Its elements could be determined by considering $\bar{z}$, the average of the z-component over a period of the (unperturbed) motion in the approximation which is linear in the field strength **E**; hence the $\alpha(j, m)$ could be obtained from a special case of Kramers' dispersion formula (in the limit of vanishing frequency of the incident radiation). Thus, Mensing and Pauli found the following matrix elements:

$$\alpha(0,0) = \frac{1}{3}\,\frac{8\pi^2 A\mu^2}{h^2} \qquad (\text{for } j = 0), \tag{209a}$$

$$\alpha(j,m) = \frac{1}{(2j-1)(2j+3)}\left[\frac{3m^2}{j(j+1)} - 1\right]\frac{8\pi^2 A\mu^2}{h^2} \qquad (\text{for } j \neq 0), \tag{209b}$$

where A is the moment of inertia of the molecule (equal to $m_0 a^2$ in Mensing's

[247] This equation does not contradict Mensing's statement that the diatomic molecule exhibits no linear Stark effect because the perturbation term in the Hamiltonian becomes $-\frac{1}{2}\alpha(\mathbf{E})^2$, which is quadratic in the electric field strength **E**.

Note that in their treatment Mensing and Pauli left out the oscillational degree of freedom of the molecule.

former notation) and μ is the (constant) dipole moment of the molecule.[248] Now Mensing and Pauli substituted expressions (209) into the equation defining the quantity $\overline{\overline{\mu}}_z$, the statistical average of the induced electric moment in equilibrium at temperature T, that is,

$$\overline{\overline{\mu}}_z = \frac{\sum_{j=0}^{\infty} \sum_{m=-j}^{+j} \bar{\mu}_z(j,m)\exp(-W_j/kT)}{\sum_{j=0}^{\infty} \sum_{m=-j}^{+j} \exp(-W_j/kT)}, \tag{210}$$

where W_j $(= (h^2/8\pi^2 A)j(j+1))$ denotes the energy eigenvalue of the rotating molecule and k is Boltzmann's constant. Then they observed that only the lowest state with angular momentum index $j = 0$ contributed to the right-hand side of Eq. (210), the reason being that the sums, $\sum_{m=-j}^{+j} \alpha(j,m)$, vanished for $j \neq 0$ (because $(3/(2j+1))\sum_{m=-j}^{+j} m^2 = j(j+1)$ and, therefore, the terms within the square brackets in Eq. (209b) cancelled). In the limit of high temperatures $\overline{\overline{\mu}}_z$ became equal to $\frac{1}{3}\mu^2 kT \cdot E$, which agreed perfectly with the classical result.[249]

[248] The matrix elements of $\boldsymbol{\alpha}$ follow from the formula

$$\alpha(j,m) = \frac{2}{h}\left[\frac{|ez(j,m;\, j+1,m)|^2}{\nu(j+1,j)} - \frac{|ez(j,m;\, j-1,m)|^2}{\nu(j,j-1)}\right],$$

where $z(j,m;\, j \pm 1,m)$ are the matrix elements of the position coordinate describing the unperturbed motion and $\nu(j, j-1)$ the rotational frequency $(= (h/4\pi^2 A)j)$. (To arrive at the above equation, one takes Kramers' dispersion formula for zero frequency of the incident radiation, and notes that in each of the sums over emission and absorption transitions due to the selection rule $\Delta j = \pm 1$ only one term has to be kept. In case of zero angular momentum, of course, just the first term of the sum exists.) Now the numerators had already been calculated by Lucie Mensing in her previous paper (Mensing, 1926a, pp. 821–822, Eqs. (30) and (34)) as

$$|ez(j,m;\, j+1,m)|^2 = \frac{(j+1)^2 - m^2}{(2j+1)(2j+3)}\,\mu^2$$

and

$$|ez(j,m;\, j-1,m)|^2 = \frac{j^2 - m^2}{(2j-1)(2j+1)}\,\mu^2,$$

where μ^2 is a constant (which Mensing denoted by $r^2(n,n')$) in case the oscillatory motions are neglected. The quantity μ can be identified with the electric dipole moment of the diatomic molecule.

[249] In the classical theory, $(\overline{\overline{\mu}}_z/|\mathbf{E}|)N$ (where N is the number of molecules per unit volume), which is equal to the quantity $(3/4\pi)((\epsilon-1)/(\epsilon+1))$ (i.e., proportional to refractivity, where ϵ is the dielectric constant), can be written as $(\text{const.} + C\mu^2/kT)N$, and C assumes the value $\frac{1}{3}$ (Debye, 1912a, pp. 98–100, Eqs. (7) and (11); see also the remark about units on p. 100). In quantum theory, according to Mensing and Pauli, $(\overline{\overline{\mu}}_z/|\mathbf{E}|)\,N$ becomes equal to

$$\frac{1}{3}\,\frac{\mu^2 N}{kT}\left(\sigma \sum_{j=0}^{\infty}(2j+1)\exp[-\sigma j(j+1)]\right)^{-1},$$

where $\sigma = h^2/8\pi^2 AkT$. For large temperatures, σ is a small quantity, and the sum in the above expression may be changed into the integral, $\int_0^\infty \sigma(2j+1)\exp[-\sigma j(j+1)]\cdot dj$, which assumes the value $1\,(= \int_0^\infty \exp(-x)dx)$; thus $(\overline{\overline{\mu}}_z/|\mathbf{E}|)\,N$ becomes $\frac{1}{3}\mu^2 N/kT$, yielding the classical result $(C = \frac{1}{3})$.

Although Pauli's earlier result had not yet been disproved empirically, Mensing and Pauli were convinced that the data would be represented correctly by the quantum-mechanical result.[250] They concluded: 'Hence it becomes obvious that here, as in some other cases, the new quantum mechanics comes—with respect to the statistical average values—closer to the results of classical mechanics than the erstwhile quantum theory' (Mensing and Pauli, 1926, p. 512).

Independently of Lucie Mensing and Wolfgang Pauli, the problem of treating the diatomic molecule according to the new quantum mechanics was taken up in 1926 by several authors in Europe and the United States. They not only worked out the energy states of the rotator in matrix mechanics (see, e.g., Tamm, 1926, and Landau, 1926), but also arrived at the above discussed formulae for the elements of the polarization matrix α and the quantity $\overline{\overline{\mu}}_z$ in the limit of high temperatures (see, e.g., Van Vleck, 1926a; Kronig, 1926d).[251] Thus, the rotator became the standard system in which the theoreticians learned to use and appreciate the tools of matrix mechanics.

[250] In the previous atomic theory, Pauli had obtained the following expression for the elements $\bar{\mu}_z(j,m)$ (see Pauli, 1921a, p. 324, Eq. (17)):

$$\bar{\mu}_z(j,m) = \frac{8\pi^2 A\mu^2}{h^2} \cdot \frac{1}{4j^2}\left(3\frac{m^2}{j^2} - 1\right).$$

The elements $\bar{\mu}_z(j,m)$—excluding the term with $j = 0$—had contributed to $\overline{\overline{\mu}}_z$, in the limit of high temperatures T, the amount $1 \cdot 5367kT \cdot |\mathbf{E}|$, rather than $\frac{1}{3}kT \cdot |\mathbf{E}|$ (Pauli, 1921a, p. 327, Eq. (26)).

In a review of recent data on the dielectric constant, Otto Blüh stated that they yielded only a lower limit for the constant C in Debye's law, which was far below the classical value of $\frac{1}{3}$ (see Blüh, 1926, p. 235); hence one could not rule out the older result.

[251] The problem of determining the polarizability of diatomic molecules was taken up, shortly after Mensing and Pauli, by Charles Manneback in Zurich (Manneback, 1926). He applied the methods of wave mechanics, arriving at the same results as Mensing and Pauli had done with matrix methods. Thus, he contributed to demonstrating the equivalence of the two formulations of quantum mechanics.

Chapter V
Modifications and Extensions of Matrix Mechanics

The fact that matrix methods were available for the task of formulating the ideas, which had emerged from the penultimate sharpening of the correspondence principle, must be considered to be fortuitous. It enabled Born and his collaborators to write down a scheme of the new quantum theory without having to invent new mathematical tools. However, the theory thus developed—matrix mechanics—did not satisfy the authors completely. Especially, when they investigated its physical applications they quickly realized that a wider, more general scheme was necessary in order to deal with problems in which nonlibrating variables played a crucial role. Fortunately for the further development of quantum mechanics such modifications were soon discovered, still before the end of the year 1925. Interestingly enough, the ideas leading to the generalization of matrix theory were brought in by outsiders: they were proposed by Cornelius Lanczos in Frankfurt and by Norbert Wiener at the Massachusetts Institute of Technology. These outsiders knew certain mathematical methods which they considered to be superior to the matrix methods for handling the problems of atomic theory; their methods—the field-like formulation of quantum mechanics (Lanczos) and operator mechanics (Wiener)—would have led to far-reaching consequences if they had been properly pursued. However, these authors themselves did not possess the necessary knowledge of atomic physics in order to arrive at those consequences. The experts like Born, Pauli or Heisenberg, on the other hand, made use only of a small part of the possibilities offered by the different mathematical schemes. Born was primarily interested in treating the problem of uniform motion in a straight line, while Heisenberg and Pauli's minds were totally occupied by the efforts to introduce action and angle variables into matrix mechanics for the purpose of calculating such specific problems as the relativistic fine structure and the anomalous Zeeman effects.

In proceeding with their programmes, Heisenberg and Pauli discovered another important point. The existing quantum mechanics, even if generalized to allow for the representation of nonlibration coordinates, did not suffice to describe the atomic situation. The scheme had to be further extended to include a new concept—the spin of the electron—which had again been proposed by two outsiders: George Uhlenbeck and Samuel Goudsmit in Leyden. But the incorporation of electron spin into the theory did not occur smoothly: in the beginning it seemed to create more difficulties than it removed. However, after a few months, following several careful analyses, the difficulties with electron spin disappeared and a satisfactory explanation of all atomic problems—with the exception of the

helium problem—could be reached. It took a further thorough examination of certain consequences of the quantum-mechanical scheme, especially a deeper look at the degenerate case leading to the recognition of an exchange phenomenon, until it became clear that the energy states of helium would also eventually come out correctly.

V.1 Nonmechanical Stress versus Spin

The brilliant and successful application of matrix mechanics to the problem of calculating the energy states of the hydrogen atom raised great hopes that it might soon become possible to solve the outstanding problems of atomic structure. Two sets of atomic phenomena, which had become intimately connected with each other, required an explanation urgently: first, the anomalous Zeeman effects; second, the multiplet structure, in general, of spectral lines and, in particular, the doublet structure of hydrogen and alkali lines. Since Heisenberg and Pauli were intimately familiar with the problems which had arisen in dealing with these phenomena during the previous years and had now mastered the tools of matrix mechanics, they became involved in investigating these tricky aspects of atomic structure according to the new scheme.

The development began already before Pauli attacked the hydrogen problem. In a letter to Ralph Kronig, dated 9 October 1925, Pauli discussed the so-called intercombination lines, that is, the spectral lines arising from the combination of, say, the singlet and triplet terms in alkaline earth atoms, and he suggested that they might be explained on the basis of a heuristic ('*Ersatz*') mechanical model and the application of 'sharpened' correspondence arguments.[252] '*Ersatz* models' of atoms had been constructed, especially by Heisenberg and Pauli since fall 1924, in order to be able to deal with the 'classically non-describable duplicity' ('*klassisch nicht beschreibare Art von Zweideutigkeit,*' Pauli 1925a, p. 385) of the states of the series electron in many-electron atoms. In spite of the successes achieved in this way in explaining the available wealth of spectroscopic data, Pauli noted:

> The substitute models really suffer from all the deficiencies [against] which I have always preached in the past years: 1. The "stress" ("*Zwang*") is not interpreted physically, but just described formally by reformulating the relativistic energy differences as such, which depend on the [relative] orientation of the [angular momentum] vectors and are characterized by the notorious double magnetism. 2. Hence one has here always two kinds of modification of the classical theory, which appear to be logically independent of each other. One may now characterize these as follows: (a) transition from the classical mechanics to Heisenberg's mechanics; (b) replacement of the classical Hamiltonian $H(p, q)$ by the one describing the substitute model. (Pauli to Kronig, 9 October 1925)

[252] The intercombination lines belonged to those phenomena which were the most difficult to explain on the basis of atomic models. In his *Handbuch* article, Pauli had discussed them twice (see Pauli, 1926b, Sections 36 and 44).

Thus, even though Pauli accepted the new quantum mechanics, which involved the reformulation of the classical Hamiltonian in terms of the corresponding quantum-theoretical quantities of matrices, he believed that the scheme developed so far could neither describe the anomalous Zeeman effects nor even the fact that the helium atom possessed singlet *and* triplet states. Moreover, he claimed, one had to incorporate certain essential features which had hitherto been expressed by the '*Ersatz*' models.

Pauli had asked Kronig to show his letter to Heisenberg, who was also in Copenhagen at that time. Heisenberg, however, did not agree with the views Pauli had expressed, for he wrote in his next letter: 'Further, I really do not believe any word anymore of your theory of stress ("*Zwang*"). Although it [the stress] also certainly exists in the case of hydrogen, nevertheless I don't believe that it has the least bit to do with the nucleus; perhaps it is already present within the mould ("*Schimmel*")' (Heisenberg to Pauli, 12 October 1925). Heisenberg tended to believe that the replacement of the mechanical models of atoms—which he and Pauli had designated as '*Schimmel*' for a while—by the matrix representation might already be sufficient to account for the observed structure of spectral lines. He especially expected that some special feature of the quantum-mechanical scheme might prove to be helpful. Some indication was already provided by the fact that the Hamiltonian matrices describing the dynamical behaviour of atoms had both discrete *and* continuous energy eigenvalues. From this he concluded: 'According to quantum mechanics, Larmor's theorem, for instance, is not valid in general. For hydrogen, for example, there exists—even for the orbits in the ground state—a relatively large probability to pass over by absorption [of radiation] into free electrons, and for free electrons Larmor's theorem certainly does not apply [because of the gyromagnetic ratio]' (Heisenberg to Pauli, 12 October 1925). Heisenberg was aware of the fact that Larmor's theorem was valid in the matrix treatment of the harmonic and anharmonic oscillators, but in these cases only a discrete spectrum existed. For realistic atomic systems, however, there was a good chance that the theorem would break down and, as a result, the complex structure of spectra and the anomalous Zeeman effects might perhaps follow.

Within a few weeks of this exchange with Kronig and Heisenberg, Pauli arrived at two important conclusions, which emerged from his investigation of the hydrogen atom: first, by quantizing the total angular momentum one would not, at least in the case of the one-electron atom, obtain half-integral quantum numbers for the electron; second, it became inconceivable that Larmor's theorem could be violated at all in quantum mechanics. Heisenberg immediately conceded to Pauli: 'With your objection against introducing hyperbolic orbits for explaining the Zeeman effect, I surely agree. It also occurs to me fairly improbable that these hyperbolae [which describe the orbits of free electrons] would be sufficient to explain all the ugly features of the coupling [which gives rise to the anomalous Zeeman effects]' (Heisenberg to Pauli, 3 November 1925). However, Heisenberg was not yet convinced that his earlier speculation could be ruled out entirely, and he wrote again to Pauli: 'I have also thought a bit about the "stress"

("*Zwang*"), and I still have the hope again and again that it might come in only in the case of the coupling of two (or several) electrons (and, indeed, through your exclusion of the equivalent orbits)' (Heisenberg to Pauli, 16 November 1925). Pauli did not share Heisenberg's opinion and, in writing to Bohr, stated quite emphatically:

> I could join in the opinion that atoms with one electron are fundamentally different from atoms with several electrons only if one were able to establish a theory on a unified basis, from which would follow, on one hand, no stress ("*Zwang*") for the hydrogen atom and, on the other, the empirically-obtained anomalous Zeeman effect correctly for the alkali atoms. In the present form of Heisenberg's theory, however, this is not the case; it always leads to Larmor's theorem, and I believe I can state with certainty that it [Heisenberg's theory] will lead in the case of helium to one rather than two term systems. (Pauli to Bohr, 17 November 1925)

Pauli believed firmly that the occurrence of two term systems of the helium atom (i.e., ortho-helium and parhelium) was closely connected with the fact that already the ground state of the hydrogen atom (i.e., of a one-electron atom) in an external field was twofold. 'In particular, I believe,' he remarked to Bohr, 'that one would, if one performed the experiment of deflecting hydrogen atomic beams in inhomogeneous magnetic fields . . . , see the same thing as in the case of the Ag-atom, namely ± 1 magneton and *not* non-magnetic atoms' (Pauli to Bohr, 17 November 1925). In the matrix treatment of the hydrogen atom just performed, the ground state possessed the angular momentum zero (i.e., no angular momentum whatever), and there was no experiment available which contradicted this conclusion. Later, Erwin Wrede, a student of Otto Stern's, began to work on a Stern–Gerlach experiment with atomic hydrogen. On 17 July 1926 he presented the results of his measurements in a talk entitled 'Atomic Hydrogen Beams and their Magnetic Splitting' ('*Wasserstoffatomstrahlen und deren magnetische Aufspaltung*') at the eleventh meeting of the *Gauverein Niedersachsen* (Lower Saxony Section) of the German Physical Society in Göttingen. He had found that under the action of an inhomogeneous magnetic field a beam of hydrogen atoms split up into two distinct parts, consistent with the fact that each atom possessed a magnetic moment of one Bohr magneton.[253] Then, however, this result, which confirmed Pauli's earlier speculation, did not imply any difficulty concerning the theoretical understanding, for the spin of the electron had meanwhile been postulated and accepted as a fundamental property.

But in November 1925 the conceptual situation in this regard appeared to the pioneers of quantum mechanics to be quite unpleasant. As Heisenberg recalled: 'We had this doubling [i.e., "two-valuedness"] and we really did not know what it meant. That was disagreeable. We were a bit unhappy, but then so many things

[253] E. Wrede's talk was listed among the talks given at the meeting (see *Verh. d. Deutsch. Phys. Ges.* (3) **7**, p. 37). Wrede had reported his conclusions already earlier to Bohr (Wrede to Bohr, 6 July 1926). He obtained his doctorate with a thesis including the results of his hydrogen measurements (Wrede, 1927).

were still not understood that we thought, "Let's wait and see what will come later." Then the paper of Uhlenbeck and Goudsmit came' (Heisenberg, Conversations, p. 377).

On 17 October 1925, exactly a month before Pauli wrote to Bohr about his suspicion that the ground state of hydrogen ought to be twofold, George Uhlenbeck and Samuel Goudsmit had signed their note for *Naturwissenschaften* in which they proposed to replace the hypothesis of the 'non-mechanical stress' ('*unmechanischer Zwang*') of an electron (Uhlenbeck and Goudsmit, 1925). On the basis of a reanalysis of the existing models for explaining the spectroscopic data on multiplet structure and the anomalous Zeeman effects, they came to the conclusion that the entire empirical data—including relativistic and X-ray doublets—could be accounted for in a simple and satisfactory way by ascribing a 'proper (i.e., intrinsic) rotation' to the electron such that the angular momentum of this rotation assumed the value $\frac{1}{2}(h/2\pi)$. Heisenberg read Uhlenbeck and Goudsmit's note immediately on its appearance in *Naturwissenschaften*—it came out in its issue No. 47 on 20 November—and wrote to Pauli about it on a postcard: 'Since I have bothered you anyway with so many questions, a further one certainly does not matter; hence I would like to hear from you, as the master of criticism, what you think about [Uhlenbeck and] Goudsmit's note in *Naturwissenschaften*.' After mentioning a few difficulties arising from the intrinsic rotation hypothesis in atomic physics, he concluded: 'It is funny that in the model of [Uhlenbeck and] Goudsmit [they] can indeed get the Z^4-law with the magnetic hypothesis; hence I am not *a priori* against [Uhlenbeck and] Goudsmit's theory. However, I would like to hear your objections' (Heisenberg to Pauli, 21 November 1925).

With the 'magnetic' origin of the Z^4-law Heisenberg had in mind the idea proposed by Alfred Landé one and a half years previously. In discussing the origin of the so-called relativistic X-ray doublets, Landé had considered a magnetic coupling between the radiating electron and the atomic core of a many-electron atom. He assumed that there existed two different states of the atom, in which the core possessed the magnetic quantum numbers m and $m + \Delta m$, respectively, in the direction of the magnetic field produced by the orbiting electron at the centre of the atom; then the energy difference of the two spectral terms associated with the quantum numbers m and $m + \Delta m$ was given by $h\Delta\nu_{\text{mag}} = (\text{const.})(Z^3\Delta m/n^3k^2)$, Z being the atomic number and n the principal quantum number of the electron orbit. (See Landé, 1924a, p. 95; the constant factor has the value $8\pi^4 m_e e^8 h^{-4} c^{-2}$, where m_e and $-e$ are the mass and charge of the electron, h is Planck's constant and c is the velocity of light *in vacuo*.) Heisenberg now imitated Landé's procedure; i.e., he considered a one-electron atom, having a nucleus with charge Ze, and he attributed to the electron, in accordance with Uhlenbeck and Goudsmit's hypothesis, an 'intrinsic' angular momentum of $\frac{1}{2}(h/2\pi)$ and a magnetic moment of one Bohr magneton. Then, for terms with principal quantum number n (and orbital angular momentum quantum number k of the electron), he found for the energy difference according

to the Bohr–Sommerfeld theory the expression

$$h\Delta\nu = \frac{8\pi^2 m_e e^8}{h^4 c^3} \cdot \frac{Z^4}{n^3 k^2} \cdot 2, \tag{211}$$

whose value, in the case of hydrogen and the helium ion, was twice as large as the one calculated from Sommerfeld's relativistic theory (Sommerfeld, 1915c). 'That k^2 occurs in the denominator instead of $k(k-1)$,' Heisenberg noted, 'would not bother anybody, but I cannot get rid of the factor 2. If one could get rid of this factor 2, then the model [of Uhlenbeck and Goudsmit] would agree indeed with your model 1925–10 [—this is how Heisenberg teasingly referred to Pauli's *Ersatz* model, about which Pauli had reported to Kronig in the above-mentioned letter (Pauli to Kronig, 9 October 1925); there was obviously no dearth of such models—] and all difficulties would be removed with a magic wand ("*Zauberschlag*")' (Heisenberg to Pauli, 21 November 1925). In his first reaction to the spin hypothesis, Heisenberg was quite positive. He thought that a quantum-mechanical treatment of an atomic system with intrinsic rotation might —besides yielding some 'averaging' over the angular momentum quantum number k—retain the essential features brought in by the new concept, especially the explanation of the multiplet structure of spectral terms. However, he noticed that besides the difficulty concerning the factor 2, there was another problem. 'In addition one must however investigate,' he remarked to Pauli, 'where the really existing relativity disappears in the case of hydrogen' (Heisenberg to Pauli, 21 November 1925). Three days later, in his next letter to Pauli, Heisenberg was much more reserved; he had had time to reflect about certain implications of Uhlenbeck and Goudsmit's proposal and had found further difficulties which had to be removed before intrinsic rotation or spin—as it was to be called later—could be accepted in atomic theory. Thus he wrote:

> Nevertheless, I myself do have very great reservations against the attempt of [Uhlenbeck and] Goudsmit to establish zoology [i.e., the term structure of many-electron atoms] in such a simple way: 1. Also the nucleus must [then] have angular momentum, and then the simplicity of the zoology could hardly be understood. 2. It would of course be simpler if the electron possessed only charge and mass, but no angular momentum; in principle, however, one cannot argue against the angular momentum, but the idea that the electron has a structure (in particular, that several types of electrons exist) appears to me to be detestable. 3. In the case of the alkaline-earth atoms the energy difference between singlet and triplet atoms having the *same* l [in the erstwhile atomic theory the quantum number l had been defined as $k-1$, where k was the quantum number of the orbital angular momentum (see, e.g., Pauli, 1926b, p. 219, Eq. (38)); l assumed the values 0, 1, 2, etc., for the s-terms, p-terms, d-terms, etc.] should then arise only through different interactions of the magnets [i.e., magnetic moments] of the two electrons with each other. This yields a totally wrong order of magnitude. (Heisenberg to Pauli, 24 November 1925)

These three objections were certainly serious. Two of them, (1) and (3), arose from the analysis of the spectroscopic data, while (2) expressed the difficulty of

understanding why the electron should have an internal structure. At a time when one believed that the atomic nucleus consisted of protons and electrons, it seemed evident that the nucleus must possess an intrinsic angular momentum due to the electrons.[254] With his remark concerning the difference of singlet and triplet terms having the same quantum number l, Heisenberg drew attention to the following fact: if one explained it, as appeared most obvious, as being due to the magnetic interaction of the electrons—i.e., as arising from the interaction of the magnetic moments of two or more electrons—then the term differences would come out several orders of magnitude less than the observed ones. In spite of these objections, however, Heisenberg felt that one should at least consider the electron spin hypothesis as providing a reasonable possibility of arriving at a *unified* explanation of the phenomena which had been attributed thus far to the existence of a 'non-mechanical stress.' 'Nevertheless, Goudsmit has stimulated me so far towards zoology,' he concluded the letter containing the above objections, 'that I want to put his model once through the matrix-mill and see whether at least the g-formulae, intensities, etc., come out correctly' (Heisenberg to Pauli, 24 November 1925).

Pauli received the proposal of Goudsmit and Uhlenbeck with much greater reservation than had Heisenberg. After all, he had already turned down a similar idea of Ralph Kronig's in early 1925, although it provided a mechanical model for the extra quantum number involved in the exclusion principle. Now, in November 1925, he began systematically to marshal arguments against the spin hypothesis. He did not regard the factor 2 in Eq. (211) as the most serious difficulty that had arisen; he had a deeper conceptual reason for turning down the idea of electron spin.[255] Especially, he 'strongly doubted the correctness of the idea because of its classical mechanical character' (Pauli, 1946, p. 214). He rather believed, as he had pointed out in his letters to Kronig (9 October) and Bohr (17 October), in the necessity of extending the quantum-mechanical scheme of Born, Heisenberg and Jordan in such a way that the features attributed before to 'non-mechanical stress' ('*Zwang*') would be incorporated. In Pauli's opinion, this basically quantum-theoretical extension could not simply be replaced by endowing the electron with an extra mechanical property, the 'intrinsic angular momentum' of Uhlenbeck and Goudsmit. He completely agreed with Bohr when the latter remarked that 'the definition of stress ("*Zwang*") really rests exactly on

[254]The protons would contribute a much smaller intrinsic momentum because of their larger mass. In fact, Pauli, in a short note submitted August 1924 (Pauli, 1924c), had already discussed the origin of the satellites accompanying the components of line multiplets that arose from the interaction of the angular momentum of the nucleus with the angular momentum of the series electrons. This interaction should give rise to a very small splitting of the multiplet components, which was called 'hyperfine structure'; hence it followed that, if the nuclear angular momentum was responsible for hyperfine structure, then it had to be several orders of magnitude (exactly three orders, i.e., the ratio of electron mass over nuclear mass) smaller than the angular momenta connected with electrons.

[255]In fact, Pauli at first did not notice the difficulty concerning the factor 2 and requested Heisenberg to give him the steps leading to Eq. (211). Heisenberg, in his letter to Pauli dated 24 November 1925, gave the derivation and mentioned: 'The above calculations have been given already by Landé (*Z. f. Phys.* **24**, 88, 1924); there, however, Z^3 always replaces Z^4.'

the proof that mechanical pictures are inadequate for atoms having several electrons' (Bohr to Pauli, 25 November 1925). With great relief, Pauli replied to Bohr: 'I am very happy that we are essentially in agreement concerning our judgment about the present situation in atomic theory. Also Heisenberg, with whom I recently had an active exchange of letters, has similar views. Hopefully we shall soon make progress' (Pauli to Bohr, 1 December 1925).

The agreement between Bohr, Heisenberg and Pauli in their skeptical attitude towards the spin hypothesis did not last long. On 9 December 1925 Niels Bohr travelled to Leyden in order to attend the celebrations connected with the fiftieth anniversary of Hendrik Antoon Lorentz' doctorate. In Leyden he not only made the acquaintance of Uhlenbeck and Goudsmit and their ideas, but also had occasion to discuss the spin hypothesis with Albert Einstein. A few months later, Bohr recalled the situation in a letter to Ralph Kronig, who had returned from Copenhagen to the United States before the Lorentz jubilee. He wrote:

> Since you left Copenhagen we have often missed you. Indeed, I have been deeply interested in just these spectral problems to which you have given so much thought . . . . From the discussion after your beautiful lecture on the Zeeman effect you know how unhappy and desperate I felt over the state of things. Certainly I had heard then that in a letter to *Naturwissenschaften* Uhlenbeck and Goudsmit had proposed to provide a basis for a new quantum number by ascribing a vector to the electron itself; but not understanding how this vector could materially influence the motion of the electron to which it was attached, I had thought so little about this idea, that I even quite forgot to question you about this possibility [during] the discussion that evening. When I came to Leyden, however, to the Lorentz festival, Einstein asked the very first moment I saw him what I believed about the spinning electron. Upon my question about the cause of the necessary mutual coupling between the spin axis and the orbital motion, he explained that this coupling was an immediate consequence of the theory of relativity. This remark acted as a complete relivation [*sic*, revelation] to me, and I have never since faltered in my conviction that we at last were at the end of our sorrows. (Bohr to Kronig, 26 March 1926)

After leaving Leyden, Bohr visited Göttingen for several days and spoke, especially in conversations with Heisenberg, in favour of the electron spin hypothesis. Then, on his way back to Copenhagen, he passed through Berlin where he met Pauli. 'We spent a beautiful day in Berlin,' Bohr reported to James Franck, his host in Göttingen. 'I have delivered your message to Planck, and I have succeeded with Pauli at least to the extent that he is himself extremely interested to find out whether or not the matrix reformulation of the electron-magnet gospel correctly represents experience' (Bohr to Franck, 22 December 1925). At that time Heisenberg, assisted by Jordan, was treating the anomalous Zeeman effect in matrix mechanics by including the electron spin hypothesis. On 24 December 1925 he wrote to Pauli about the initial results, and remarked:

> As agreed, we have attacked the quantum mechanical calculations of this theory; however, "due to lack of time and spirit" we have not proceeded beyond the quasi-trivial results. The g- and γ-formulae are really entirely trivial, nearly so is

also the sum-rule and it comes out as was to be expected; for the doublets there results Voigt's formula, and for triplets a cubic equation (which I believe agrees with the cubic equation contained in my first paper [Heisenberg, 1922a]?!), etc. [Then he continued:] What is still missing in order to calculate the absolute values of the splitting, are "only" the two average values $\overline{r^{-2}}$ and $\overline{r^{-3}}$ [i.e., their matrix reformulations]; the first one for [calculating] the relativity, and the second one for the magnetism [i.e., for the matrix calculation of the relativistic splitting and of the magnetic splitting of the terms caused by the interaction of electron spins, which would yield the quantum mechanical formula replacing the classical Eq. (211)]. As soon as one has these average values, one can decide whether or not [Uhlenbeck and] Goudsmit's theory is a possible description, but I do not yet know how to compute these average values. (Heisenberg to Pauli, 24 December 1925)

While Heisenberg expressed some optimism that one might eventually obtain the right answer concerning the relativistic fine structure by calculating the inverse powers of the radial variable—especially a correction of the wrong factor 2 in Eq. (211)—Pauli, towards the end of 1925, again fell back into his previous negative attitude. 'After thinking about it again, I am still skeptical, and I believe that Goudsmit is on the wrong track,' he wrote to Bohr on 30 December 1925. He then pointed out, first, the difficulty of understanding, at least in the limit of high quantum numbers k, the existence of screening doublets. 'And secondly,' Pauli continued, 'when one looks at it more closely, it becomes difficult to understand why an electron with angular momentum $\frac{1}{2}(h/2\pi)$ does not possess a much larger energy (mass) than the one that is actually determined.' That is, Pauli expected that the existence of an intrinsic rotation would contribute considerably to the self-mass of the electron. 'Though I admit,' he went on, 'that one will be able to come to a final conclusion only on the basis of detailed quantum mechanical calculations, and that one cannot see at the moment any other way of understanding the failure of Larmor's theorem and the duplicity, yet I can so far only shake my head about Goudsmit's idea and say: "I don't like the thing!"' (Pauli to Bohr, 30 December 1925).

In spite of their slightly different attitudes, Pauli and Heisenberg agreed on one point: the next task consisted in calculating in matrix theory the averages of the inverse powers of the radial matrix quantities, $\overline{\mathsf{r}^{-2}}$ and $\overline{\mathsf{r}^{-3}}$, over full periods. On returning to Göttingen after his Christmas vacation, Heisenberg reported to Pauli: 'I have used the railway travel from Munich to think a little about the famous average values. It is true that I have not arrived at any definite result, but perhaps you may be able to do something with the following idea' (Heisenberg to Pauli, 7 January 1926). Then he sketched a description of the electron's motion in the hydrogen atom in terms of polar coordinates. In particular, he proposed to introduce a quantum-theoretical angle ϕ through the reformulated classical relation

$$\phi = \arctan\left(\frac{\mathsf{y}}{\mathsf{x}}\right), \tag{212}$$

where x and y were the Cartesian coordinate matrices. He added: 'That this

quantity cannot be expanded in a [quantum-theoretical] Fourier series, is of no significance, according to a recent paper of Born and Wiener' (Heisenberg to Pauli, 7 January 1926).

A week later Heisenberg informed Pauli that he might pass through Hamburg on 20 January in order to discuss the problem of evaluating the averages of the radial matrix quantities, $\overline{r^{-2}}$ and $\overline{r^{-3}}$, with the help of an action-angle formalism. On the same postcard he wrote: 'Herewith I am sending you the paper of Born and Wiener and a similar paper of Lanczos for [our] discussion on Wednesday' (Heisenberg to Pauli, 14 January 1926). The papers of Born and Wiener (1926b) and of Lanczos (1926a), to which Heisenberg referred, involved certain modifications of the matrix scheme; with these modifications one could try to establish a consistent quantum-mechanical scheme of action *and* angle variables, which seemed to be necessary in order to solve problems of relativistic fine structure and of the anomalous Zeeman effects.[256]

V.2 Field-Like Representation of Quantum Mechanics

The new theory of quantum mechanics of Heisenberg, Born and Jordan, is related most closely with the theory of integral equations. The equations of motion and the quantum conditions can be written down in the form of integral equations. Thus there results a continuum interpretation, which must be placed side by side with the discontinuum interpretation [of Heisenberg, Born and Jordan] as having equal rights, since there exists a one-to-one relation between the two. For the fundamental interpretation of the theory, the integral formulation has the advantage that it is in immediate harmony with the usual field-like conceptions of physics. (Lanczos, 1926a, p. 812)

These programmatic lines constituted the abstract of the first paper on quantum mechanics that was written by a physicist outside the Göttingen and Copenhagen circles. It was entitled '*Über eine feldmäßige Darstellung der neuen Quantenmechanik*' ('On a Field-Like Representation of the New Quantum Mechanics') and its author was Kornel (or Cornelius, as he called himself later on) Lanczos, then an assistant at the Institute for Theoretical Physics of the University of Frankfurt-am-Main. Who was Cornelius Lanczos and how did he come to work on the new theory?

Lanczos was born on 2 February 1893 in Székesfehérvár, Hungary.[257] From

[256] Heisenberg had received from Born a copy of the paper of Born and Wiener and read it during the Christmas vacation; it was recieved by *Zeitschrift für Physik* on 5 January 1926 and appeared in its issue No. 3 of 12 March 1926. On his return to Göttingen after the Christmas vacation he found a copy of Cornelius Lanczos' paper on a field-like representation of the new quantum mechanics (which, evidently, Lanczos had sent to Göttingen); this paper was received by *Zeitschrift für Physik* on 22 December 1925 and later appeared in its issue No. 11/12 of 26 February 1926.

[257] Lanczos used the Hungarian spelling of his first name, Kornel (without the accent on e), in most of his German publications, changing in 1929 to Cornel and later to Cornelius. It is of interest to note that his family name, Lanczos, appeared in an early publication of 1923 as Lancius. Finally, it should be mentioned that his father's name was Carolus Loewy—'Lánczos' was the Hungarian translation of 'Loewy.'

1911 to 1916 he studied physics and mathematics at the University of Budapest, where his teachers were Loránd Eötvös and Lipot (Leopold) Fejér; Eötvös was known for his experimental researches connected with establishing the equivalence of gravitational and inertial masses and Fejér for his contributions to the theory of Fourier series. Upon graduation Lanczos became an assistant at the Polytechnical University of Budapest where, for the next four years, he worked on experimental physics. In 1921 he finally received his doctorate from the University of Szeged, which was also situated in Budapest at the time, with a thesis on 'The Function-Theoretic Behaviour of Maxwell's Vacuum Equations,' supervised by Rudolph Ortvay, a former student of Sommerfeld's. Adverse political circumstances prevailing in Hungary made it difficult for Lanczos to pursue his plans there upon completing his doctorate. Like many another young man at the time with an ambition to embark upon a career in scientific research, Lanczos went to Germany. He found a position at the Physics Institute of Freiburg-im-Breisgau and stayed there for three years.[258] In fall 1924 he went to the University of Frankfurt, where the prospects for his scientific career seemed to be more promising.

In spite of the fact that he worked in places like Freiburg and Frankfurt, which did not count themselves among the principal centres of theoretical physics in Germany, Lanczos soon managed to achieve a good name for himself by his work on general relativity theory.[259] During a visit to Berlin he had come under the spell of Albert Einstein, the founder of this theory. 'I had my first discussion with Einstein when I was in Germany in 1921,' Lanczos recalled forty-five years later. 'I had a chance to talk to him for a few minutes between two lectures. Of course, I had already had a tremendous admiration for him in my student days, and so I was in a great palpitation of heart when this opportunity came' (Lanczos in Whitrow, 1972, p. 48). Lanczos had at once admired the relativity theory because of its conceptual power and, as he expressed it in 1923 during a discussion, because of 'its logical consistency [*Folgerichtigkeit*], in that it takes the foundations of differential geometry—which

[258] In the list of members of the German Physical Society (which Lanczos had joined in spring 1920), published at the end of the year 1921 (see *Verh. d. Deutsch. Phys. Ges.* (3) **2**, p. 87), Lanczos' home address was given as Berlin. Hence he spent some time in late 1921 in Berlin before going to Freiburg. In Freiburg the leading physicist was Franz Himstedt who, in 1924, was succeeded by Gustav Mie. Another member of the faculty was Emil Cohn, formerly professor at the University of Strasbourg (he was known for his research on the electrodynamics of moving bodies) and since 1920 '*ordentlicher Honorarprofessor*' ('honorary full professor') of theoretical physics; he had a deep interest in the physics of space and time. Among the mathematicians in Freiburg was Alfred Loewy.

[259] In none of the papers which Lanczos submitted from Freiburg, and in hardly any of the ones submitted later from Frankfurt, did Lanczos thank one of the professors or the senior people at these universities. Hence one may conclude that he worked on his papers entirely by himself, guided only by the ideas of Einstein and a few others. Gustav Mie, with whom some of his interests may have overlapped, came to Freiburg only in fall 1924 when Lanczos had already gone to Frankfurt.

On the other hand, Lanczos already had heard about Einstein's work on relativity theory in Hungary since he studied with Eötvös, whose experimental results had provided one of the fundamental principles (i.e., the equivalence of inertial and gravitational masses) that was a cornerstone of general relativity theory, and Ortvay, who had introduced the recent ideas of modern physics to his students.

are free of any arbitrariness—as the general foundation of the description of nature, so that the imperative of thought becomes the necessity of nature' (Lanczos in Raschevsky, 1923, p. 441). Thus, Lanczos, the former disciple of Eötvös, became a devoted researcher in general relativity theory during the following years, treating such topics as the equations of gravitation, the properties of de Sitter space, the problem of rotation, the theory of the perihelion shift of Mercury, and cosmology.[260]

It was while investigating the behaviour of weak fields in Einstein's theory of gravitation that Lanczos discovered an interesting fact (Lanczos, 1925a). He had to solve a differential equation for a tensor field with certain boundary conditions imposed on the solutions; he converted the differential equation into an integral equation and discovered that the latter was of the Fredholm type. This discovery pleased him very much, especially since he had arrived at the integral equation not by mathematical but by physical argumentation. Indeed, Lanczos did not know then that Hilbert had established the equivalence of differential equations (with boundary conditions) and integral equations nearly twenty years before.[261] It was perhaps a charming accident that Lanczos, who was well-trained mathematically, had not come across this result earlier; yet his rediscovery of the old methods of integral equations had a certain advantage which showed itself very soon in his work.

In fall 1924 Lanczos went to Frankfurt where Erwin Madelung, the Director of the Institute of Physics, was just preparing the second edition of his book, *Die mathematischen Hilfsmittel des Physikers*, a compendium of the mathematical tools of physics (Madelung, 1925). Madelung added a number of new topics, such as Green's functions, tensor analysis and general relativity theory in the new edition, and Lanczos helped him with their presentation.[262] In the course of this

[260] In the 1920s, when debates on relativity theory often took place in the discussions at physics meetings, Lanczos took a strong stand in favour of Einstein's theory. (By the way, Lanczos served as a secretary—*Schriftführer*—of the German Physical Society from 1922 to 1924.) At the annual meeting of the Society in Bonn in September 1923, Lanczos expressed his opinion in support of relativity theory. Although the atmosphere then was less controversial than at the Bad Nauheim meeting three years earlier, some people still raised doubts about the validity of general relativity theory. New results, such as the measurement of the deflection of light in the gravitational field of the sun made by the *American Einstein Expedition* of 1922, seemed to be compatible with a different theoretical description than the one provided by Einstein's theory (Hopmann, 1923). Lanczos, however, argued that any numerical agreement obtained from a nonrelativistic theory had to be considered an extreme coincidence ('*kolossaler Zufall*,' Lanczos in Hopmann, 1923, p. 484).

[261] Lanczos discovered this only after he had submitted the above-mentioned paper and added the following remarks in proof:

> In writing this paper I have unfortunately not taken enough notice of the existing mathematical literature. Therefore, I wish to point out here that already Hilbert, in his fundamental memoirs on integral equations (which appeared together in one volume, Teubner, 1912), applied "Green's tensor" to solve simultaneous differential equations. (See pp. 206–212 [of Hilbert, 1912c]; see also the book of Courant–Hilbert, *Methoden der mathematischen Physik I*, Springer, 1924, p. 316, which provides an excellent review of the theory of integral equations.) He deals naturally only with Euclidean domains. The presentation given above should be closer to the line of thought of the physicist interested in theory. (Lanczos, 1925a, p. 124, footnote 1)

[262] Madelung, in the preface to the new edition, acknowledged the help of his assistant. One finds that in the section on relativity some of Lanczos' recent results on the theory of gravitation were included.

work for his new boss Lanczos had to study the mathematical literature, in particular the subjects treated in the book of Courant and Hilbert (1924), which had just been published and to which Madelung referred frequently. A short section on quantum theory was also included in the *Mathematische Hilfsmittel*, and it was more than a coincidence that, at about the same time as he assisted Madelung on the book, Lanczos began to address himself to an outstanding problem of quantum theory.

The current status of quantum theory had been reviewed recently, in September 1924, at the *Naturforscherversammlung* in Innsbruck, where Hendrik Kramers and Arnold Sommerfeld presented differing views concerning the structure of atoms. The most controversial question, as revealed by the discussions in which Einstein participated, was the status of the radiation problem. Two theories of radiation competed with each other at that time, each describing a certain set of phenomena adequately: one was the theory of Bohr, Kramers and Slater, which had just been successful in accounting for the dispersion of light by atoms; the other was Einstein's old theory of light-quanta, which had recently been supplemented by the statistical method of the Indian physicist Satyendra Nath Bose. However, the two theories seemed to be completely at variance, and Lanczos became interested in the question whether general relativity theory might help in resolving the discrepancies. He hoped, especially, that perhaps some of the quantum features of radiation would follow by considering radiation in a closed universe.

Lanczos concerned himself with this question in the first half of 1925, making ample use of the methods of integral equations. The latter implied that the electromagnetic four-potential Φ_i ($i = 0, 1, 2, 3$; Φ_0 denoting the electric potential, and Φ_1, Φ_2 and Φ_3 forming the magnetic vector potential) satisfied an integral equation of the form

$$\Phi_i = \int \sum_{i=0}^{3} K_{ij}(x, x')\Phi_j(x')\,dx', \qquad i = 0, 1, 2, 3, \tag{213}$$

where $K_{ij}(x, x')$ is a kernel which depends on two points, x and x', of the four-dimensional space, and the path of integration extends over the entire universe. As a special case, Lanczos treated the situation in a spherical universe of finite, constant radius containing an atom at its origin (Lanczos, 1925b). The eigenfunctions solving Eqs. (213) in this example could be expressed as products of periodic functions in the spherical (spatial) coordinates (i.e., the distance from the origin and two angles). Every eigenfunction, $\Phi_i^{(n)}$, was then characterized by a number ν_n, which Lanczos interpreted as the frequency of the 'standing ether waves' (Lanczos, 1925c, p. 129, p. 137). The 'ether waves' could be excited by an electron of the atom, which performs a periodic motion with frequency ω. If integral multiples of ω coincide with the frequency of one of the ether oscillations, then the amplitude of the latter would become infinite according to the usual dynamical considerations (because the denominator, $\nu_n^2 - k^2\omega^2$, is zero),

unless one assumed zero coupling between the electron's motion and the ether-vibration. Lanczos now made this assumption and claimed that it provided an explanation for the stationary orbits of electrons in atoms as postulated by Bohr's theory in the following sense:

> Of all orbits, which are dynamically possible under the action of the Lorentz force, there exist only certain definite ones in which charge and current density possess such a distribution that they are in accord with the condition of stationarity [of the orbits]. Only on such orbits are permanent revolutions [of the electrons] possible, for then no transfer of energy to the ether occurs. In all other situations the electron emits radiation, probably not in the classical sense by the overflow of energy into space, but by the process that the amplitude of the forced ether vibrations increases steadily due to resonance [with the motion of the electron]. Because of the reaction of this radiation [i.e., of the ether vibrations] on the electron a new force, which is not of a conservative nature, arises in addition to the Lorentz force; and, as a consequence of this force, we have to understand [the fact] that the electron is always again driven back into the stationary orbits. (Lanczos, 1925c, pp. 140–141)[263]

Two types of radiation, had, therefore, to be distinguished in Lanczos' theory. The first, consisting of all the standing vibrations of the ether of constant amplitude, were described by the orthogonal and normalized eigenfunctions which were solutions of Eqs. (213). In the latter, the kernel $K_{ij}(x, x')$ took into account the motion of *all* charged particles (electrons and others) in the universe. Although Lanczos was aware of the fact that these 'eigenoscillations' of the ether had not been experimentally detected, he referred to the concept of 'virtual oscillator fields' in the theory of Bohr, Kramers and Slater (1924) as giving support to the existence of a permanent radiation of the ether. The second type of radiation was produced by the transitions of an electron from one stable orbit to another stable orbit. In this case, the energy of transition was regarded as being transferred completely to the ether, exciting in it a large number of eigenoscillations which interfered in such a way as to give rise to the monochromatic emission line of an atom.

With the help of these considerations, Lanczos reformulated the standard problem of describing the radiation emitted by atoms in a completely new manner. Instead of studying the mechanics of the atom one had now to investigate the excitation of the eigenoscillations in the ether (or the universe) by the motion of electrons in atoms. The fundamental difficulty of Lanczos' programme was how to construct the universal kernel function entering Eq. (213), and he therefore concluded: 'On the basis of this conception the solution of the quantum problem therefore appears to be reduced to the knowledge of a Hamiltonian principle, which for the time being still lies far in the dark, and its

[263] Lanczos proved himself again to be a follower of Albert Einstein's ideas about the origin of quantized electron orbits in quantum theory. Since many years Einstein had speculated that the stationary orbits may be obtained as specific solutions of some generalized field equations (see, in particular, Einstein, 1923d).

connection with the eigenfunctions' (Lanczos, 1925c, p. 141). He suggested that the action principle of general relativity could possibly give a hint for finding the appropriate action principle in radiation theory. Realizing how difficult this task might be, Lanczos hoped to be able to throw new light on further details by a 'closer investigation of the powerful ("*weitgreifende*") "method of eigenfunctions," even without an explicit use of a Hamiltonian principle' (Lanczos, 1925c, p. 141).

While he was pondering about the continuation of this programme, Lanczos read the papers of Heisenberg (1925c) and Born and Jordan (1925b). He immediately grasped that these authors had made important progress and, in the introduction of his next paper, he noted:

> The far-reaching thoughts of Heisenberg might be of path-breaking importance for research in quantum theory. In a logical continuation of the new ideas, Born and Jordan succeeded in giving Heisenberg's conceptions an adequate mathematical expression by a large-scale generalization; thus they discovered the general formal foundations of the new theory. One arrives at a logically established discontinuum theory, for which the classical concepts are of importance merely through correspondence and as heuristic guidelines; however, the new theory proceeds from the very beginning along its own directions and associates with the old symbols a completely new meaning. The fundamental principles of quantum theory have thus obtained an unexpected deepening. (Lanczos, 1926a, p. 812)

To Lanczos the new development in quantum mechanics came as a big surprise, for it broke completely with the physical conceptions he had grown up with, especially the field theory of matter. However, he quickly saw that, in spite of the apparent differences in formulation, the Heisenberg–Born–Jordan quantum mechanics made some connection with his earlier work on relativity and quantum theories. And the key to this connection was provided by the method of integral equations which had become his favourite tool. Since Lanczos had become familiar with Hilbert's work, he knew perfectly about the equivalence of matrix equations, such as the ones that appeared in the paper of Born and Jordan, and integral equations, which he had used earlier. To exploit this mathematical equivalence now, in late 1925, became his goal. Within a few weeks he was able to accomplish this task, and before Christmas 1925 he submitted a paper on it to *Zeitschrift für Physik*.[264] He summarized its content in the following words:

> In the following we wish to point to an extraordinarily simple and beautiful connection that can be established between the Heisenberg–Born–Jordan theory and the theory of integral equations. We shall see that all results of the theory may also be expressed in the form of integral equations; thus they obtain a formulation, which may appear more natural to the physicist—who is used to analytical methods

[264] Born and Jordan's paper (1925b) on quantum mechanics was published in late November 1925. Lanczos' acticle on the field-like representation of quantum mechanics was received by *Zeitschrift für Physik* on 22 December 1925 (Lanczos, 1926a).

—than the matrix formulation. Simultaneously, there arises with it also a continuum interpretation, which, as far as the *description* of empirical facts is concerned, can stand with equal rights beside the discontinuum interpretation, since there exists a unique correspondence between the two interpretations. However, what concerns the *physical interpretation* of the facts, that is, the real nature of the quanta, it is not excluded that the integral formulation might even be superior to the matrix formulation; the reason is that it has the advantage of being immediately consistent with the field conceptions, in fact it is even built on them, while the concept of the field is apparently far removed from the discontinuum interpretation. (Lanczos, 1926a, pp. 812–813)

For the purpose of establishing the relation between matrix methods and the methods of integral equations, Lanczos introduced a finite, bounded mathematical space of arbitrary extension and dimensions, whose points he denoted by s and σ. In this space he assumed a complete orthogonal set of eigenfunctions, $\phi^i(s)$ $(i = 1, 2, 3, \ldots)$, to be given; this set belongs to a *fundamental* kernel function, $K(s, \sigma)$, which is not specified for the moment. Thus, each eigenfunction, $\phi^i(s)$, solves an integral equation of the type of Eq. (213), i.e.,

$$\phi(s) = \lambda \int K(s, \sigma)\phi(\sigma)\, d\sigma, \tag{214}$$

with a particular eigenvalue λ_i. Moreover, two different eigenfunctions, $\phi^i(s)$ and $\phi^k(s)$, with different eigenvalues (i.e., $\lambda_i \neq \lambda_k$), satisfy the orthogonality relations

$$\int \phi^i(s)\phi^k(s)\, ds = \begin{cases} 1 & \text{for } i = k \\ 0 & \text{otherwise.} \end{cases} \tag{215}$$

(In Eqs. (214) and (215), the integrations extend over the entire space.) With the help of the functions $\phi^i(s)$ any given matrix, (a_{ik}), could be expressed in terms of a specific kernel function $f(s, \sigma)$—the latter being a function of the two variables s and σ, which is at least piecewise continuous—by the equation

$$a_{ik} = \int f(s, \sigma)\phi^i(s)\phi^k(\sigma)\, ds\, d\sigma, \qquad i, k = 1, 2, \ldots . \tag{216a}$$

Conversely, $f(s, \sigma)$ could be obtained from the elements of its associated matrix, (a_{ik}), as

$$f(s, \sigma) = \sum_{i,\, j=1}^{\infty} a_{ik}\phi^i(s)\phi^k(\sigma). \tag{216b}$$

Equations (216a, b) established a one-to-one correspondence between the infinite matrices, (a_{ik}), and the piecewise continuous kernel functions, $f(s, \sigma)$. Lanczos was thus able to re-interpret without difficulty *all* the matrix operations and

equations, which Born and Jordan had used to establish quantum mechanics, by means of integral equations containing 'field-like' kernel functions.[265]

As the most important matrix operations to be translated, Lanczos considered the operations of taking the diagonal sum of a matrix, the product of two matrices, the differentiation of a matrix function with respect to one of its matrix variables, and the definition of the time derivative of a matrix. The first of these operations was easy to express because taking the diagonal sum of a matrix, (a_{ik}), was just equivalent to integrating the associated kernel for equal points, $s = \sigma$, over the entire space.[266] The matrix product, on the other hand, was represented in Lanczos' formalism by folding the corresponding kernels, that is,

$$h(s,\sigma) = fg(s,\sigma) = \int f(s,\tau)g(\tau,\sigma)\,d\tau, \tag{217}$$

where $h(s,\sigma)$ is defined as

$$h(s,\sigma) = \sum_m a_{im}b_{mk}\phi^i(s)\phi^k(\sigma), \tag{217a}$$

and $f(s,\tau)$ and $g(\tau,\sigma)$ correspond to the matrices (a_{im}) and (b_{mk}). Also Born and Jordan's symbolic differentiation, i.e., the differentiation of a matrix function with respect to a matrix, Lanczos was able to translate properly (see Lanczos, 1926a, pp. 814–815). Finally, he turned his attention to the definition of the time derivative of a dynamical variable. In matrix mechanics, where a variable had been represented by a Hermitian matrix, the operation of taking the time derivative always produced a matrix whose elements were equal to the elements of the original matrix multiplied by the factor $2\pi i\nu(ik)$, where $\nu(ik)$ was the transition frequency (which was related to the difference of the energy terms, $W_i - W_k$). Lanczos claimed that in his field-like representation, where the quantum-mechanical dynamical variables were described by kernel functions, $f(s,\sigma)$, the time differentiation operation had to be expressed as

$$\dot{f}\,(s,\sigma) = (Kf - fK)(s,\sigma). \tag{218}$$

Equation (218) again possessed the form of an integral equation due to the folding occurring on its right-hand side. It involves the fundamental kernel function, $K(s,\sigma)$, which determines the orthogonal set of eigenfunctions, $\phi^n(s)$, with the associated eigenvalues, λ_n. Evidently, the λ_n must be related to the

[265] In the case of Hermitian matrices with complex elements, one had to replace the products $\phi^i(s)\phi^k(\sigma)$ in Eq. (215) by the product $[\phi^i(s)]^*\phi^k(\sigma)$, and the same in Eq. (216a) and (216b). The kernel functions, such as $f(s,\sigma)$, would then be Hermitian, satisfying the relation $f(s,\sigma) = f^*(\sigma,s)$, where the asterisk denotes the complex conjugate. Since the consideration of Hermitian matrices leads only to this small formal modification, Lanczos carried out his proofs always for real matrices. We shall follow his treatment below.

[266] The reason for this result is the validity of the orthogonality relation, Eq. (215), which ensures that $\int f(s,s)\,ds = \sum_{i,k} a_{ik}\int\phi^i(s)\phi^k(s)\,ds$ is equal to the diagonal sum $\sum a_{ii}$.

energy terms W_n, the eigenvalues of the energy matrix of the system under investigation, through the equation[267]

$$\lambda_n^{-1} = \frac{2\pi i}{h} W_n . \tag{219}$$

Hence the kernel $K(s,\sigma)$, apart from a constant factor, could replace the function $H(s,\sigma)$ representing the Hamiltonian in Lanczos' scheme. Equation (218) led to important consequences concerning the dynamical equations which had to be developed.

Lanczos did not have the slightest doubt that the dynamics in quantum theory had to be derived from a Hamiltonian principle. He was satisfied that the relevant Hamiltonian principle could be related to the fundamental integral equation (214) in the following way. He considered the special kernel function, $p\dot{q}(s,\sigma) - H(s,\sigma)$, which represented the Langrangian in his scheme, and expressed the Hamiltonian principle by the equation

$$I = \int [\, p\dot{q}(s,s) - H(s,s)\,]\, ds = \int [\,(pKq - pqK)(s,s) - H(s,s)\,]\, ds = \text{extremum}, \tag{220}$$

where I is the action integral and $H(s,\sigma)$ denotes the Hamiltonian kernel function. This formulation implied, of course, that Lanczos now *identified the (so far) purely mathematical space of points s with the physical space of four-dimensional (space-time) points x*. Then Eq. (220) was simply equivalent to the extremum principle, Eq. (88b), of Born and Jordan.[268] Evidently, it is possible to derive the Hamiltonian equations of motion from the extremum principle, Eq. (220), by variation of the action-integral with respect to $p(s,\sigma)$ and $q(s,\sigma)$, the kernel functions representing the quantum-mechanical momentum and position variables. For this purpose, Lanczos had just to consider $H(s,\sigma)$ as a function of $p(s,\sigma)$ and $q(s,\sigma)$; then he obtained equations of motion which were equivalent to those in matrix mechanics.[269]

In varying the action-integral for obtaining the equations of motion the quantum conditions must be satisfied. Lanczos reformulated the Born–Jordan quantum condition, Eq. (67), as the integral equation

$$(pq - qp)(s,\sigma) = \frac{h}{2\pi i} E(s,\sigma), \tag{221}$$

[267] Lanczos omitted the factor $2\pi i$ (in Lanczos, 1926a, p. 816, Eq. (21)) because he defined $\dot{f}(s,\sigma) = (2\pi i)^{-1} \cdot df(s,\sigma)/dt$.

[268] Since the diagonal sum of a matrix is replaced in Lanczos' scheme by the s-integral of the corresponding kernel function if taken for equal points $s = \sigma$, it follows that $D\,[p\dot{q} - H(p,q)]$ is equal to $(2\pi i/h)\int[(pKq - pqK)(s,s) - H(s,s)]\, ds$. Hence the two extremum principles, Eq. (220) and (88b), are equivalent.

[269] Thus, for instance, the equation for $\partial H/\partial p$ was replaced by the equation $(\partial H/\partial p)(s,\sigma) = (Kq - qK)(s,\sigma)$, which agrees with the corresponding matrix equation if the relation between the eigenvalues of $K(s,\sigma)$ and $H(s,\sigma)$ is taken into account.

where $E(s,\sigma)$ is the 'unit kernel' ('*Einheitskern*,' Lanczos, 1926a, p. 818), which has the property that all functions $\phi(s)$ are eigenfunctions (satisfying a linear integral equation of the type of Eq. (214) with $E(s,\sigma)$ as the kernel) with the eigenvalue $\lambda = 1$. He observed immediately that the function $E(s,\sigma)$ was not regular in its variables; it had, in particular, the property that it was zero everywhere except when $s = \sigma$, where it became infinite in such a way that the σ-integral assumed the value unity for every given s.[270] The occurrence of the irregular kernel function $E(s,\sigma)$ bothered Lanczos, and he stated in his paper: 'From this we see that the functions p and q cannot stay finite in the entire region, since otherwise their products would also become finite everywhere. Thus we arrive at the following interesting conclusion: *If the functions* $p(s,\sigma)$ *and* $q(s,\sigma)$ *should remain finite everywhere, then the Born–Jordan quantum condition cannot be valid rigorously, but rather only in an arbitrarily good approximation*' (Lanczos, 1926a, p. 819).[271]

In order to weaken the quantum condition Lanczos proposed two possibilities. The first was to replace the unit kernel $E(s,\sigma)$ by a regular kernel function having the following property: for fixed s this function would be finite in a small sphere around the central point s such that the σ-integration would yield unity (Lanczos, 1926a, p. 819). Hence the quantum condition of Born and Jordan would not be satisfied in the strict sense for sharp points s and σ, but only if the space resolution was coarse enough that the small sphere did not dissolve into discrete points. In a note added in proof to his paper, Lanczos described another idea for obtaining regular kernel functions in quantum mechanics (Lanczos, 1926a, pp. 827–830). In particular, he proposed to modify the quantum condition, Eq. (67), in such a way that it would remain valid only for finite submatrices of $\mathsf{pq} - \mathsf{qp}$; that is, the diagonal elements with higher indices would deviate appreciably from the standard value $h/2\pi i$ and go to zero. This would mean that $(h/2\pi i)E(s,\sigma)$ is replaced by the regular function $U(s,\sigma)$, which has the property that the integral, $\int U(s,s)\,ds$, is zero.[272] After the only irregular function of the scheme was thus eliminated, the functions $p(s,\sigma)$ and $q(s,\sigma)$ as well as all other functions describing quantum-mechanical variables could be treated as regular functions.

Lanczos did not worry about the violation of the Born–Jordan quantum condition in the case of matrix elements with high indices. Arguing from the point of view that in a finite universe, with only regular mass and charge distributions, the energy is always finite, he assumed that higher energy values

[270] For example, one also has $\lim_{\epsilon\to 0}\int E(s,s+\epsilon)\,d\epsilon = 1$. Hence $E(s,\sigma)$ had similar properties as the expression which Paul Dirac would later call a 'delta-function' (Dirac, 1927a).

[271] Another way out of this difficulty seemed to be to use the quantum condition (221), not directly but only when both sides of it were folded with a regular kernel, say $K(s,\sigma)$. Lanczos did not find this satisfactory, for it did not guarantee that the p and q were regular functions in his scheme.

[272] Evidently, the matrix corresponding to $U(s,\sigma)$, according to Eq. (216a), has the diagonal sum zero. This can be achieved, e.g., when one approximates the matrix by a finite, diagonal $(n \times n)$-matrix, having all elements equal to $h/2\pi i$, except the last one which assumes the value $(1-n)h/2\pi i$, and then takes the limit $n \to \infty$. (See Lanczos, 1926a, p. 829.)

corresponding to very high indices of the matrices do not occur in reality. All matrices, he concluded therefore, had to be finite and the regular function $U(s,\sigma)$ should always replace—up to a factor $2\pi i/h$—the unit kernel $E(s,\sigma)$ in Eq. (221).[273]

Lanczos' quantum-mechanical scheme may be summarized by two fundamental equations, the action principle and the quantum condition. Using the regular kernel $U(s,\sigma)$ in both instead of $E(s,\sigma)$, these equations assume the form

$$I = -\int \left[KU(s,\sigma) + H(s,\sigma) \right] ds = \text{extremum} \tag{220a}$$

and

$$(pq - qp)(s,\sigma) = U(s,\sigma), \tag{221a}$$

where $K(s,\sigma)$, $U(s,\sigma)$ and $H(s,\sigma)$ have to be considered as functions that are given.[274] Disregarding, for the moment, the modification of the quantum condition envisaged by Lanczos, there existed a complete equivalence between his scheme and that of Born and Jordan. Thus, in principle, it did not matter whether quantum systems were described by means of equations involving matrices or integrals over field-like kernel functions. However, in practical cases the choice of one method or the other would depend upon the nature of the problem under investigation.

And yet, there was a considerable conceptual difference between the two formulations. In matrix mechanics a mechanical system was completely defined by its Hamiltonian matrix and the commutation relations for its dynamical variables. In Lanczos' integral equation scheme one had to know the corresponding kernels $H(s,\sigma)$ and $U(s,\sigma)$ and, in addition, the fundamental kernel $K(s,\sigma)$. 'For the matrix interpretation,' Lanczos noted, 'only the eigenvalues of this kernel are of importance, while the orthogonal system of eigenfunctions associated with it remains completely undetermined' (Lanczos, 1926a, p. 820). Indeed, Eq. (219) meant only that the eigenvalues of the kernels $K(s,\sigma)$ and $H(s,\sigma)$ had to be, apart from a constant factor, exactly the same.[275] But from the mathematical theory it followed that different integral equations of the type of Eq. (214)—i.e., equations which have different kernels, $K(s,\sigma), K'(s,\sigma), K''(s,\sigma), \ldots,$ and a different set of eigenfunctions, $\phi^i(s), \phi'^i(s), \phi''^i(s), \ldots,$

[273] In drawing the above conclusion, Lanczos thought of the harmonic oscillator as a typical system in quantum mechanics. There are other systems, however, such as the hydrogen atom, in which infinitely many states occur without W_n going to infinity in the limit of large n.

Beginning with his earliest papers on general relativity theory, Lanczos had always argued in favour of nonsingular field solutions. His insistence on the regular solutions, $p(s,\sigma)$ and $q(s,\sigma)$, in quantum mechanics may be explained very well from this background.

[274] Equation (220a) is obtained by rewriting the first term, $\int[pKq - pqK](s,s)\,ds$, as $\int[K(qp - pq)](s,s)\,ds$, which can be done if all factors in the products are regular (i.e. in that case, the factors in the products pKq and pqK can be cyclically permuted to give Kqp and Kpq, respectively).

[275] Note that in the integral equation approach the eigenvalues of $H(s,\sigma)$ are W_n^{-1} rather than W_n.

may yield the same eigenvalues λ_i. If, however, the eigenvalues of different integral equations were identical, the associated sets of eigenfunctions, $\phi^i(s)$ and $\phi'^i(s)$ $(i = 1, 2, \ldots)$, would be connected by orthogonal transformations. Now it was clear why the eigenfunctions played no role in matrix mechanics because there two quantities related by orthogonal transformations had to be considered identical. Lanczos concluded: 'Therefore one can never assert anything from matrix equations about the basis functions $\phi^i(s)$' (Lanczos, 1926a, p. 820).

In a previous paper on the quantum problem, Lanczos had claimed that the eigenfunctions should have a physical meaning if the integral equation approach were to make any sense at all. This claim had grown out of his deeply felt scientific belief, for he had declared: 'The author confesses his adherence to that Pythagorean view of the world, which sees in numbers more than just a formal and convenient tool for the purpose of describing [empirical] facts, which rather recognizes in mathematics an unavoidable symbolic language of nature, and which assumes that a certain harmony exists between mathematical aesthetics and Reality' (Lanczos, 1925b, pp. 148–149). He was convinced that 'the progress of mathematical knowledge is simultaneously accompanied by an increasingly more deeply penetrating knowledge of Reality' (Lanczos, 1925b, p. 149). One of the most fundamental mathematical objects in the theory of integral equations was evidently the kernel, $K(s, \sigma)$, and the question arose whether nature provided any information about it beyond the eigenvalues λ_i. Lanczos now argued in his discussion of quantum mechanics:

> If all physical facts are of such a nature that they can always, in principle, only yield the coefficients of matrices, then the matrix representation deserves to be preferred (at least from the positivistic point of view!), for it does not introduce any fundamentally unrealizable element in the description of [empirical] facts. The situation changes, however, if the kernel possesses a physical meaning. In that case the field-like representation must be considered as the more adequate one, since the matrix formulation provides less information insofar as it can only furnish the eigenvalues of the kernel but leaves the system of eigenfunctions completely undetermined. (Lanczos, 1926a, pp. 820–821)

Evidently Lanczos, in agreement with his work on the electromagnetic radiation in a closed universe (Lanczos, 1925b, c), favoured the possibility that the kernel function $K(s, \sigma)$ occurring in his fundamental equation (214) had a physical significance and he made an effort to grasp it. He recalled the method he had employed earlier in general relativity theory for determining the kernel functions of the gravitational field (Lanczos, 1925a), and which he had later taken over in the case of the electromagnetic field (Lanczos, 1925c): it was to write down the differential equations for the fields, including all the source terms, and thereby constructing the kernels of the equivalent integral equations, which (i.e., the kernels) are identical to the Green's functions of the problem. By analogy to these cases, the kernel $K(s, \sigma)$ of Lanczos' field-formulation of quantum mechanics could be interpreted as a *universal* Green's function of the theory; that is, the systems which it described embraced all the atomic systems

existing in nature (each of which could be regarded as a subsystem in the approximation in which the interactions between the atoms were neglected). Hence λ_i, the eigenvalues of $K(s,\sigma)$, and the corresponding eigenfunctions $\phi^i(s)$ at once gave the energy states of all the atoms in mutual interaction.

Lanczos was further aware of the fact that, in constructing the fundamental kernel $K(s,\sigma)$ in quantum mechanics, one had to take into account the peculiar features of the quantum-theoretical fields, such as their dependence on two space variables, that were not common to classical fields. 'However, it must be noted,' Lanczos remarked, 'that the kernel functions $p(s,\sigma)$ and $q(s,\sigma)$ are associated with the system as the totality of all *possible* states, while a definite *real* quantum state is characterized by the usual field functions $p_i(s)$ and $q_i(s)$' (Lanczos, 1926a, p. 823, footnote 2). The latter could be obtained by a space integration over the fields as

$$p_i(s) = \int p(s,\sigma)\phi^i(\sigma)\,d\sigma,$$
$$q_i(s) = \int q(s,\sigma)\phi^i(\sigma)\,d\sigma. \tag{222}$$

Thus, the variables in the actual quantum states, $p_i(s)$ and $q_i(s)$, were obtained by folding the field kernels $p(s,\sigma)$ and $q(s,\sigma)$ with $\phi^i(s)$, the eigenfunctions of the fundamental kernel $K(s,\sigma)$.

Another unusual situation occurred with respect to the description of the space coordinate, which appeared in Lanczos' theory in two ways: first, as a dynamical variable $q(s,\sigma)$; second as a geometrical variable s or σ. Although this was a completely new, nonclassical feature of the theory that required proper interpretation, and although he could not yet provide any, Lanczos was not discouraged at all. Rather, he concluded his paper on the field-like representation of quantum mechanics with the words:

> This much, however, we do believe that we are allowed to conclude from the possibility of bringing the entire theory [quantum mechanics] in full accord with our field-like conceptions: that the modifications, which we must apply to our classical views in order to reach an understanding of the quantum problems, must lie in a totally different direction than could be characterized simply by the contrast between continuum and discontinuum; and that the solution of the quantum mystery should have hardly anything to do with a quantum-like re-interpretation of geometry or infinitesimal calculus. (Lanczos, 1926a, p. 823)

With this conclusion he stated clearly that the new development in quantum theory deviated from the line which it had taken during the previous years; not so much the discretization of the classical laws had to be achieved, but a more general modification which would *exhibit both discrete and continuous features*. Lanczos himself expected firmly to be able to contribute to the establishment of such a quantum mechanics.

In his second paper on quantum mechanics, submitted to *Zeitschrift für Physik* in the beginning of March 1926, Lanczos urgently pursued the problem of determining the fundamental kernel function $K(s,\sigma)$ or the fundamental, complete set of orthogonal eigenfunctions $\phi^i(s)$ (Lanczos, 1926b). He replaced this task by the equivalent one: If one starts out from the variational principle, Eq. (220a), in which $K(s,\sigma)$ only occurs folded with the kernel $U(s,\sigma)$, one may consider the fundamental set of eigenfunctions, $\phi^i(s)$, to be as much the solutions of the integral equation (214) as of a similar equation with $U(s,\sigma)$ as the kernel.[276] Lanczos, therefore, concluded: 'It is not the energy function [i.e., $K(s,\sigma)$] which we must consider as the given basis kernel, but the quantum function [$U(s,\sigma)$], and it must yield that system according to which the other functions should be expanded' (Lanczos, 1926b, p. 408).

The advantage of reformulating the problem as the determination of the fundamental set of eigenfunctions solving an integral equation with $U(s,\sigma)$ as the kernel function was evident. First it was the nonzero kernel, $U(s,\sigma)$, the so-called quantum function, which made the system under consideration, i.e., the universe, a quantum system; hence $U(s,\sigma)$ had to be treated as the primary concept in quantum mechanics. Second, unlike the energy function, which depends on the details of what constitutes the universe, the kernel $U(s,\sigma)$ was a global function, which was rather insensitive with respect to the specific properties of the systems constituting the universe.[277] Lanczos therefore believed that he had simplified the problem considerably. Moreover, the new situation agreed better with the principal idea which had guided his previous work on the quantum problem (Lanczos, 1925b, c): namely, the assumption that the structure of the universe, as expressed by the set of eigenfunctions $\phi^i(s)$, was responsible for the deviations from the classical description that were accounted for by the quantum functions $U(s,\sigma)$. As for the explicit determination of $U(s,\sigma)$ or $\phi^i(s)$, Lanczos was still not able to report any progress when he submitted his third paper on quantum mechanics to *Zeitschrift für Physik* in April 1926 (Lanczos, 1926c).[278]

[276] In matrix mechanics this situation is reflected by the fact that the matrices corresponding to the kernels $K(s,\sigma)$ and $U(s,\sigma)$ are simultaneously diagonal. Evidently, Lanczos took the kernel function $U(s,\sigma)$ as being different from the unit kernel, $E(s,\sigma)$, because the latter did not allow one to single out a fundamental set of eigenfunctions.

[277] For example, the quantum function $U(s,\sigma)$ was the same for all systems having one degree of freedom. For systems having f degrees of freedom, Lanczos proposed the same quantum function, i.e., $U(s,\sigma) = \sum_{k=1}^{f} (p_k q_k - q_k p_k)(s,\sigma)$. Born, Heisenberg and Jordan—as well as Kramers, Weyl and Pauli—had introduced the quantum condition for each degree of freedom; this corresponded to $U(s,\sigma) = (p_k q_k - q_k p_k)(s,\sigma)$, $k = 1, \ldots, f$. Lanczos believed that his definition had to be preferred (Lanczos, 1926b, p. 407).

[278] In this paper Lanczos investigated, especially, the possibility of representing the position and momentum variables in matrix mechanics by real matrices. he found that in this case the quantum function $U(s,\sigma)$ had positive and negative imaginary eigenvalues and its 'diagonal sum,' i.e., $\int U(s,s)\,ds$, was automatically zero. As a consequence, however, to each energy state W_n there corresponded another energy state $-W_n$. This seemed to be at variance with the spectroscopic observation, since one would also expect term sums, $W_n + W_m$, besides term differences, $W_n - W_m$, determining the frequencies. In the example of the perturbed harmonic oscillator, Lanczos proved that the transition elements from states with positive energy to states with negative energy were zero; thus, he showed that in this case no contradictions to the observation followed.

By the time Lanczos realized that his programme on a modified formulation of quantum mechanics had become stuck, Erwin Schrödinger's first paper on 'Quantization as an Eigenvalue Problem' appeared in *Annalen der Physik* (Schrödinger, 1926c). In his paper Schrödinger proposed a treatment of atomic systems based on continuous field functions similar to those of Lanczos' integral equation scheme. By imposing a boundary condition, Schrödinger obtained the orthogonal set of eigenfunctions for a specific atomic problem, the nonrelativistic hydrogen atom, as solutions of a differential equation. In his further papers, which appeared in a rapid sequence, Schrödinger was able to treat other atomic systems as well by solving their particular differential equations (Schrödinger, 1926d, f). Schrödinger's differential equation method was found to be equivalent to Lanczos' integral equation method; especially, Eq. (214) could be rewritten as a Schrödinger equation *if the fundamental kernel* $K(s,\sigma)$ *were identified with the Hamiltonian function* $H(s,\sigma)$. Lanczos immediately recognized that by means of this step his own field-like representation of quantum mechanics was completed in some sense, and he refrained from further work on this theory. 'Afterwards it was too trivial,' he recalled several decades later. 'It was no longer of interest, because Schrödinger came along and he did it. As it often happens, it is the second man who hits the nail on the head and not the first one' (Lanczos in Mehra, 1973b, p. 287).

The Schrödinger-equation method completed Lanczos' integral-equation reformulation of quantum mechanics in the sense that it made a specific assumption about the nature of the fundamental kernel, $K(s,\sigma)$, which determines the orthogonal set of basic eigenfunctions $\phi^i(s)$ according to Eq. (214). That is, one now described different physical quantum systems by different sets of orthogonal, normalized eigenfunctions, $\psi_H^i(s)$, where the latter obeyed a Schrödinger equation involving a given Hamiltonian operator H (the equivalent of Lanczos' Hamiltonian kernel function $H(s,\sigma)$). Thus, the Schrödinger functions, $\psi_H^i(s)$, characterize only the quantum states of a given system, like the hydrogen atom, provided it can be considered as being isolated from the rest of the universe. Since in our actual universe matter and energy are distributed very thinly, it was plausible to regard atomic systems as being isolated from the rest of the world to a certain extent. Therefore, the replacement of the kernel, $K(s,\sigma)$, by $H(s,\sigma)$, the Hamiltonian of the atomic system under investigation, in Eq. (214) provided a suitable description of the actual situation, and the success achieved by using the Schrödinger equation confirmed this assumption.

Lanczos, however, was not psychologically prepared to pursue his programme in this direction. For him the quantum problem had a cosmic dimension. In the beginning of 1925, in his search for a field theory unifying gravitation and electricity, he had concluded the existence of eigenoscillations in a closed universe as offering the explanation of certain quantum effects observed in electromagnetic radiation. Stimulated by the development of matrix mechanics in Göttingen he had then thought of formulating the quantum problem on the assumption that the universe as a whole was responsible for the behaviour of atomic systems. This explains why he chose to identify the kernel determining the

basic set of eigenfunctions with a universal quantum function, $U(s,\sigma)$, rather than with specific Hamiltonian functions, $H(s,\sigma)$. Schrödinger's wave mechanics did not achieve the goal Lanczos had set himself: it offered no way of calculating $U(s,\sigma)$, and assumed it to be given instead.[279] Lanczos, on the other hand, did not entirely give up his plan and returned to it several years later.[280]

Lanczos' work on quantum mechanics attracted Heisenberg's attention when he was looking for a way of dealing with the position variables that do not change periodically with time. But neither Heisenberg nor Pauli, who were both interested in enlarging the formalism of matrix mechanics, made use of the integral equation scheme to treat angular variables. And, surprisingly, Lanczos himself never considered any systems other than the purely periodic ones which had already been successfully treated by matrix methods. One reason why Lanczos restricted himself in this way was that his picture concerning the motion of particles stemmed from a generalized field theory; and *the* difficulty in that theory was how to explain the existence of something which corresponded in classical theory to the very small closed electron orbits in atoms.

Practically the most serious obstacle in Lanczos' approach was that it could not help in solving the actual problems of atomic physics until the fundamental kernel and the basic system of eigenfunctions had been properly defined. As soon as Schrödinger's work helped to replace this kernel by the Hamilton function or operator, the preferred approach to follow was based on his differential equation rather than on Lanczos' integral equation. The direct impact which Lanczos' work made on the quantum physics of the day was negligible; his idea of representing matrices by continuous field variables was not acknowledged

[279] In the Schrödinger equation the commutation relation is automatically satisfied on account of the structure of the momentum operator ($p = (h/2\pi i)(\partial/\partial q)$, where q is the position variable).

[280] In the years following 1926 Lanczos turned again to general relativity theory. Later, he became interested in the Dirac equation and worked on it especially during the year in Berlin that he spent close to his idol Einstein. The opportunity had arisen in 1928 because Einstein's assistant left and Einstein was looking for somebody to assist him in his mathematical work. Leo Szilard, a Hungarian physicist who was close to Einstein, intervened on behalf of Lanczos. 'At that time I was very much interested in the equation of motion in general relativity,' Lanczos recalled later. 'It was a sideline of his [Einstein's] investigations, so he thought it might be a good occasion for us to work together' (Lanczos in Whitrow, 1972, p. 49). Lanczos was awarded a fellowship by the *Deutsche Notgemeinschaft* to spend one year in Berlin, where he had the 'greatest experience' of his life, seeing Einstein 'almost daily and have discussions with him of both a scientific and a non-scientific nature.' (Lanczos in Whitrow, 1972, p. 49).

During his stay in Berlin, Lanczos considered the possibility of deriving the universal quantum function or quantum condition from a nonlinear generalization of the Dirac equation. In August 1929 his hopes of completing *his* field-like representation of quantum mechanics ran high. 'If the possibilities surmised here should be capable of being realized, then quantum mechanics would cease to be an independent discipline,' Lanczos remarked in conclusion of his renewed effort. 'It would fuse with a deepened "theory of matter", which must be built upon regular solutions of nonlinear differential equations, that is, ultimately merge into the "world equations" of the universe. The dualism of "matter and field" would then be overcome just as the dualism of "corpuscle and wave" ' (Lanczos, 1929, p. 493). Had Lanczos succeeded in deriving the quantum condition from field theory, he would have been able to answer at the same time Einstein's question, namely, whether field theory offered the possibility of solving the quantum riddle, in the affirmative sense (Einstein, 1923d). He never did, nor did anyone else after him.

beyond a mention in footnotes.[281] However, in the wider perspective of the development of quantum theory, Lanczos' was the first field-like approach to quantum mechanics; it raised many questions which, to some extent, were answered only later in the context of the physical interpretation of quantum mechanics and the systematic formulation of quantum field theory. In Lanczos' long scientific career, his concern with quantum mechanics remained a brief episode which lasted about one and a half years.[282] He first made use of the Schrödinger equation as late as 1930, when he applied it to calculate the effect of very strong electric fields on atoms; in this problem the usual perturbation methods failed and Lanczos developed an elegant mathematical method for obtaining an asymptotic solution (Lanczos, 1930).

V.3 The Operator Mechanics

At the end of October 1925 Max Born left Göttingen for the winter semester to give a series of lectures at the Massachusetts Institute of Technology in Cambridge, Massachusetts.[283] From 14 November 1925 to 22 January 1926 he

[281] For example, Schrödinger referred to him in his paper on the equivalence of matrix and wave mechanics (see Schrödinger, 1926e, p. 754, footnote 1). Note, however, that Schrödinger misunderstood Lanczos' approach, claiming that Lanczos' theory would lead to the wrong eigenvalues (i.e., to inverse energy eigenvalues). The same misunderstanding occurred when Pauli wrote to Jordan: 'As for Lanczos, my considerations [concerning the relation of wave mechanics to matrix mechanics] have only a few points of contact with his. He looks at a problem in which the energy values are the reciprocal eigenvalues, while here the energy values are the eigenvalues themselves. Further, in his treatment the functions, which like the Green's functions depend on two points, play an essential role; functions like that are not used here. I really believe that Lanczos' *Ansatz* is not of much use' (Pauli to Jordan, 12 April 1926). Apparently, both Schrödinger and Pauli did not think of the fact that the eigenvalues of Hilbert's integral equations correspond to the inverse eigenvalues of the Schrödinger equation.

[282] Lanczos' later scientific career developed as follows. He became *Privatdozent* at the University of Frankfurt during the winter semester 1926–1927 and was appointed *Extraordinarius* for theoretical physics at the same place by the end of 1932. Being on leave of absence at that time, however, he did not return to Germany, but remained at Purdue University, Lafayette, Indiana, where he had been a professor of mathematics since 1931. Lanczos remained at Purdue until 1946, working on polynomials, Fourier analysis and mathematical methods. (In the year 1943–1944 he was on the staff of the National Bureau of Standards, contributing to the *Mathematical Tables Project*.) In 1946 he joined Boeing Aircraft Company in Seattle, Washington, as senior research engineer. From 1949 to 1952 he worked at the National Bureau of Standards' Institute for Numerical Analysis at the University of California, Los Angeles. During 1953–1954 he returned to industry as a specialist in computing with North American Aviation. Having spent the year 1952–1953 as a visiting professor at the Institute for Advanced Studies in Dublin, Ireland, he was offered a permanent position as a senior professor there in 1954. From then on he made his home in Ireland, marrying Ilse Hildebrand in 1954. (He had been married earlier to Maria Rupp from 1927 until her death in 1939; they had a son, Elmar, who was born in 1933.) From 1954 onwards he travelled frequently, lecturing at many places, especially in the United States and in Hungary, the land of his birth. He died in Budapest on 24 June 1974.

[283] Max Born had mentioned his plan of going to Boston as early as April 1924 in a letter to Niels Bohr. At that time Born expected that he 'probably will spend half of the winter semester [1924–1925] in America (Boston)' (Born to Bohr, 16 April 1924). The plan was finally realized a year later. In this

delivered thirty lectures: the first twenty dealt with the structure of the atom and the last ten with the lattice theory of rigid bodies (i.e., of the solid state). They were all published under the title: *Problems of Atomic Dynamics* (Born, 1926d). In his lectures Born included accounts of the most recent discoveries of matrix mechanics that had been made by Heisenberg, Jordan and himself and by Pauli. At M.I.T. Born again met a young mathematician, Norbert Wiener, who had visited Göttingen during the previous summer and whose work 'had attracted a certain amount of attention' (Wiener, 1956, p. 108). Thirty years later Wiener recalled Born's arrival as follows:

> He was enormously excited about the new basis Heisenberg had just given for the quantum theory of the atom. This theory was an essentially discrete one, and the tools for its study consisted in certain square arrays of numbers known as matrices. The separateness of the lines [i.e., rows] and columns of these matrices was associated with the separateness of the radiation lines in the spectrum of an atom. However, not all parts of the spectrum of an atom are made of discrete lines, and Born wanted a theory which would generalize these matrices or grids of numbers into something with a continuity comparable to that of the continuous part of the spectrum. The job was a highly technical one, and he counted on me for aid . . . . I had the generalization of matrices already at hand in the form of what is known as operators. Born had a good many qualms about the soundness of my method and kept wondering if Hilbert would approve of my mathematics. Hilbert did, in fact, approve of it, and operators have since remained an essential part of quantum theory. (Wiener, 1956, pp. 108–109)

The new stage of development of quantum theory was particularly favourable for the collaboration between Born and Wiener, and just a few weeks later Born reported to Bohr about its success: 'I have also, I believe, made fundamental progress with it [i.e., quantum theory] with the help of a young mathematician, Wiener, who has just recently developed the mathematical methods that one needs for the theory of aperiodic motions' (Born to Bohr, undated, most probably early December 1925).

Norbert Wiener, Born's new collaborator, was born on 26 November 1894 in Columbia, Missouri. His father, Leo Wiener, an immigrant to the United States from Byelostok in White Russia, soon after Norbert's birth obtained a position at Harvard University, becoming a respected professor of Slavic language and literature there. Leo Wiener was keenly interested in his son's education and instructed him early in languages and mathematics.[284] A child prodigy, Norbert

---

context, it should be mentioned that at M.I.T. the custom existed 'to invite from time to time distinguished scientists and engineers to address its classes and students' (C. L. Norton in the foreword to Born, 1926d). Thus, for example, Peter Debye had spent some time there in the beginning of the year 1925 as a Foreign Lecturer.

[284]Though Leo Wiener and his wife Bertha, *née* Kahn, had three other children, two girls and a boy, the father's attention was very much devoted to Norbert. This resulted in a one-sided education for the boy who, in addition, was plagued by nearsightedness and physical clumsiness. Norbert inherited a somewhat aggressive character from his father.

made rapid progress in his studies, entering high school at the age of nine and graduating in 1906 at the age of eleven. He then entered Tufts College in Medford, near Boston, where he studied Greek, German, philosophy, mathematics, chemistry and biology, and obtained his bachelor's degree in 1909. After having enrolled in the Harvard Graduate School to study zoology, Wiener soon changed both his field and his university and began to study philosophy at Cornell University, where his father's friend, Frank Thilly, occupied the chair of ethics. In September 1911 Wiener returned to Harvard as a candidate for a doctorate in philosophy; for his thesis he chose a topic in mathematical logic and obtained his Ph.D. in spring 1913 at the age of eighteen.

After his doctorate Wiener was appointed by Harvard University to a travelling fellowship. He first went to Cambridge, England, where Bertrand Russell became his 'chief teacher and mentor' with whom he 'studied mathematical logic and a good many more general matters concerning the philosophy of science' (Wiener, 1956, p. 21).[285] On Russell's advice Wiener learned more about mathematics, both in Cambridge and later in Göttingen, where he spent the summer semester of 1914 while Russell was invited to lecture at Harvard.[286] In Göttingen, Wiener was exposed to an atmosphere of serious and sustained intellectual work, and he retained many fond memories of his stay there: of the meetings of the *Mathematische Gesellschaft* (The Mathematical Society) and the following bull sessions ['*Nachsitzungen*'] in *Café Rohn's*, of the excellent reading room established by Felix Klein, of the people he met in Göttingen and, above all, of his appreciation of David Hilbert's universal outlook.[287]

Upon leaving Göttingen, Wiener continued his studies in logic, first with

[285] Wiener attended two courses of Bertrand Russell, a reading course on *Principia Mathematica*, and another one, in which Russell presented his views of sense data (i.e., the data received by human senses) and in which he introduced such topics as Einstein's theory of special relativity. 'Besides his consciousness of Einstein's importance,' Wiener recalled, 'Russell also saw the present and future significance of electron theory, and he urged me to study it, even though it was very difficult for me at that time, in view of my inadequate preparations in physics' (Wiener, 1953, p. 194).

Besides Russell's lectures Wiener attended, for instance, a course given by Godfrey Harold Hardy, starting out from mathematical logic and theory of assemblages and extending up to the Lebesgue integral. Wiener found this course particularly satisfying because he had not received any training in rigorous mathematics earlier. 'Hardy, however,' recalled Wiener, 'led me through the complicated logic of higher mathematics with such clarity and in such detail . . . and he gave me a real sense of what is necessary for a mathematical proof.' (Wiener, 1956, p. 22).

[286] Wiener had already been in Germany during a European trip with his family in summer 1901, which took him to Cologne, Vienna and London. This time again his family came along; however, unlike Norbert, the other family members spent most of their time in Munich.

[287] Russell had suggested that in Göttingen Wiener might study mathematics with David Hilbert and Edmund Landau and philosophy with Edmund Husserl. While Wiener at least tried to follow a course given by Hilbert on differential equations and another by Landau on finite groups, he did not know enough German to follow Husserl's courses on Kant and phenomenology.

Later, Wiener did not remember whether he had seen Felix Klein in summer 1914. However, he became acquainted with two young mathematicians, Felix Bernstein and Otto Szasz. The latter especially took care of him in Göttingen, and Wiener was happy to return this help many years later when Szasz came to the United States.

Russell in Cambridge from fall 1914 to spring 1915, and afterwards, when war conditions made it advisable to return to America, with John Dewey at Columbia University in New York. During the academic year 1915–1916 he was invited as a Docent Lecturer at Harvard to give a series of lectures on mathematical logic, but he was not able to secure a position for himself in future there. Instead, after spending three years on various jobs at various places, he finally obtained, with the help of William Fogg Osgood—a colleague and friend of his father's at Harvard—the position of an instructor of mathematics at the Massachusetts Institute of Technology.[288] Thus, in fall 1919, Wiener began his academic career in earnest.

M.I.T. was not regarded as one of the best places for mathematics in those days; mathematics served mainly to implement engineering training there. However, this situation ideally suited the stranded logician Norbert Wiener, for he began to appreciate the importance of applying mathematics to solve engineering problems. Whenever he needed it he could, of course, draw upon the advice of pure mathematicians from neighbouring Harvard and other universities. Thus, for example, he obtained the first mathematical problem he worked on from Irving Barnett of the University of Cincinnati. Barnett directed his attention to the recent fruitful work on 'the generalization of the concept of probability to cover probabilities where the various occurrences being studied were not to be represented by points or dots in a plane or in a space but by something of the nature of path curves in space' (Wiener, 1956, p. 36). Wiener became interested in this and he investigated various possibilities of extending Lebesgue's integral for that purpose, looking simultaneously for the physical applications of the new technique. 'The papers which I wrote on this subject,' he remarked later, 'were, I believe, the first to discuss anything new—combining the Lebesgue technique of integration with the physical ideas of Gibbs' (Wiener, 1956, p. 39). Hence the work on Brownian motion constituted the new instructor's first successful contribution to mathematical physics.[289]

During his early years at M.I.T., Wiener frequently travelled to Europe and collaborated with some colleagues there. For example, in Strasbourg he worked with Maurice Fréchet for a few weeks before the beginning of the International

[288] As Wiener related in his autobiography (Wiener, 1956), his lectures at Harvard (on a postulational and constructive treatment of *analysis situs*) did not leave a favourable impression on George David Birkhoff, who pointed out to Wiener many difficulties in his treatment.

In the years following the docentship at Harvard, Wiener held for a year an instructorship of mathematics at the University of Maine in Orono; then he worked as a staff writer for the *Encyclopedia Americana* in Albany, New York; in 1918 he joined the ballistics staff, directed by Oswald Veblen, at the Aberdeen Proving Ground in Maryland; finally, at the end of the war, he did some journalistic work for the *Boston Herald*.

[289] Wiener published his early scientific papers on mathematical logic during his Cambridge stay in British journals; he continued to publish on various topics of logic, philosophy and mathematics without, however, drawing much attention. Wiener's work on the path integration (Wiener, 1921a, and later papers), on the other hand, was taken up and continued later by Bernard Koopman, John von Neumann and G. D. Birkhoff.

Congress of Mathematicians in September 1920. He also presented a paper at the Congress dealing with a normed vector space (Wiener, 1921b); this paper became the first publication in a field that was later called Banach spaces.[290] In 1922, 1924 and 1925 Wiener again visited Europe, renewing his acquaintances and establishing new contacts with mathematicians in France and Germany. Thus, he obtained a reputation abroad, especially through some work on the boundary value problem in potential theory; and Leo Lichtenstein, in an addendum to his *Encyklopädie* article on recent developments in the theory of elliptical partial differential equations, remarked with respect to the boundary value problem: 'With these problems . . . a series of papers by G. Bouligand as well as H. Lebesgue and, in particular, by N. Wiener is concerned' (Lichtenstein, 1924, p. 1334).[291]

Besides potential theory Wiener became more and more involved in another mathematical problem that led to some of his more important work. The problem in question arose when the electrical engineers at M.I.T. wanted to have a proper mathematical foundation of communication theory, that is, of the

[290] The Polish mathematician Stefan Banach presented the idea of Banach spaces already in June 1920 in his doctoral thesis; however, he published his work only in 1922. Wiener, on the other hand, took up the problem shortly before the International Mathematical Congress at Strasbourg, at which he presented his results (Wiener, 1921b); he was stimulated to do this work by Maurice Fréchet's postulational approach to topology. Wiener conceded later that Banach had slight priority in developing the concept. In 1920 the Banach spaces did not have any physical application; but later Wiener discovered that the differential space which describes Brownian motion was very closely related to them.

[291] On his visits to Europe Wiener got to know especially Henri Léon Lebesgue and George Louis Bouligand in France and Leo Lichtenstein, a distant cousin of his father, in Berlin. All three of them worked on potential theory in which Wiener had become interested. Wiener had noticed that the method of generalized integration, which had led him to work on Brownian motion, was also applicable to the problem of boundary conditions in potential theory. (For the origin of this work, see the story which Wiener told in his autobiography, Wiener, 1956, pp. 79–88, 90–93.) Lebesgue, however, already had studied the problem of boundary conditions in 1913. In December 1923 Bouligand submitted a note to the Academy of Sciences, Paris, on the solution of the Dirichlet problem, i.e., the problem of finding a solution of an elliptic differential equation with boundary values which connect continuously to the values in the interior of the domain (Bouligand, 1924a). Lebesgue took up the problem immediately and presented his own communication to the Paris Academy in its session of 21 January 1924, in which he pointed out the conditions under which the Dirichlet problem had no solution (Lebesgue, 1924a). In January 1924 Wiener also published a paper on potential theory (Wiener, 1924a) in the *Journal of Mathematics and Physics of the Massachusetts Institute of Technology* (which had been founded in 1922 and directed during its early years by Wiener and then by his friend Philip Franklin). When he saw the notes of the French mathematicians in print in the *Comptes Rendus*, Wiener knew that he had to hurry if he were to compete successfully in that field. Thus, he prepared a short note in French, assisted by a student who was better in French than he, and mailed it to Lebesgue for publication in *Comptes Rendus*. It arrived just in time, for Bouligand had meanwhile also obtained similar results and sent them to his former teacher Lebesgue. In keeping with the best traditions of the Paris Academy the two papers were presented at the same session of 24 March 1924 (Wiener, 1924b; Bouligand, 1924b), and Lebesgue added remarks on both notes (Lebesgue, 1924b). Unlike many other scientific competitions, this one ended with Wiener and Bouligand becoming friends and collaborators over many years. 'Bouligand was even more prepared than I was to admit the somewhat greater completeness of my paper,' Wiener noted thirty years later, 'and we made arrangements to meet when I should come to France for a visit' (Wiener, 1956, p. 93).

theory describing the electrical oscillations which account for the transport of messages along a telephone line, and Dugald C. Jackson, Chairman of the Electrical Engineering Department at M.I.T., sought the help of his mathematical colleagues. Until then the mathematical basis for dealing with this problem had been the so-called harmonic analysis based on Fourier series and Fourier integrals. It enabled one to deal *either* with strictly periodic oscillations *or* with oscillations which had a very general structure, provided their amplitudes could be regarded as being very small in the remote past and in remote future. The electrical oscillations which occurred in a telephone line were of neither type; hence it became Wiener's task to extend harmonic analysis for that purpose. Wiener knew that the engineers had previously used with considerable success another method in the same context: *the operator calculus* of Oliver Heaviside. Wiener now combined the two methods: he generalized harmonic analysis, and with its help he put the Heaviside formalism on a more secure foundation.

In order to understand Wiener's task we must briefly describe the operator methods, especially those of Heaviside. Gottfried Wilhelm Leibniz and Joseph-Louis Lagrange may be considered as the founders of operator methods.[292] Their ideas were taken up and further developed in the late eighteenth and early nineteenth centuries by numerous mathematicians; however, the lack of care in handling the mathematical operations involved gave rise to many errors in these early papers, and it became imperative to establish formal rules which were satisfied by the specific operations in question. This was achieved during the first half of the nineteenth century, principally by François Joseph Servois, Robert Murphy and George Boole.[293] As a particular example, the action of the nth power of the operator $D\ (= d/dt)$, i.e., the differentiation with respect to the variable parameter t, on a function $u(t)$ was considered, with n assuming both positive and negative values or even rational values. It was this operator which Heaviside used primarily in the electromagnetic theory with great success but, at

[292] Leibniz noted the resemblance between the formulae of the n-fold differentiation of the product $x(t)y(t)$ of two functions, $x(t)$ and $y(t)$, with respect to the common variable t, on one hand, and the nth power of the algebraic sum of x and y, on the other (Leibniz, 1710). The observation of the '*modus operandi*' in $d^n(xy)$ is usually identified with the invention of operator calculus, although its first explicit formulation is found in a much later paper of Lagrange (1772). Lagrange especially proposed to use the calculus symbol in the same manner as an algebraic quantity.

[293] While Servois studied the properties of the operations used (Servois, 1814), Murphy analyzed the logical structure of the operator method (Murphy, 1837). In particular, Murphy distinguished between what he called the 'subject,' i.e., the symbol on which a certain operation had to be performed, the 'operation' itself represented by its own symbol, and the 'result' which may be connected with the former two by the algebraic sign of equality. Thus, for example, in the algebraic equation, $ab = c$, he considered the number a to be the symbol of the subject, b the symbol for the operation of multiplication by the number b, giving rise to c as the result. Murphy also noted that some of the operations he had treated commuted with each other while others did not. Finally, Boole not only turned operator methods into an effective tool for solving linear differential equations (Boole, 1844), he also founded mathematical logic on them, thereby seeking to explain all of mathematics on the basis of a few fundamental operations.

the same time, to the discomfort of his mathematically more rigorously oriented contemporaries.[294]

Heaviside's operator methods were as unusual as his education. Oliver Heaviside, a self-made scientist, had published a large number of papers on various aspects of electromagnetic theory before he recognized the importance of operator calculus to the investigation of the transients and discussed the use of 'operators in physical mathematics' in detail in two papers published in *Proceedings of the Royal Society* (Heaviside, 1893a, b).[295] In the application of his operator methods to solve problems in electrical circuit theory, Heaviside went beyond everything which had been established in this field before him. The main point was that in order to study the response to a given circuit having the (complex) *impedance* Z,

$$Z(i\omega) = R + i\omega L, \tag{223}$$

where R denotes the (real) *resistance*, L the (real) *inductance* and ω the frequency impressed upon the system (i being the imaginary unit), he introduced the step-function, $H(t)$, defined by

$$H(t) = \begin{cases} 0 & \text{for } t < 0 \\ 1 & \text{for } t > 1, \end{cases} \tag{224}$$

[294] The history of operator calculus is discussed, for instance, in the *Encyklopädie* article of Salvatore Pincherle, who contributed himself many papers on that subject (Pincherle, 1906). It should be mentioned that the early work on operator methods was frequently motivated by the desire of mathematicians to remove a metaphysical interpretation of the differential calculus, i.e, to avoid talking about infinitely small quantities, such as the differential dt; this goal was certainly achieved by treating, say, the symbols dt and d/dt like finite algebraic quantities.

The applications of operator methods, as discussed by Pincherle in his review article, were confined to the theory of differential and difference equations and to selected problems of number theory. If one operated with inverse powers of the differential operator D ($= d/dt$), one could also treat linear integral equations. Heaviside's operator calculus was not mentioned by Pincherle.

[295] Oliver Heaviside was born on 13 May 1850 in Camden Town, London, being the youngest son of the artist Thomas Heaviside and his wife Rachel Elizabeth, *née* West. Rachel's sister had married the physicist and inventor Charles Wheatstone. However, Wheatstone did not influence his nephew's education, and young Oliver learned physics entirely on his own. After leaving school, Heaviside served as a telegraph operator with the Great Northern Telegraph Company in Newcastle-on-Tyne, but he retired from this appointment in 1874 because of increasing deafness. Supported by his brother and later on by well-wishers and a pension from the Government, Heaviside lived in a seaside cottage in Devonshire. Although he had not had the benefit of a formal higher education, he trained himself in mathematical physics and played an important role in the development of Maxwell's electromagnetic theory and its practical applications. He published his first paper already in 1873 in *Philosophical Magazine* and continued to contribute regularly in the following years. The value of his work was only gradually appreciated by engineers and physicists—such as Lord Kelvin, who made public some of Heaviside's results in his Presidential Address to the Institution of Electrical Engineers in 1889—but his reputation was firmly established after his papers were collected and published as separate books.

Heaviside was elected Fellow of the Royal Society in 1891. The Institution of Electrical Engineers made him Honorary Member in 1908 and awarded him the first Faraday Medal in 1921. He received an honorary doctorate from the University of Göttingen. Heaviside became a legendary figure long before he died in his seaside cottage at Torquay on 3 February 1925. His name has been attached to one of the reflecting ionized layers (of the ionosphere) surrounding the earth, whose existence he had predicted in 1902.

as a characteristic measure of action. That is, $H(t)$, the so-called 'Heaviside function,' describes an external potential difference of one unit (volt), impressed from the time $t = 0$ onwards. The response, $A(t)$, called the '*indicial admittance*' —i.e., the current resulting in the circuit because of the unit action—is then given by the solution of the linear differential equation

$$Z(D)A(t) = H(t) \qquad \text{with } D = \frac{d}{dt} . \tag{225}$$

The differential operator $Z(D)$ in Eq. (225) is obtained from the impedance function $Z(i\omega)$, Eq. (223), by writing the operator D instead of the variable $i\omega$. Heaviside claimed that Eq. (225) could be replaced by the operator equation

$$A(t) = \Omega H(t), \tag{226}$$

where the operator Ω is just the inverse of the operator $Z(D)$, and he proposed several rules for the practical solution of this equation. For example, if it is possible to expand Ω into a negative power series in D, say as $\sum_{r=n}^{\infty} b_r D^{-r}$ (with constant coefficients b_r), then one obtains as solution for $A(t)$ a positive power series in the time variable t, i.e., $\sum_{r=n}^{\infty}(1/r!)b_r t^r$ for $t > 0$ (with the same coefficients b_r). By carefully applying this rule and some others, Heaviside and his followers were able to solve the differential equation (225) for almost every electrical circuit under consideration.[296]

For a long time, Heaviside's operator calculus remained suspect for the mathematicians. One of the reasons was that the step-function $H(t)$ was not an analytic function; hence the replacement of the differential equation (225) by the operator equation (226) did not seem to be allowed under any conditions whatsoever. Those who insisted on mathematical rigour, on clear definitions of the operators and well-defined equations obeyed by them could not take Heaviside's solutions seriously.[297] Against these objections Heaviside held that 'mathematics is an experimental science, and definitions do not come first, but later on' (Heaviside, 1893b, p. 121). He was completely vindicated, for soon after his death in 1925 John Renshaw Carson and, independently, Norbert Wiener provided a complete justification of his operator methods.[298]

[296] Of course, in practical cases the observation helped that if Ω is a sum of two operators, Ω_1 and Ω_2, and the solutions of Eq. (226) with these operators, namely, $A_1(t)$ and $A_2(t)$, are known, then the response $A(t)$ may be obtained as the sum $A_1(t) + A_2(t)$.

[297] By the turn of the century two methods had been developed to define operators in a rigorous mathematical sense. First, linear operations of 'subjects' were considered as being similar to linear transformations of a point in an n-dimensional (or even infinite-dimensional) space. Otherwise, one used the method of 'Laplace transforms.' However, both methods were not applicable to deal with Heaviside's problem.

[298] Carson (1887–1940), an engineer of the American Telegraph and Telephone Company, worked mainly on transmission theory. In a series of papers, ending with a long memoir (Carson, 1925), he proved the validity of Heaviside's Eq. (226) by replacing it by the integral equation, $[Z(p)]^{-1} = p\int_0^{\infty} A(t)\exp(-pt)\,dt$, with $p = i\omega$. (The equivalence of this equation with Eq. (226) is obtained by observing that the integral $p\int_0^{\infty} H(t)\exp(-pt)\,dt$ is unity; furthermore, it is assumed that all derivatives of $A(t)$ vanish at $t = 0$.) Then $[Z(p)]^{-1}$ represents a function similar to a Laplace transform. With the help of the above equation Carson succeeded in establishing Heaviside's rules.

The earliest attempts to establish Heaviside's operator calculus rigorously was based upon the use of integral transforms. It was found that, with certain restrictions, the operator Ω in Eq. (226) could be replaced by a well-defined integral, and from it the response function $A(t)$ could be obtained. For this purpose the concept of the Laplace transform had to be extended in such a way that it made sense to talk about the transform of the nonanalytic Heaviside function.[299] The possibility of an appropriate extension appeared to be likely in the early 1920s, but a systematic mathematical theory had yet to be developed. At this stage Norbert Wiener became interested in supplying the proof of Heaviside's methods and results. He decided to achieve his goal by making use of harmonic analysis because, in principle, the Fourier integral allows one to deal with nonanalytic functions. However, he noticed at once that 'the functions to which the Fourier integral applies are subject to severe restrictions at infinity' and 'these restrictions are much more severe than those to which Heaviside applies his method' (Wiener, 1926a, p. 560). He had, therefore, to *generalize* harmonic analysis.

In seeking to formulate a consistent theory of electric circuits, Wiener had to solve two problems: First, he had to extend the method of Fourier integrals in such a way that all time-dependent functions (e.g., the external function $H(t)$ and the response function $A(t)$) could be represented as harmonic integrals, say as $\int_{-\infty}^{+\infty} f(\omega)\exp(i\omega t)\,d\omega$; second, he had to define adequately the operators occurring in the theory, especially the differential operator $Z(D)$ appearing in Eq. (225). If the former task were achieved, the latter would be straightforward, simply because the differential operator D acts on a given harmonic component, say $f(\omega)\exp(i\omega t)$, like a multiplicative factor $i\omega$. However, it is difficult to understand that the inverse operator $[Z(D)]^{-1}$ (which is identical with Ω), when applied to the step-function $H(t)$, should yield results like the ones claimed by Heaviside. The reason is that, for instance, the solution of Eq. (226) for Ω—being a series in inverse powers of D—involves integrals in which $[Z(i\omega)]^{-1}$ has to be expanded over the entire range of ω. In spite of this difficulty Wiener succeeded in his task. In particular, he was able to generalize harmonic analysis in such a way that it applied to all functions $f(t)$, which are integrable and square integrable over any finite range of the variable t and, for large t, do not grow faster than a finite power of t.[300] And he showed further that a well-defined meaning could be given to the action of all the operators on the functions that Heaviside had envisaged (Wiener, 1926a).

The main idea, which enabled Wiener to carry out his programme, was to approximate any function $f(t)$ by functions $f_\lambda(t)$, whose expansion contained only frequencies up to a given one, ω_λ. The action of the operator Ω on $f_\lambda(t)$ was

[299] See, e.g., the method of Carson described in footnote 298.

[300] Evidently the step-function $H(t)$ is a function of this class since for large t (>0) it is constant (and equal to unity).

then given by the equation (Wiener, 1926a, p. 565)

$$\Omega f_\lambda(t) = \int_{-\infty}^{+\infty} f_\lambda(\tau) G(\tau - t, \lambda)\, d\tau, \tag{227}$$

where

$$G(\tau - t, \lambda) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} F_\lambda(i\omega)\Omega(-i\omega) \exp\big[\, i\omega(\tau - t)\,\big]\, d\omega. \tag{227a}$$

(The function $\Omega(-i\omega)$ is obtained by replacing the operator D in Ω by the variable $-i\omega$.) The function $F_\lambda(i\omega)$ restricts the absolute values of the limits of the frequency integration; $F_\lambda(i\omega)$ is basically unity for $|\omega| < \omega_\lambda$ and approaches zero in a smooth way for $|\omega| = \omega_\lambda$.[301] By substituting the appropriately approximated step-function for $f_\lambda(t)$, Wiener obtained the above-mentioned rule of Heaviside; that is, the application of the inverse powers of the operator D could be justified and yielded the well-known result. A difficulty still remained if one wanted to have the operator D acting directly on the step-function; this operation led to an object which, in those days, was considered as being ill-defined.[302] On the other hand, any positive noninteger power of D could be applied to $H(t)$, yielding a well-defined function. Therefore, Wiener factorized any operator Ω containing a term proportional to D as a product operator, $\Omega/\sqrt{D} \cdot \sqrt{D}$, and applied it to $H(t)$ in two steps: first, he calculated the result of the application of $\sqrt{D}$ on $H(t)$; then he expanded the operator $\Omega/\sqrt{D}$ in a Fourier series, finally obtaining the result (Wiener, 1926a, p. 579)

$$\left(\frac{\Omega}{\sqrt{D}} \cdot \sqrt{D}\right) H(t) = \frac{1}{\sqrt{\pi t}} \left[M(0) - \frac{1}{2} \frac{M'(0)}{t} + \frac{1}{2} \cdot \frac{3}{2} \frac{M''(0)}{t^2} - + \ldots \right], \tag{228}$$

where

$$M(i\omega) = \frac{\Omega(i\omega)}{\sqrt{i\omega}}\,. \tag{228a}$$

($M'(0)$ and $M''(0)$ denote the first and second derivatives of the function $M(i\omega)$ with respect to the variable $i\omega$ taken at the point zero.) From Eq. (228) Wiener concluded that the positive powers of D always operate on the step-function to give a small response in the distant future (for t very large), in agreement with an earlier assumption of Heaviside's.[303]

[301] One may replace $F_\lambda(i\omega)$ in Eq. (227a) by the operator $F_\lambda(D)$. This operator played an important role in Wiener's considerations. For example, the action of the product operator $F_\lambda(D)D$ on an arbitrary function can always be shown to give a well-defined result.

[302] Paul Dirac later on identified such an object as the delta-function (Dirac, 1927a).

[303] Since the function $\sqrt{i\omega}$ possesses a cut on the axis of positive imaginary $i\omega$-values, Wiener distorted the path of the frequency-integration in Eq. (227a) slightly to the right of the cut, calling this procedure a 'backward' or 'retrospective' operation. He noticed that all operators Ω, stemming from electrical circuits, were retrospective; that is, the corresponding functions $\Omega(i\omega)$ became analytic in the complex right-hand plane of the variable $(i\omega)$.

After successfully applying his generalized harmonic analysis to establish Heaviside's operator calculus, Wiener continued to extend its use in different directions, both physically and mathematically. In this way he discovered, for example, that rather irregular functions, such as those which describe the time independence of noise in an electrical line, or the intensity of white light emitted by a continuous source, could be treated by means of the generalized harmonic analysis (Wiener, 1926b, c). Then, however, arose the problem of investigating the frequency spectrum of the function involved. In particular, Wiener became interested in the problem of describing a continuous spectrum, and he succeeded in doing so by again using the mathematical tool of Lebesgue's integral in a functional space of paths, the latter having properties similar to the space which he had introduced in the theory of Brownian motion (Wiener, 1926b, c). Not only was he able to examine in many cases the 'spectral intensity' of the functions $f(t)$, but he also showed how an essential *statistical feature* enters into the analysis of phenomena like white light.[304] Although the work on generalized harmonic analysis had a considerable impact on mathematical physics later on, it did not influence the development of quantum mechanics; only by helping Born to apply operator methods did Wiener contribute to the new theory in an important way. This was the more surprising as in summer 1925 when Wiener had passed through Göttingen and had presented there a seminar on his recent attempts at generalizing harmonic analysis it had seemed that this new mathematical tool might be appropriate for formulating quantum-theoretical problems.[305]

[304] The problem of the analysis of white light in terms of the frequency spectrum had been studied already a long time before Wiener, especially by Lord Rayleigh and Arthur Schuster. A major difficulty had turned up when one wanted to explain why the intensities of two independent light sources added, despite the fact that in electromagnetic theory it is the amplitudes of the waves that add up. Lord Rayleigh had proposed that the reason was the following: different parts of a source emitting a continuous spectrum are not correlated in phase, hence the interference terms in the intensities drop out (Rayleigh, 1880). Schuster, on the other hand, had introduced special methods to perform the analysis of light and its intensities with the help of the so-called 'periodogram' (Schuster, 1894, 1897).

In 1926 Wiener finally put Schuster's intuitive procedure on a secure mathematical foundation. In agreement with Schuster's intent, he considered all Lebesgue measurable functions $f(t)$ which were square integrable and derived their spectral intensities.

[305] Wiener had revisited Göttingen, the place where he had studied in 1914, on his European trip in 1924. At that time he had found that his results on potential theory interested the Göttingen mathematicians, especially David Hilbert and Richard Courant. The interaction between Wiener and the Göttingen scientists became stronger a year later. As Wiener recalled: 'In 1925, after a little mountain climbing with [J. W.] Alexander, of Princeton, I again visited Göttingen, where my work on generalized harmonic analysis was beginning to attract real attention' (Wiener, 1956, pp. 95–96). Wiener gave a seminar on this subject. There is an interesting episode connected with this seminar. When, after the session, people gathered in *Café Rohn's* to have supper, Hilbert is reported to have said:

> The speeches that are given nowadays are so much worse than they used to be. In my time there was an art to giving speeches. People thought a lot about what they wanted to say and their talks were good. But now the young people cannot give good talks any more. It is indeed exceptionally bad here in Göttingen. I guess the worst talks in the whole world are given in Göttingen. This year especially they have been very bad. There have been—no, I have heard no good talks at all. Recently it has been especially bad. But now, this afternoon, there was an exception . . . . This afternoon's talk was the worst there ever has been. (Reid, 1970, pp. 169–170)

In fact, in the beginning of the summer semester of 1925 Max Born had already turned his attention to the treatment of those phenomena in which all the atomic motions could not be considered as being periodic. Especially, he had studied, together with Pascual Jordan, the action of aperiodic electric fields on an atom by reformulating the classical expressions for the resulting electric dipole moment in agreement with the correspondence principle (Born and Jordan, 1925a). In their work they had encountered the problem of representing the coordinates occurring in collision processes—in which the ingoing and outgoing electrons and atoms move along approximately straight-line trajectories—in terms of Fourier integrals. On hearing about Wiener's new mathematical theory, Born had thought that it might provide a general method which would allow the decomposition of all variables used for describing the motion of electrons in atomic systems, including the position variables of electrons colliding with atoms. Born, who was looking forward to visiting M.I.T. to give a series of lectures during the winter semester 1925–1926, had hoped that he would be able to carry on the work on aperiodic motion with the help of generalized harmonic analysis and the personal assistance of its expert, Norbert Wiener. However, by early November 1925, when Born arrived at M.I.T., the status of quantum theory had changed considerably from what it had been earlier that summer; he now brought along with him a new quantum theory, matrix mechanics, by means of which the problems of atomic physics were approached in a manner quite different from the one he had followed with Jordan just a few months previously. In making use of the results of the theory of infinite Hermitian matrices and quadratic forms, Born had recognized the possibility of being able to describe the continuous spectra emitted by atoms; yet, in formulating matrix mechanics in collaboration with Heisenberg and Jordan, he had not succeeded in developing a suitable method on which the calculation of continuous energy states could be based. An even more serious difficulty remained: the matrix methods failed completely in dealing with the problems of atomic collisions. Born still hoped that Wiener's mathematical technique, which the latter had presented in his Göttingen seminar, might provide the key for dealing with these questions, and he counted on Wiener's assistance during his stay at M.I.T. Norbert Wiener, who had just then become interested in quantum theory, felt enthusiastic about working with Born. He immediately recommended the use of operator methods, the intimate connection of which with generalized harmonic analysis he had already explored, as the appropriate tool for treating aperiodic motions.[306]

Wiener had been led to discover the deeper mathematical justification of Heaviside's rules (in the operator theory of electrodynamics) by first recognizing

[306] When Wiener began to study operator methods he had in mind only the physical applications to problems of classical electrodynamics. (See the discussion in the introduction of Wiener, 1926a, pp. 557–562; the paper was submitted in April 1925.) Even when he stayed at Göttingen in summer 1925, he did not think of using his methods, whether of generalized harmonic analysis or of operator calculus, in quantum theory. As he recalled later in his autobiography: 'What I did not realize at that time was that my talk was closely keyed to the new ideas of physics which were about to burst into bloom at Göttingen in the form of what is now known as quantum mechanics' (Wiener, 1956, p. 97).

that it was possible to define more general operations than had been employed in the theory of integral equations up to that time. Now, in November 1925, he directed Born's attention to the fact that the Hermitian matrices of matrix mechanics 'could be regarded as operators, acting on vectors in a multi-dimensional space, and he suggested a generalization of matrix mechanics into a kind of operator mechanics' (Born, 1978, p. 226).[307] Born was attracted at once by Wiener's proposal because it connected well with the ideas he had developed in Chapter 3 of the three-man paper (BHJ, 1926). There he had discovered a considerable advantage in extending the matrix description of dynamical variables in such a way that the infinite matrix, say (q_{mn}) with $m,n = 1,2,3,\ldots,$ always occurred in equations of the type

$$y_m = \sum_{n=0}^{\infty} q_{mn} x_n\,, \qquad m = 1,2,3\ldots\,. \tag{229}$$

That is, it acted like a linear operator, as the mathematicians would say, on a sequence of variables or on a vector, (x_n), having an infinite number of components, in order to produce another such vector, (y_m). From Hilbert's theory of integral equations Born also knew that, if certain conditions were imposed upon the matrix elements, q_{mn}, and the vectors (x_n) and (y_m), Eq. (229) was equivalent to an integral equation; but the restrictions were very severe and the matrices used in matrix mechanics did not satisfy them. On the other hand, Wiener had just shown that operator equations could deal with physical problems of electrodynamics where the familiar integral equation methods failed (Wiener, 1926a). Born did not yet know much about operator methods. He somehow felt that the methods to be used had some resemblance to those of almost-periodic functions.[308] In any case, he entered into the collaboration with Wiener with optimism. They were soon able to establish an operator formulation of quantum mechanics which proved to be more powerful than the previous matrix methods, especially in dealing with such problems as those occurring in the description of atomic collisions. Before the end of the year they submitted papers on their results, both in English to the *M.I.T. Journal of Mathematics and Physics* and in German to *Zeitschrift für Physik*; these papers were identical in content and appeared a few months later in print (Born and Wiener, 1926a, b).

Their first task was to find the relation between Born's algebraic equations (229) and the corresponding operator equation that would arise in Wiener's formalism. In order to arrive quickly at the result, the authors turned to a system having only discrete eigenvalues, W_n $(n = 1,2,3,\ldots)$. Then they claimed: 'In

[307] The idea to connect infinite matrices with operators had already occurred earlier in the mathematical theory of operators. As we have mentioned in footnote 297, the operators had been interpreted as causing linear transformations of points in a finite- or infinite-dimensional space; hence they could be described in principle by infinite quadratic matrices.

[308] He expressed this feeling in a letter to Niels Bohr, where he remarked: 'Please convey my greetings to your brother Harald, and tell him that Wiener and I are using methods which are closely related to his "almost periodic functions"' (Born to Bohr, undated, probably early December 1925).

place of the transformation of an array of variables $x_1, x_2, \ldots,$ which lacks any very obvious physical interpretation, we may consider transformations of functions $x(t)$ of the time' (Born and Wiener, 1926a, p. 84).[309] The functions $x(t)$ were defined by the series

$$x(t) = \sum_{n=1}^{\infty} x_n \exp\left(\frac{2\pi i}{h} W_n t \right). \tag{230}$$

A conventional harmonic analysis into components of the frequency ν_n $(= W_n/h)$ could be carried out on $x(t)$, yielding the Fourier amplitudes x_n, due to the relation

$$x_n = \lim_{\tau \to \infty} \int_{-\tau}^{+\tau} x(t) \exp\left(- \frac{2\pi i}{h} W_n t \right) dt. \tag{231}$$

Similarly a function $y(t)$ was formed with the Fourier amplitudes y_m (connected with the frequencies $\nu_m = W_m/h$). With the help of these definitions Born and Wiener translated Eqs. (229) into a single integral equation, that is,

$$y(t) = \lim_{\tau \to \infty} \frac{1}{2\tau} \int_{-\tau}^{+\tau} q(t,s) x(s)\, ds, \tag{229'}$$

where the 'kernel' function $q(t,s)$ was given by the expression

$$q(t,s) = \sum_{m,n=1}^{\infty} q_{mn} \exp\left[\frac{2\pi i}{h} (W_m t - W_n s) \right], \tag{229'a}$$

in terms of the matrix elements q_{mn} and the frequencies W_m/h and W_n/h. Equation (229′) was finally interpreted as the operator equation

$$y(t) = qx(t), \tag{229''}$$

with the integral operator q given by

$$q = \lim_{\tau \to \infty} \frac{1}{2\tau} \int_{-\tau}^{+\tau} q(t,s)\, ds. \tag{229''a}$$

In Eq. (229″) the operator q acts on the continuous function $x(t)$, transforming it into the continuous function $y(t)$.

[309] In matrix mechanics the quantities or vectors (x_n) or (y_m) could not be related to a physical quantity. One had just identified the (x_n) as columns of the unitary transformation matrix which brings the Hamiltonian matrix to diagonal form. (Note that in matrix mechanics the lowest value of the indices m or n is 0 instead of 1.) Thus, the significance of the (x_n) had remained 'still very obscure' (Born, 1926d, p. 125). The German version of the above quote was even more definite, for Born and Wiener remarked: '*In place of the transformation of any array of variables* $x_1, x_2, \ldots,$ *which lack any obvious explanation, we may consider transformations of functions* $x(t)$ *of the time, which can be interpreted in physics more easily*' (Born and Wiener, 1926b, p. 174). They did not mean that the quantity $x(t)$ had an obvious meaning; it expressed only the hope that a physical interpretation might be easier to obtain.

With the help of this procedure a connection was established between the algebraic equations (229) and the operator equation (229″) provided, of course, the expression on the right-hand side of Eq. (229′a) existed. However, for the matrices (q_{mn}), which were considered in matrix mechanics, that was not generally true; for example, if for the q_{mn} one substituted the transition amplitudes of the harmonic oscillator, which grow with increasing indices (m and $n = m \pm 1$), then the double sums in Eq. (229′a) would diverge. Therefore, Born and Wiener dropped the requirement that the kernel $q(t,s)$ should be represented by a trigonometric series as in Eq. (229′a). At first, they decided to admit all integral operators of the form (229″a), where the kernel $q(t,s)$ is an arbitrary function, compatible with the convergence of the integral on the right-hand side of Eq. (229′). But even this generalization did not reach far enough, for Wiener knew from previous experience that very important operations in physics (notably electrodynamics) could not be represented in this way. Hence he proposed to Born that Eq. (229), instead of being generalized into Eq. (229′), should be replaced right away by an operator equation of the type of Eq. (229″); however, the operator q should be an *arbitrary linear operator*, which in general would not be written as in Eq. (229″a).[310]

Although Eq. (229″) admitted operators, which were much more general objects than the matrices which had been considered in matrix mechanics until then, Born and Wiener claimed that 'there corresponds to every operator a "matrix"' (Born and Wiener, 1926a, p. 86). In order to find the generalized matrix corresponding to a given operator, say q, they first generated the function $q(t,W)$ of time, t, and energy, W, by writing

$$q(t,W) = \exp\left(-\frac{2\pi i}{h}Wt\right) q \exp\left(\frac{2\pi i}{h}Wt\right). \tag{232}$$

The harmonic analysis of the function $q(t,W)$ was then made into components of the frequencies $(V-W)/h$; the corresponding amplitudes, $q(V,W)$, were given by

$$q(V,W) = \lim_{\tau\to\infty}\frac{1}{2\tau}\int_{-\tau}^{+\tau}\exp\left(-\frac{2\pi i}{h}Vt\right) q \exp\left(\frac{2\pi i}{h}Wt\right) dt. \tag{233}$$

The significance of these generalized harmonic amplitudes is not difficult to recognize. If, in Eq. (233), the energy parameters, V and W, assume only discrete values, W_m and W_n, respectively, with $m,n = 1,2,\ldots,$ then $q(W_m,W_n)$ be-

[310] A linear operator satisfies the equation, $q[c_1x_1(t) + c_2x_2(t)] = c_1qx_1(t) + c_2qx_2(t)$. Simple examples of linear operators are the multiplication by a constant or by a function $f(t)$.

That Eq. (229″), with q a linear operator, generalizes Eq. (229) can be seen at once from the example of the unit operator (i.e., the operator which transforms the function $x(t)$ into itself). Evidently, it is a linear operator. On the other hand, when one seeks to represent this operation by the integral equation (229′), one recognizes that no kernel, $q(t,s)$, which is a function of both time variables, t and s, in the conventional sense, can be used to describe it.

comes identical with the matrix element q_{mn}; hence all the values, which the function $q(V, W)$ can assume in this case, comprise the quantum-mechanical matrix (q_{mn})—or $(q(mn))$, as it was written in matrix mechanics—and the operator possesses a genuine matrix representation. The function $q(t, W)$ also has a straightforward meaning: its value for the particular energy parameter, $W = W_n$, becomes equal to the sum of elements in the nth column of the matrix $(q_{mn} \exp[(2\pi i/h)(W_m - W_n)])$. If, however, one or both energy parameters assume continuous values, the function $q(V, W)$ describes a matrix with (one or two) continuous indices, thus providing just the formalism which Born, Heisenberg and Jordan had searched for in vain a few months earlier. In this formalism the function $q(t, W)$—which Born and Wiener, in an obvious generalization of the case of discrete energy values, called 'the sum of the line [i.e., column] of the "matrix" of q corresponding to the value W of the energy' (Born and Wiener, 1926a, p. 87)—characterizes the system under consideration in the following manner: namely, $q(t, W)$ may be represented by a Fourier sum in case all energy states of the system are discrete; or by a Fourier integral when the energy spectrum is purely continuous; or by a sum of both when discrete and continuous energies occur.[311]

The description in terms of operators could account for even more general quantum systems than the ones treated earlier by Born, Heisenberg and Jordan. However, the function $q(V, W)$, as calculated from Eq. (233), does not exist (in the conventional sense) for all linear operators even if the associated function $q(t, W)$, Eq. (232), is perfectly regular. An example of such an operator is the differential operator D $(= d/dt)$. Evidently, its associated function, $D(t, W) = (2\pi i/h)W$, is finite for all systems having finite energies; but its harmonic decomposition, $D(V, W)$, is not a conventional function.[312] Like D, many operators, describing dynamical variables of systems having continuous energy eigenvalues, have a 'matrix' which is singular in a manner similar to $D(V, W)$. As a result, conventional harmonic analysis, using either Fourier sums or Fourier integrals, cannot be applied to the differential operator D or positive integer powers of it in such systems, and one has to use a procedure similar to the one proposed by Wiener for Heaviside's operators (see Eqs. (227) and (227a)).

Before developing their operator formulation of quantum mechanics, Born and Wiener defined the Hermiticity property of operators. To begin with they noticed that any linear operator might be expressed as a sum of terms, each of

[311] Born, Heisenberg and Jordan had already noticed a similar situation when they considered the contribution to the orthogonality relations for the columns of the transformation matrix. (See BHJ, 1926, p. 593.)

[312] Difficulties with the operator D had been noticed by Wiener already in his attempts to justify Heaviside's operator methods. In the present case one sees easily that the function, $(2\pi i/h) \cdot W \lim_{\tau\to\infty}(1/2\tau) \int_{-\tau}^{+\tau} \exp[(2\pi i/h)(W - V)t]\,dt$, exhibits a peculiar structure when the system under consideration has continuous eigenvalues. (In the case of discrete eigenvalues W_n, $n = 1, 2, \ldots$, $D(V, W)$ becomes identical with the original matrix, $((2\pi i/h)W_m\delta_{mn})$, which is a perfectly regular object.) The above function is zero everywhere except when V sharply becomes the value W.

which was a product of a function of time, $u_n(t)$, and a differential operator, $F_n(D)$, that is,[313]

$$q = \sum_n u_n(t) F_n(D). \tag{234}$$

Then they showed that the transposed operator, q^T, was $\sum_n F_n(D) u_n(-t)$. Consequently the Hermitian conjugate of q is $\sum_n F_n^*(D) u_n^*(-t)$. Thus, the authors stated the following condition of 'Hermitianness' (Born and Wiener, 1926a, p. 91, Eq. (24))[314]:

$$\sum_n u_n(t) F_n(D) = \sum_n F_n^*(D) u_n^*(-t). \tag{235}$$

After these technical preliminaries Born and Wiener turned to the equations of quantum mechanics. They remarked: 'We represent a coordinate by a linear operator q instead of by a matrix. This has the advantage of enabling us to treat cases where the matrix representation goes to pieces, even though we introduce the notion of a continuous matrix. We must seek for the meaning of the statement that an atom is represented by a number of operators in the fact that an atom is a mechanism the characteristic oscillations of which are in some way given when the coordinate operators are determined' (Born and Wiener, 1926a, p. 92). Similarly, they represented the momentum variable by an operator p and required the quantum conditions to assume the form

$$pq - qp = \frac{h}{2\pi i} 1, \tag{236}$$

where 1 is the unit operator, i.e., the operator which when applied to a function yields the same function. In agreement with matrix mechanics Born and Wiener further assumed that all operators representing the dynamical variables of a quantum system were Hermitian.

With the above formulation, however, the description of mechanical systems was not yet complete because, as Eq. (229″) showed, the operators had always to be applied to the functions $x(t)$. Born and Wiener, therefore, associated with each system a set of functions ('*zugehörige Funktionen*,' Born and Wiener, 1926b, p. 182), $\{x(t)\}$, and required that this set be closed with respect to all linear operations occurring in the system. That is, if in Eq. (229″) one replaced the operator q by any other Hermitian operator describing a dynamical variable of the same system, and applied it to any function $x(t)$ of the set $\{x(t)\}$, the resulting function $y(t)$ should also belong to $\{x(t)\}$.[315]

[313] This situation is a generalization of the one in electrodynamics where the operators Z and Ω are only functions of the differential operator.

[314] Evidently, the 'matrix elements' of a Hermitian operator q must satisfy the condition $q(V, W) = q^*(W, V)$. Born and Wiener showed that this is so provided Eq. (235) holds (Born and Wiener, 1926a, pp. 90–91).

[315] The closure relation for $\{x(t)\}$ ensures that all operations for a given mechanical system can be defined in operator mechanics. As long as the physical significance of the functions $x(t)$ is not known, the closure relation remains the only requirement that can be imposed on them.

Born and Wiener now *postulated* that the laws of quantum mechanics could be formulated in terms of the operators just as well as in terms of the matrices. In particular, the Hamiltonian equations of motion for each pair of variables, q and p, formulated as the operator equations,

$$\dot{q} = \frac{\partial H(p,q)}{\partial p} = \frac{2\pi i}{h}(Hq - qH), \tag{237}$$

$$\dot{p} = -\frac{\partial H(p,q)}{\partial q} = \frac{2\pi i}{h}(Hp - pH),$$

should determine the motion of a system whose Hamiltonian operator is H. Again, Eqs. (237) are a generalization of the corresponding matrix equations and, in agreement with matrix mechanics, H is *assumed to be an operator function of the dynamical variables, having the same functional dependence on the operators p and q as the classical Hamiltonian has on its dynamical variables*. In addition, the theory of canonical transformations and perturbation theory could be formulated as operator equations, being an obvious generalization of the corresponding matrix equations. Moreover, in analogy to matrix mechanics, the time derivative of any dynamical variable, say the operator q, might be replaced by the difference of operator products given by

$$\dot{q} = Dq - qD, \tag{238}$$

which was assumed to hold in all cases.[316] From Eq. (238) the frequency condition for the radiation emitted in the transitions between discrete states, in the transitions between discrete and continuous states, as well as in the transitions between continuous states, could be obtained. On the other hand, the conservation of energy followed from Eqs. (237) because the time derivative of the operator, $H(p,q)$, becomes identical to zero.

[316] When q describes a system having only discrete energy states, W_n, and the kernel of the operator, $q(t,s)$, is given by Eq. (229′a), Eq. (238) follows from the two equations

$$Dqx(t) = \lim_{\tau\to\infty}\frac{1}{2\tau}\int_{-\tau}^{+\tau}\frac{\partial q(t,s)}{\partial t}x(s)\,ds$$

$$= \lim_{\tau\to\infty}\frac{1}{2\tau}\int_{-\tau}^{+\tau}\sum_{m,n=1}^{\infty}\frac{2\pi i}{h}W_m q_{mn}\exp\left[\frac{2\pi i}{h}(W_m t - W_n s)\right]x(s)\,ds$$

and

$$qDx(t) = -\lim_{\tau\to\infty}\frac{1}{2\tau}\int_{-\tau}^{+\tau}\frac{\partial q(t,s)}{\partial s}x(s)\,ds$$

$$= \lim_{\tau\to\infty}\frac{1}{2\tau}\int_{-\tau}^{+\tau}\sum_{m,n=1}^{\infty}\frac{2\pi i}{h}W_n\cdot q_{mn}\exp\left[\frac{2\pi i}{h}(W_m t - W_n s)\right]x(s)\,ds.$$

The second equation holds if there are no contributions to the integrand at $\tau = -\infty$ and $\tau = +\infty$.

In Born and Wiener's operator mechanics the Hamiltonian operator exhibited properties which had no direct analogue in matrix mechanics. For instance, because of the operator identity

$$H = \frac{h}{2\pi i} D, \tag{239}$$

H could be replaced in the operator equations by the operator D up to an additive constant.[317] Then, with the help of Eq. (233), it followed for the 'matrix elements' of H that

$$H(V, W) = W \lim_{\tau \to \infty} \frac{1}{2\tau} \int_{-\tau}^{+\tau} \exp\left[\frac{2\pi i}{h} (W - V) t \right] dt. \tag{240}$$

The expression on the right-hand side becomes equal to W for $V = W$ and zero otherwise. Hence, the 'matrix,' $H(V, W)$, associated with the Hamiltonian operator, is automatically diagonal. However, this fact alone did not contribute to the solution of the fundamental problem of operator mechanics, that is, to determine the energy eigenvalues of a system. The reason was that as long as $x(t)$, the associated functions of that system, were not known, the energy eigenfunctions could not be found by simply applying the differential operator, $(h/2\pi i)D$, to $x(t)$. Since Born and Wiener were not aware of any procedure for obtaining these functions, they were obliged to calculate the eigenvalues in the same way as Heisenberg had done in the case of the anharmonic oscillator (Heisenberg, 1925c): that is, for a given system (determined, for example, by its Hamiltonian $H(p, q)$ or the Newtonian equations of motion), they integrated the equations of motion, taking into account the restrictions imposed upon the variables by the Hermitian property and by the commutation relations. The operators occurring in these equations were tacitly assumed to act on the functions $x(t)$ or on their harmonic components, $x(\omega)\exp(i\omega t)$, where the frequencies ω were related to the energy eigenvalues of the system as $\omega = 2\pi W/h$. Thus, for a given operator equation, an ordinary equation was obtained involving the still unknown 'Fourier amplitudes' $x(\omega)$.

With the generalization of matrix mechanics into operator mechanics Born and Wiener hoped that those systems in which continuous eigenvalues occurred could now be treated. Before pursuing this matter they demonstrated that the new methods yielded the same result as matrix mechanics for the simple system of the harmonic oscillator. They observed that the operator equation of motion in this case,

$$D^2 q + \omega_0^2 q = 0, \tag{241}$$

with ω_0^2 the force constant (the mass of the system having been assumed to be

[317] From a direct comparison of Eq. (238) and the first of Eqs. (237), it follows that the operators H and $(h/2\pi i)D$ coincide only up to a constant because in both equations differences occur.

unity), admitted the solution

$$q = \exp(i\omega_0 t)\phi(D) + \exp(-i\omega_0\tau)\phi^*(D - i\omega_0), \tag{242}$$

where ϕ(D) satisfied the difference equation[318]

$$|\phi(D)|^2 - |\phi(D - i\omega_0)|^2 = \frac{h}{4\pi\omega_0}\,1. \tag{242a}$$

The last equation provides the complete solution of the problem, as can be seen from letting both sides operate on the function $\exp(i\omega t)$. Then the algebraic equation

$$|\phi(i\omega)|^2 - |\phi(i\omega - i\omega_0)|^2 = \frac{h}{4\pi\omega_0} \tag{242a'}$$

follows, in which ϕ appears as a function of the variable $i\omega$. Equation (242a′) states that the absolute square of ϕ assumes finite, positive values only for discrete values of ω, i.e., $\omega_k = k\omega_0$, with $k = 0, 1, 2, \ldots$. Hence W, the energy parameter of the harmonic oscillator, becomes restricted to the discrete values $W_k = (h/2\pi)\omega_0 k$, with $k = 0, 1, 2, \ldots$.[319] Finally, for these discrete W-values, the functions ϕ take on the values

$$\phi\left(\frac{2\pi i}{h}\,W_k\right) = \left(\frac{h}{4\pi\omega_0}\right)^{1/2}(k+1)^{1/2}\exp(i\delta_{k+1}), \tag{243}$$

with the arbitrary phase δ_{k+1}.

The result, expressed in Eq. (243), indicates that the operator q (and all other dynamical operators of the harmonic oscillator) have discrete eigenvalues. All these operators act on functions $x(t)$, which can be written as infinite Fourier sums of terms, $\sum_{k=0}^{\infty} x(k\omega)\exp(ik\omega t)$. The set of functions, $\{x(t)\}$, is evidently complete because the application of the operators p and q on a term having the frequency $k\omega_0$ will produce terms with frequencies $(k\omega_0 \pm \omega_0)$, which, by assumption, occur already in the Fourier series.

The functions $x(t)$ had not appeared in matrix mechanics. To compare the results of the operator and matrix methods, Born and Wiener calculated the function $q(t, W_k)$, obtaining

$$\begin{aligned} q(t, W_k) = \left(\frac{h}{4\pi\omega_0}\right)^{1/2}\Big\{&(k+1)^{1/2}\exp\big[\,i(\omega_0 t + \delta_{k+1})\,\big] \\ &+ k^{1/2}\exp\big[-i(\omega_0 t + \delta_k)\,\big]\Big\}, \end{aligned} \tag{244}$$

[318] This result followed by taking into account the quantum condition, Eq. (236), for the operators q and $p\,(=\dot{q})$ and observing that both operators are Hermitian, i.e., they satisfy Eq. (235).

[319] The values W_k agree with the values of the energy only up to a constant (see footnote 317). Hence, there was no disagreement between the results obtained in operator mechanics and in matrix mechanics (see Eq. (106) in Section II.4).

which is indeed identical with the sum of the kth column of the position matrix of the harmonic oscillator.[320] Though the matrix methods had been developed specifically to deal with systems possessing discrete eigenvalues, Born appreciated the merit of employing operator methods for this purpose as well. In his lectures at M.I.T. he noted: 'It is advantageous to operate with this function [i.e., $q(t, W_k)$] rather than with the coefficients [i.e., matrix elements], because we then have the whole machinery of the calculus at our disposal, and differential or integral equations replace infinite sets of simultaneous equations in an infinite number of variables' (Born, 1926d, pp. 125–126).

With the application of operator methods to the harmonic oscillator no new results had been obtained. However, by treating the one-dimensional uniform motion of a particle of mass μ Born and Wiener dealt with a system which had so far remained outside the quantum-mechanical considerations, especially matrix mechanics. In operator mechanics, however, this problem could be handled without any difficulty: one had only to integrate the equation of motion,

$$\mu D^2 q = 0, \tag{245}$$

for the position operator q, under the restrictions imposed by quantum mechanics. The solution for the operator q and its conjugate momentum p are given by the equations

$$q = t\phi_1(D) + \phi_2(D), \tag{246a}$$

$$p = \mu\phi_1(D), \tag{246b}$$

where the operator functions $\phi_1(D)$ and $\phi_2(D)$ have to be determined. The Hermiticity requirement (see Eq. (235)) leads to the two equations

$$\phi_1(D) = \phi_1^*(D) \tag{247a}$$

and

$$\phi_2(D) = \left[\frac{d\phi_1(D)}{dD} + \phi_2(D)\right]^* \tag{247b}$$

which $\phi_1(D)$ and $\phi_2(D)$ have to satisfy.[321]

[320] From Eq. (244) one finds by conventional harmonic analysis of the absolute values of the matrix elements, $|q_{mn}|$, the result, $(h/4\pi\omega_0)^{1/2}[(m+1)^{1/2}\delta_{m+1,n} + m^{1/2}\delta_{m-1,n}]$, which agrees with the result of the matrix calculation (see Born and Jordan, 1925b, p. 878, Eq. (63)).

Note that in the English version of Born and Wiener's paper there were several mistakes (see, e.g., Born and Wiener, 1926a, p. 95, Eq. (37)), which were corrected in the German version.

[321] Equation (247a) follows from the Hermiticity property of the operator p. The transpose of the operator q is q^T, where

$$q^T = \phi_1(D)t + \phi_2(D) = t\phi_1(D) + \left[\frac{d\phi_1(D)}{dD} + \phi_2(D)\right].$$

In the example of the harmonic oscillator, the commutation relation or quantum condition, Eq. (236), was responsible for the fact that the energy assumed discrete values. How would the commutation relation affect the motion in the present case? On substituting the expressions for p and q, Eqs. (246) and (247), in Eq. (236), Born and Wiener arrived at an equation,

$$\frac{d\phi_1}{dD}\phi_1(D) = \frac{h}{2\pi i\mu}1, \tag{248}$$

for the operator $\phi_1(D)$.[322] Thus, the operator $\phi_1(D)$ satisfied a differential equation rather than a difference equation like Eq. (242a), as in the case of the harmonic oscillator. Therefore, the operator solution of Eq. (248), that is,

$$\phi_1(D) = \left(\frac{hD}{\pi i\mu}\right)^{1/2}, \tag{249}$$

depended on a continuous manner on the operator D.[323] Substituting this result in Eq. (246a) and taking into account Eq. (247b), Born and Wiener finally obtained the position operator representing the uniform motion in quantum mechanics as

$$q = t\left(\frac{hD}{\pi i\mu}\right)^{1/2} - \frac{1}{4}\left(\frac{h}{\pi i\mu D}\right)^{1/2} + \psi(D), \tag{250}$$

where the operator function $\psi(D)$ denotes the real part of $\phi_2(D)$, which cannot be determined from the above dynamical relations.

The right-hand side follows because of the equation

$$F(D)u(t) = \sum_{k=0}^{\infty}\frac{1}{k!}u^{(k)}(t)F^{(k)}(D),$$

with

$$u^{(k)}(t) = \frac{d^k u(t)}{dt^k} \qquad \text{and} \qquad F^{(k)}(D) = \frac{d^k F(D)}{dD^k},$$

which Born and Wiener called the 'Leibniz formula' ('*Leibnizsche Formel*,' Born and Wiener, 1926b, p. 179). Now one obtains the Hermitian conjugate by taking the complex conjugate of q^T. Since $\phi_1(D)$ is Hermitian, the equality of $(q^T)^*$ and q yields Eq. (247b).

[322] Since the operators $\phi_1(D)$ and $\phi_2(D)$ commute, the quantum condition (236) leads to the equation $(h/2\pi i)1 = \mu[\phi_1(D)t\phi_1(D) - t\phi_1(D)\phi_1(D)]$. The first term on the right-hand side may be rewritten with the help of the Leibniz formula, quoted in footnote 321, as $\mu[t\phi_1(D)\phi_1(D) + d\phi_1(D)/dD \cdot \phi_1(D)]$. Thus one term remains, which may be written as $\mu(d\phi_1(D)/dD)\cdot\phi_1(D)$ or $\mu\phi_1(D)\cdot(d\phi_1(D)/dD)$. Evidently, $d\phi_1(D)/dD$ is zero in the classical limit; consequently $\phi_1(D)$ is a constant then (independent of D) and $\phi_2(D)$ is a real function, due to Eq. (247b).

[323] Equation (248) is equivalent to the operator equation, $(h/2\pi i)\,dD = d[\phi_1(D)]^2$, whose integration yields the solution, $(h/\pi i\mu)D = \phi_1^2(D) + \text{const}$. By dropping the constant and choosing the positive sign of the root, one arrives at Eq. (249).

From the operator solutions for p and q the authors calculated the functions, $q(t, W)$ and $p(t, W)$, obtaining

$$q(t, W) = t\left(\frac{2W}{\mu}\right)^{1/2} + i\frac{h}{4\pi(2\mu W)^{1/2}} + \psi\left(\frac{2\pi i}{h}W\right) \tag{251a}$$

and

$$p(t, W) = (2\mu W)^{1/2}. \tag{251b}$$

Both assumed real values, apart from a time-dependent imaginary part in $q(t, W)$, which, according to Born and Wiener, could be neglected because 'if W is large, $h/4\pi(2\mu W)^{1/2}$ is small'; moreover, 'in this case the sum of the terms in a column really represents the motion' (Born and Wiener, 1926a, p. 98).[324] The 'matrix' $q(V, W)$ of the position operator is degenerate because, considered as a function of the variables V and W, it is zero everywhere except when V assumes sharply the value W. The function $q(t, W)$ cannot, therefore, be represented by a Fourier sum or integral, and the motion is free of any periodicity. Since the real parts of $q(t, W)$ and $p(t, W)$ showed the same dependence on time t and energy W as the corresponding classical variables of an accelerated particle, Born and Wiener considered them as the proper quantum-mechanical representations of position and momentum variables. They concluded that: 'Our representation is as close to the classical treatment as seems to be consistent with the foundations of quantum theory' (Born and Wiener, 1926a, p. 98).

The quantum-mechanical solution of the problem of uniform motion, notably Eqs. (251a) and (251b), which had been obtained by a straightforward application of the operator methods, cleared an important puzzle of the theory of atomic systems. This puzzle consisted in the observation that many features pertaining to moving electrons and atoms could be described perfectly by means of classical equations, though one expected that quantum theory would alter the description completely. Max Born, who was deeply concerned with these questions and had discussed them at length with Niels Bohr during the previous year, had good reason to be pleased with the results of his collaboration with Norbert Wiener; in particular, they had been able to demonstrate how, in spite of the restrictions imposed by quantum theory—especially, the commutation relations, Eqs. (236)—an unaccelerated mass-particle followed a classical trajectory. The position function, $q(t, W)$ in Eq. (251a), seemed to possess all the observed properties: in particular, it yielded the correct space-time interpretation, as can be seen by examining the structure of the 'kernel' function, $q(t, s)$, entering into the integral equation of the type of Eq. (229′). Born discussed this point in a letter to Niels

[324] In Eq. (251a), we have corrected the second term, i.e., an error of a factor 2 in the denominator under the square root. We have also replaced in the quotation the word 'row' by 'column,' in agreement with the (correct) German text ('*Spaltensumme*,' Born and Wiener, 1926b, p. 187).

Bohr: 'This $q(t,s)$ possesses even a much more beautiful physical significance than the matrix,' he wrote. 'It is closely connected with the probability that the coordinate q makes at the time t a quantum jump up and at time s a jump down. In the case of the uniform motion in a straight-line, for example, $q(t,s)$ is zero everywhere except at $t = s$ (one must formulate it mathematically correctly, which is not difficult); that is, if a jump up occurs, then it is followed by a jump down, hence no jump occurs at all' (Born to Bohr, undated, most probably early December 1925).[325] Born proposed to interpret this situation in the following manner: 'In this, and only in this, case can one state that $q(t,s)$ is not a probability of transition [jumping], but also a coordinate, a position in the sense of geometry. In all other cases the "positions" can be determined only approximately; in those regions, where the probabilities of transition are appreciable, one cannot fix any "position"' (Born to Bohr, *ibid.*). Thus, Born believed that the operator formulation had provided the key to an understanding of the relationship between quantum mechanics and the description of phenomena in space and time. By pursuing similar ideas he hoped to resolve the difficulties which still remained in incorporating the conception of the light-quantum in a consistent atomic mechanics. He concluded that 'for the moment, however, our mathematical formulae are in most cases wiser than we are; the formulae arise entirely automatically, the interpretation is often [more] difficult' (Born to Bohr, *ibid*).

Born and Wiener's operator formalism seemed to open a new avenue for approaching the problems of atomic physics. Thus, it should have been only a matter of time and skill before the operator equations describing realistic atomic systems would be solved. However, the events developed differently than Born had foreseen. Almost simultaneously with the paper of Born and Wiener, Schrödinger's paper on the wave equation appeared in print (Schrödinger, 1926c), and this equation proved to be a remarkably adequate tool for treating atomic problems, which finished off any attempt of further pursuing the methods of Born and Wiener. Many years later Born regretted the fact that he and

[325] The 'kernel' function $q(t,s)$ can be obtained in the following way. One first calculates the 'matrix' $q(V,W)$ as

$$q(V,W) = \left[\left(\frac{2W}{\mu}\right)^{1/2} \frac{h}{2\pi i} \frac{\partial}{\partial W} + \frac{ih}{4\pi(2\mu W)^{1/2}} + \psi\left(\frac{2\pi i}{h}\right)W\right] \cdot \lim_{\tau\to\infty} \frac{1}{2\tau}\int_{-\tau}^{+\tau}\left[\frac{2\pi i}{h}(W-V)t\right]dt.$$

The right-hand side of this equation describes an irregular function having the same structure as the matrix of the differential operator D; i.e., $q(V,W)$ is zero everywhere except when $V = W$. By substituting $q(V,W)$ into the equation for the kernel $q(t,s)$ which is obtained from Eq. (229′a) by changing the m, n-sums into integrals over V and W, one finds (apart from normalization factors) the double integral $\int dV \int dW\, q(V,W)\exp[(2\pi i/h)(Vt - Ws)]$. One integration, say over V, is trivial because it leads only to replacing the variable V by W. Hence $q(t,s)$ becomes proportional to $\int dW\, q(W,W)\exp[(2\pi i/h)W(t-s)]$. The final step, the W-integration, gives the function $q(t,s)$; evidently, it assumes the value zero everywhere except when $t = s$.

Note that we have replaced Born's notation of $Q(t,\tau)$ by $q(t,s)$, in agreement with our previous notation (which is the one used in the paper of Born and Wiener).

Wiener had not developed operator mechanics in the direction of wave mechanics. He remarked then: 'First we did the right thing. We expressed the energy as d/dt and we wrote the commutation law for the energy and time as an identity by applying $(td/dt - dt/dt)$ to a function of t. It was absolutely the same as for p and q, but we did not see that. I shall never forgive myself, because if we had done this, we would have had the whole wave mechanics at once, a few months before Schrödinger' (Born, Conversations, p. 16). Born even went on to blame Wiener for the clumsy parts of their formalism and lamented that: 'In one case he saw that it was a simple differential representation [i.e., for the energy in terms of d/dt], but in the other case [i.e., for the operators p and q] he insisted on the very clumsy integral equation [Eq. (229′)], which I did not quite understand. And we could understand only after having the wave theory [of Schrödinger]. Then we knew why the exponential factors [occurring, e.g., in the function $x(t)$] came into these integrals' (Born, Conversations, p. 17).

Though one may sympathize with Born and Wiener's failure to see, in November 1925, the connection between p, the momentum operator, and d/dq, the operator denoting the derivative with respect to the position operator, which plays an important role in wave mechanics, several reasons can be given for their failure. To begin with, one finds that nowhere in the Born–Wiener paper or in Born's letter to Niels Bohr of December 1925 was there any mention of a commutation relation between energy and time; Born and Wiener were therefore not likely to notice the analogy of this relation with the standard pq-commutation relations.[326] Second, in order to consider a relation between the momentum p and d/dq as being analogous to the relation between the energy W and d/dt, one has to regard the position variable q on the same footing as the time parameter t. To Born, such a treatment would have been unacceptable because it seemed to contradict the principle on which matrix mechanics had been founded; namely, the fundamental difference between dynamical variables and mere parameters. However, Schrödinger, who developed wave mechanics on the basis of concepts quite different from those of matrix mechanics, introduced the coordinate q and the time t on the same footing. In Schrödinger's theory, the ideas of de Broglie concerning matter waves played an important role. It is surprising that Born, who knew about the possibility that mass-particles might exhibit wave properties—e.g., from Einstein's work on quantum statistics (Einstein, 1925a), Louis de Broglie's thesis (1924e, 1925), and Walter Elsasser's explanation of the Ramsauer effect (Elsasser, 1925)—did not think of it when he dealt with the uniform motion of a particle in operator mechanics. Had he tried

[326] Of course, it would not have been difficult to find such a relation from the formulae of Born and Wiener. For example, one might have just substituted for q the time t in Eq. (238) to obtain the relation, $Dt - tD = 1$. And, with the help of Eq. (239), the equation, $Ht - tH = (h/2\pi i)1$, would then have followed.

A commutation relation of energy and time was first considered by Wolfgang Pauli after he had learned the results of Born and Wiener (Pauli to Heisenberg, 31 January 1926).

to incorporate the wave aspects into his formalism he might have been led to consider operators which could be applied to functions periodic in both the time and the position variables. However, in winter 1925–1926 de Broglie's waves did not occupy his mind; in fact, he ignored a hint concerning them that had occurred in the solution, Eqs. (250) and (251), for the position variable of the uniform motion.[327]

As for Wiener, he had only recently become interested in quantum theory and was not fully acquainted with its problems. The physical ideas involved in formulating operator mechanics came from Born; Wiener executed the mathematical parts of Born's programme on the basis of his knowledge of operator methods.[328] The path to wave mechanics and the Schrödinger equation was thus not close to Wiener's thinking. Moreover, in Heaviside's electrodynamics it had not been necessary to introduce—in addition to the pair of operators t and D—another pair consisting of the position operator q and the differential operator $\partial/\partial q$. And the examples, which he treated in quantum mechanics with Born, did not require such an extension either. Wiener always remained pleased with the fact that, at the right historical moment, he had developed the mathematical methods by which matrices could be generalized, and that with their help he had assisted Born in solving an important physical problem in quantum mechanics. He was also happy with the fact that 'operators have since remained an essential part of quantum theory' (Wiener, 1956, p. 109). But in 1926 and 1927 when 'quantum mechanics went into an active phase of growth in which young men like Heisenberg himself, Dirac, Wolfgang Pauli and John von Neumann were making new discoveries almost every day,' he stayed out of it. He found that: 'This feverish atmosphere is not one in which I function well, nor did I feel the need of intervening in a subject which was already so completely handled' (Wiener, 1956, p. 109).[329]

Nevertheless, it is surprising that in 1926, when the full power of operator methods began to be displayed in wave mechanics, Wiener's name was no longer associated with the advances in quantum mechanics. Wiener did go to Göttingen in April 1926 and stayed there during the summer. Though this visit turned out to be less successful than Wiener and his hosts had hoped, he did have a fruitful

[327] As mentioned above, Born and Wiener, in the position function, Eq. (251a), neglected the imaginary term, $(i/4\pi)(h/(2\mu W)^{1/2})$, because they considered it to be too small if W was large. However, by introducing the de Broglie wavelength λ $(= h/(2\mu W)^{1/2})$, the second factor in the imaginary term is recognized to be identical with it. Hence the uniform motion is described by a real position variable only when λ can be neglected as being negligibly small.

[328] Born recalled later that at M.I.T. he was kept very busy with his lectures and was involved in lots of discussions with other people, while Wiener spent much time in working out the results. Said Born: 'He showed it to me and I made my remarks and suggestions' (Born, Conversations, p. 16).

[329] In spite of Wiener's personal explanation, quoted above, it seems quite remarkable that Wiener, who had had the opportunity of contributing so early to the mathematical parts of quantum mechanics, and whose later work on harmonic analysis was closely connected with the foundation of the theory, left the field entirely to John von Neumann. The latter began to contribute to quantum mechanics only in spring 1927.

association with Born.[330] Born, who was just then treating another problem of aperiodic motion, the collision of atomic particles, though now with the methods of wave mechanics, again received Wiener's assistance. He acknowledged it in a footnote in the extended paper on the statistical interpretation of Schrödinger's wave function in the following words: 'In the mathematical realization of this work I have been assisted most kindly by Professor N. Wiener of Cambridge, Mass.; I wish to thank him here and acknowledge that I would not have arrived at the goal without him' (Born, 1926b, p. 805). This footnote of July 1926 described the actual role of Norbert Wiener in the formulation of quantum mechanics more adequately than the complaints which Born began to harbour against him later. At a time when just a few physicists struggled to develop a consistent theory of quantum mechanics, the Born–Wiener collaboration not only indicated the way for handling the problem of aperiodic motion but also contributed to the physical interpretation of the theory. Wiener himself had no reason to look back with regret on his work with Born, and he continued to return to the problems of quantum mechanics whenever his kind of mathematics seemed to offer an interesting and original application.[331]

[330]On Richard Courant's suggestion, made in summer 1925, Wiener had applied for a fellowship of the John Simon Guggenheim Memorial Foundation. Upon receiving the award of the fellowship for one year, he had planned to spend it entirely in Göttingen, where he expected a warm welcome. However, when he arrived, he found that Courant was 'somewhat hostile' towards him. In his autobiography, Wiener explained this change in Courant's behaviour by the fact that, before leaving the United States, he had given an interview to an American newspaper, and Courant had learned about its contents as well as Leo Wiener's anti-German activities during World War I. As Wiener recalled; 'Courant scolded me for my newspaper publicity and was disposed to deny me the assistant and the complete official recognition he had promised to me . . . . Nevertheless he allowed me to continue in Göttingen in an unofficial way, and after some persuasion found an able young mathematician to aid in the preparation of my lectures, provided I should pay him out of my own pocket' (Wiener, 1956, p. 114).

In addition to the difficulties with Courant, Wiener's lectures were not successful at all. Born recalled: 'When the students complained about his teaching, there was trouble between him and Courant, which I tried to smooth out' (Born, 1978, p. 227). After the summer semester Wiener and his wife (whom he had married just before leaving the United States) left Göttingen for an extended honeymoon trip to Italy. They did not return to Göttingen; Wiener decided to spend the remaining months of his fellowship with Harald Bohr in Copenhagen. He received a warm welcome from both Harald and Niels Bohr.

[331]Wiener kept up his interest in quantum mechanics for many years. For example, after his return to M.I.T. in 1927 he collaborated with a new colleague, the Dutch mathematician Dirk Jan Struik, applying the methods which the latter had used in differential geometry to differential equations, especially the Schrödinger equation (see, e.g., Struik and Wiener, 1927, and further papers). Wiener continued to publish papers on quantum theory occasionally, but then he stopped suddenly for more than twenty years. In 1953 he returned to the subject and tried to formulate, together with his student Armand Siegel, a new approach to the statistical interpretation of quantum mechanics, based on the mathematical concept of 'differential space' which he had used in the early 1920s to describe Brownian motion (Wiener and Siegel, 1953, and later papers).

Wiener spent many years of his career to extending harmonic analysis. He had numerous students and collaborators. For example, in 1932 he worked with the English mathematician R. E. A. C. Paley, and they wrote a book on complex Fourier transforms (Paley and Wiener, 1934). Wiener had another successful collaboration with Eberhard Hopf on certain specific integral equations, the so-called Wiener–Hopf equations. The latter played, together with harmonic analysis, an important role in the development of communication theory which occupied Wiener during the 1940s. This work not only

V.4 Multiply Periodic Systems: Action-Angle Variables and the Method of Complex Integration

Born and Wiener's paper on operator mechanics had an immediate impact on Heisenberg who saw it about the same time as it reached *Zeitschrift für Physik* on 5 January 1926.[332] Since November 1925 Heisenberg and Pauli had been trying unsuccessfully to calculate the energy states of the hydrogen atom, taking into account *relativistic* effects and perhaps the possibility that the electron possessed, as proposed by George Uhlenbeck and Samuel Goudsmit, an intrinsic magnetic moment. In order to achieve this goal, one had to compute the averages of certain negative powers of the radius variable of the electron, and this could be done only by introducing an angle variable ϕ, as in classical theory. Pauli, in his matrix theory of the nonrelativistic hydrogen atom, had made use of the 'Lenz vector' just because it had not been possible to express an angle coordinate as a Hermitian matrix. Now Born and Wiener's work suggested to Heisenberg that it might be possible to treat the angle, which is an aperiodic position variable, as an operator. He was very happy about it and he proposed immediately, in a letter written to Pauli on 7 January 1926, to treat the two-dimensional hydrogen atom in terms of polar coordinates, the radius r and the angle ϕ. While r could be considered, in its quantum-theoretical reformulation, as a genuine matrix, like any other dynamical variable which had so far been used in matrix mechanics, the angle ϕ was certainly not a Hermitian matrix, although it could be defined, according to Eq. (212), as a function of the matrices representing the Cartesian coordinates of the electron (which, in classical theory, moves in a plane around the nucleus). Furthermore, Heisenberg defined the azimuthal momentum p_ϕ, the canonically conjugate variable to the angle ϕ, by the equation

$$p_\phi = xp_y - yp_x \,. \tag{252}$$

The variable p_ϕ could, of course, be represented by a Hermitian matrix like the variables x, p_x, y and p_y. But since he was dealing with a formalism, in which the variables (including the angle ϕ) were represented by generalized matrices—

---

influenced the practical construction of computers, it also led Wiener to create a new mathematical field, 'cybernetics,' dealing with the problems of communication and control in artificial machines as well as in living organisms.

Wiener's accomplishments received increasing recognition. In 1929 he was invited to lecture on his researches at Brown University. He spent the academic year 1931–1932 in Cambridge, England, where his former teacher G. H. Hardy asked him to give a course on Fourier integrals. During 1935–1936 he visited China, lecturing at the Tsing-Hua University in Peking. In 1951 he served as Fulbright Lecturer at the *Collège de France* in Paris and spent the academic year 1955–1956 as a guest professor at the Indian Statistical Institute in Calcutta. In 1933 Wiener received the Bôcher prize for analysis. Shortly before his death President Lyndon Baines Johnson awarded him the National Medal of Science. Norbert Wiener died in Stockholm on 18 March 1964.

[332] Heisenberg mentioned the 'recent work of Born and Wiener' ('*neuere Arbeit von Born und Wiener*') in his letter to Pauli, dated 7 January 1926, which he wrote from Göttingen shortly after his return from Christmas vacation in Munich.

notably the operators of the type which Born and Wiener had employed in their paper—Heisenberg dropped the matrix notation completely.[333] From the matrix commutation relations for the variables, x and p_x, y and p_y, respectively, and using the formal expansion of the function defining ϕ in Eq. (212), Heisenberg obtained the commutation relation

$$p_\phi \phi - \phi p_\phi = \frac{h}{2\pi i} , \tag{253}$$

for ϕ and p_ϕ. Equation (253), which had to be considered as an operator equation (including an operator 1 on the right-hand side), confirmed the correctness of the definition of p_ϕ through Eq. (252).

With the help of these results, Heisenberg was able to proceed with the calculation of the average values of the negative powers of r $(=\sqrt{x^2 + y^2})$, the radius variable of the electron (for motions in a plane). He noted that p_ϕ could be replaced, as in classical theory, by the expression $m_e r^2 \dot{\phi}$, where m_e is the mass of the electron. Assuming that $\dot{\phi}$, the time derivative of the angle, depends (again as in classical theory) only on the principal quantum number of the atom, he concluded that $\overline{r^{-2}}$, the average over a full period of the inverse square radius, depended on the orbital momentum quantum number k in the same way as $(1/p_\phi)$. Since the values of p_ϕ were known from the spectroscopic data, the quantity $\overline{r^{-2}}$, which entered into the calculation of the relativistic correction to the hydrogen terms, could be evaluated. Heisenberg was very encouraged by this result and continued to work along the same lines. On 14 January 1926 he wrote again to Pauli, informing him about the progress he had made during the previous week. Heisenberg had accepted to give a talk on 19 January before the *Kegelklub* in Stade, close to Hamburg, and he planned to pass through Hamburg the following day before returning to Göttingen.[334] 'I hope,' he wrote in anticipation of his intended meeting with Pauli, (on 20 January), 'to spend many hours discussing physics with you' (Heisenberg to Pauli, 14 January 1926). One week after this meeting Heisenberg presented, in another letter to Pauli, a final report on what he had learned about 'cyclical variables' ('*das Kapitel "zyklische Variable"*') (Heisenberg to Pauli, 27 January 1926).

In his letter Heisenberg outlined the treatment of the problem of periodic motion with one degree of freedom, which he formulated in terms of J, the action variable, and w, the angle variable, satisfying the commutation relation

$$Jw - wJ = \frac{h}{2\pi i} . \tag{254}$$

[333] Born and Wiener used their quantum-mechanical operators to describe dynamical variables, employing the same symbols as were used in classical theory. By applying the operator methods the difference between the quantum-theoretical and the classical variables was so obvious that it did not have to be emphasized by a different notation. We shall follow the same procedure in the following sections and use the matrix notation only when absolutely necessary.

[334] See the letters from Heisenberg to Pauli, 24 December 1925 and 7 January 1926.

Again Heisenberg considered the variables J and w in this equation as operators in the sense of Born and Wiener. If it is assumed that H, the Hamiltonian operator, depends only on the action variable J, then due to the equation of motion (i.e., $\dot{w} = \partial H/\partial J$) the time derivative of the angle w does not depend on time either. Thus, Heisenberg could write down the expression

$$w = t \cdot \omega(D) + w_1(D), \tag{255}$$

for w, where the operator $\omega(D)$ denotes the frequency of the atomic system (multiplied by 2π) and $w_1(D)$, the angle at time $t = 0$, both quantities being independent of t. The Cartesian coordinates, x and y, representing the motion, as well as their conjugate momenta, p_x and p_y, are periodic functions in time, being proportional to the exponential function $\exp(iw\tau)$, where $\tau\ (= \pm 1, \pm 2, \ldots)$ is a positive or negative integer. The action variable J then assumes only the discrete values, $J_n = \text{const.} + n(h/2\pi)$, n being a positive integer. By considering the time derivatives of the periodic variables, Heisenberg obtained the relation

$$\frac{d}{dt}\left[\exp(iw\tau)\right] = \frac{2\pi i}{h}\left[H(J) - H\left(J - \tau\frac{h}{2\pi}\right)\right] \cdot \exp(iw\tau), \tag{256}$$

which includes, provided (τw) is interpreted as the transition frequency from the stationary state J to the state $J - \tau h/2\pi$, Bohr's frequency condition.[335]

Equation (256) and the fact that the action variable J assumes discrete values, separated by integral multiples of $(h/2\pi)$, showed that an operator scheme of action-angle variables, as proposed by Heisenberg, would indeed yield the desired quantum-theoretical results for periodic systems. The important question still remained whether it was justified to regard the Hamiltonian operator as a function of the action variable only, that is, $H = H(J)$. If the answer was affirmative, and Heisenberg had little doubt that it was, he could prove his earlier assumption concerning the average value of $\dot{\phi}$ (namely, that it depended on the principal quantum number alone). Then he would be permitted to calculate the average values of r^{-2} and r^{-3} for the hydrogen atom as he had indicated in his letter to Pauli on 7 January.

[335] To prove Eq. (256), Heisenberg started from the expression

$$\frac{d}{dt}\exp(iw\tau) - \exp(iw\tau)\frac{d}{dt},$$

which according to Eq. (237), the equation of motion in operator mechanics, gives the time derivative of the variable $\exp(iw\tau)$ (taking into account the operator identity, Eq. (239)). Then he expanded $\exp(iw\tau)$ in powers of w, and differentiated the powers of w with respect to time. Thus, he finally obtained for the above expression the result

$$\left[i\tau\frac{\partial H}{\partial J} + \frac{h}{2\pi i}\frac{\tau^2}{2}\frac{\partial^2 H}{\partial J^2} - i\left(\frac{h}{2\pi i}\right)^2\frac{\tau^3}{3!}\frac{\partial^3 H}{\partial J^3} + \ldots\right]\exp(iw\tau),$$

which could be reformulated as the right-hand side of Eq. (256).

In the meantime, Pauli, who shared Heisenberg's interest in developing a complete theory of the hydrogen atom (that is, a relativistic hydrogen atom with or without electron spin), 'had not remained idle' ('*Ich war inzwischen auch nicht müßig*,' Pauli to Heisenberg, 31 January 1926). He had himself pursued a systematic formalism for treating action-angle variables. Unlike Heisenberg, he did not employ the operators of Born and Wiener. 'The introduction of the "operators,"' he commented in a letter to Heisenberg, 'is at any rate superfluous for many purposes. They are related to the calculus with functions of "time" and energy—which I shall sketch in the following—just as the theory of principal-axes transformation [is related] to your matrix calculus. I am glad to have the operators behind me, and shall now happily go on with my calculations' (Pauli to Heisenberg, 31 January 1926).[336]

The principal idea involved in Pauli's formulation was to represent the dynamical variables in quantum mechanics by functions of time t and energy W; for instance, a position coordinate by $q(t, W)$. Argued Pauli: 'The only essential thing is that the quantities $q(t, W)$ must be considered as continuous functions of the energy, as soon as one deals with non-periodic quantities besides periodic quantities' (Pauli to Heisenberg, 31 January 1926). Although the functions $q(t, W)$ were formally identical with Born and Wiener's generalized column-sums (which they had denoted by the same letters), Pauli gave them a quite different meaning. In particular, he remarked: 'In general the "time" t has, however, no real physical significance, but only a formal one. For nonperiodic quantities $q(t, W)$ the connection with observable phenomena remains an open question; there is no harm in this, if finally everything reduces again to periodic quantities x, with the $x(t, W)$ representing the sums of columns of matrices' (Pauli to Heisenberg, 31 January 1926). Pauli emphasized the different, formal significance of the 'time' variable in his scheme by *postulating the fundamental commutation relation*

$$Wt - tW = \frac{h}{2\pi i} 1 \tag{257}$$

between t and the energy W. Pauli conceived Eq. (257) as a generalization of a matrix equation, as is evident from the use of the factor 1 (denoting something like a unit quantity) on the right-hand side. The energy variable, W, embraces all possible energy values of the system. If it could be represented by a matrix, W would have to be identified with $(W_m \delta_{mn})$, the matrix obtained by diagonalizing the Hamiltonian matrix of the system, and Eq. (257) would be a genuine matrix equation.[337]

[336] There is little doubt that Pauli thought that the operator formalism was again part of Born's formal erudition. He believed that he could develop a method serving the same purpose (namely, to deal with aperiodic variables), but 'simpler and better than that of Born and Wiener' (Pauli to Bohr, 5 February 1926).

[337] Equation (257) can be derived in Born and Wiener's operator mechanics from the commutation relation of t and D $(= d/dt)$ and the relation between the Hamiltonian operator H and D, Eq. (239).

The idea that the time variable t might enter into quantum mechanics in a more 'formal' manner than in classical mechanics occupied Pauli for some time. He mentioned it first in a letter to Niels

In contrast to Born and Wiener, who had embarked upon the formulation of a completely new mathematical scheme of quantum mechanics, Pauli wanted to stick as closely to matrix mechanics as the physical situation would allow, for he believed that the relations governing the dynamical variables, as represented by matrices, were well defined, and their interpretation in terms of empirical facts was clear. Thus, he proceeded to develop his formalism (or '*Kalkül*') first for his description of a system which had the discrete energy eigenvalues W_n, and whose position coordinate was represented by all possible matrix elements, $q(n, n-\tau) \cdot \exp[2\pi i\nu(n, n-\tau)t]$, where τ assumed positive and negative integral values. 'In order also to be able to calculate with non-periodic quantities (as, for instance, polar angles), we must pay the price that the discontinuous index n is replaced by as many continuous variables as the number of degrees of freedom which the system possesses' (Pauli to Heisenberg, 31 January 1926). Restricting himself to systems of one degree of freedom, he replaced the discrete quantities, $x(n, n-\tau)$ and $\nu(n, n-\tau)$, denoting respectively, the amplitudes and transition frequencies in the matrix scheme, by $x_\tau(W)$ and $\nu_\tau(W)$, i.e., by functions of the (discrete and continuous) energy W. Then he represented x, say the position variable of the system, by the sum

$$x(t, W) = \sum_{\tau=-\infty}^{+\infty} x_\tau(W)\, \underline{\exp}\left[2\pi\nu_\tau(W)t\right], \tag{258}$$

where the underlined exponential function had the special properties

$$\frac{d}{dt}\left\{\underline{\exp}\left[f(W)t\right]\right\} = f(W)\, \underline{\exp}\left[f(W)t\right] \tag{258a}$$

and

$$\frac{d}{dW}\left\{\underline{\exp}\left[f(W)t\right]\right\} = \frac{df(W)}{dW}\, \underline{\exp}\left[f(W)t\right] \cdot t. \tag{258b}$$

Pauli emphasized that 'only the x_τ(W) possess physical significance via their transition probabilities,' and he thought that t and $\underline{\exp}[2\pi i\nu_\tau(W)t]$ were 'for the

---

Bohr on 17 November 1925. There he suggested that the time t did not have the usual meaning in matrix mechanics, and that the frequencies in the exponential factors of the matrix elements were also rather 'formal' quantities; in order to find the *real* time variable, he proposed to *define* it with the help of energy rather than introduce it as an *a priori* given variable. Pauli came back to this idea in fall 1926 when he discussed with Heisenberg the interpretation of the quantum-mechanical scheme. Several years later, however, Pauli realized that the description of time in quantum mechanics by a Hermitian operator made no sense; the difficulty arose, especially, when one wanted to describe a system having both discrete and continuous values of the energy W (see Pauli, 1933c, p. 140, footnote 1).

Finally it should be mentioned that Cornelius Lanczos also arrived independently at the observation that t, the time variable of the matrix scheme, had a formal character. In his first paper on quantum mechanics Lanczos stated, in particular, that the time derivative of Born and Jordan in matrix mechanics did not refer to real motion, but rather served as the definition of a special property of the Hamiltonian matrices describing dynamical variables; he, therefore, wrote the time derivative in quotation marks (Lanczos, 1926a, p. 815).

moment symbols without physical significance' which just served to permit the multiplication of two quantities, $x(t, W)$ and $y(t, W)$. Making use of the commutation relation between W and t, Eq. (257), he derived the multiplication law

$$xy(t, W) = \sum_{\tau} (xy)_{\tau}(W)\, \underline{\exp}\left[2\pi i \nu_{\tau}(W)t\right], \tag{259}$$

where the coefficient of the exponential factor was given by the equation

$$(xy)_{\tau}(W) = \sum_{\sigma} x_{\sigma}(W) y_{\tau-\sigma}(W - \nu_{\sigma}(W)h). \tag{259a}$$

Equation (259a) agreed perfectly with the multiplication law for the matrix amplitudes. In addition, the relation

$$\nu_{\sigma+\tau}(W) = \nu_{\sigma}(W) + \nu_{\tau}(W - \nu_{\sigma}(W)h), \tag{259b}$$

representing the frequency combination law, was valid.

Having thus shown that his formalism, based on the functions $x(t, W)$ provided a suitable extension of the matrix description of periodic variables, Pauli turned to the problem of introducing action-angle variables for periodic systems of one degree of freedom. Assuming that the energy of the system, following a canonical transformation of the variables, could be written as a function of J, the action variable, alone, that is, $W = W(J)$, he sought to represent the angle variable w by the equation

$$w = \omega(J) \cdot t + w_1(J), \tag{260}$$

where the factor $\omega(J)$—which, apart from a factor 2π, is a frequency—is determined by the equation of motion

$$\omega(J) = \dot{w} = \frac{\partial W(J)}{\partial J}\,. \tag{261}$$

Equation (260) was completely equivalent to Heisenberg's operator equation (255). Pauli also proved Heisenberg's commutation relation between J and w, Eq. (254), with the help of Eq. (257), his commutation relation between t and W. Finally, he concluded that $\nu_{\tau}(W)$, the frequencies occurring in Eq. (258), which in the case of a periodic system of one degree of freedom, depended on the action variable alone—that is, $\nu_{\tau}(W) = \nu_{\tau}(J)$—, were the quantum-theoretical transition frequencies.

Pauli's formalism for treating action-angle variables was not yet complete, for he did not have a method that would allow him to transform the Hamiltonian of a given periodic system into a function $W(J)$. However, it was important for him that 'now the non-periodic coordinates have been stripped of their horror' (Pauli to Heisenberg, 31 January 1926). And he was anxious to apply the new methods to the hydrogen atom to see whether they would work in this case. For the actual

calculation, in particular to obtain the energy eigenvalues as functions of the action variable J, Pauli found it preferable to extend the perturbation scheme of matrix mechanics. Before, however, he was ready to publish any of his results, he came across a paper of Paul Dirac's dealing with the same problems (Dirac, 1926a). Pauli was disappointed. He had worked hard and fast enough to obtain an elegant formalism for treating action-angle variables, but Dirac had beaten him in the race. Pauli complained to Kramers: 'In the March issue of the *Proceedings of the Royal Society* there is a very beautiful paper of Dirac [Dirac, 1926a], which contains all the results that I have found since Christmas concerning the generalization of the matrix calculus to non-periodic quantities (e.g., polar angles). I regret that in this way I have lost so much time; I could well have done something else meanwhile' (Pauli to Kramers, 8 March 1926). Though Pauli felt that Dirac's paper had rendered most of his efforts on action-angle variables superfluous, he found that 'Dirac's calculations on the H-atom are, however, unnecessarily complicated. It can be done much more simply according to the methods of Wentzel and myself' (Pauli to Kramers, 8 March 1926).

Gregor Wentzel, to whose work Pauli referred to Kramers, spent the winter semester 1925–1926 at the University of Hamburg. He had been a member of Sommerfeld's Institute since 1920, staying on there after obtaining his doctorate in 1921, and had become a *Privatdozent* in 1922. He had worked mainly on the X-ray spectra of atoms, which had also been the subject of his thesis (Wentzel, 1921a), and he had become an expert in this aspect of atomic theory as well as in the theory of spectra of complex atoms.[338] For instance, in a paper submitted to *Annalen der Physik* in February 1925, Wentzel had thoroughly analyzed the difficulties arising from the current theoretical explanation of doublet spectra (Wentzel, 1925b), and his former teacher, Arnold Sommerfeld, had not hesitated to call this account 'the best that has been written on the *doublet-mystery* of optical and X-ray spectra' (Sommerfeld to Pauli, 18 June 1925). And Pauli, who, having just expounded his exclusion principle, was struggling with the same problem at the same time, had agreed with Sommerfeld (Pauli to Sommerfeld, 22 June 1925). In fall 1925, when Wentzel arrived in Hamburg, Pauli had completed the calculation of the energy states of the hydrogen atom with the help of matrix methods; he was very excited about the Göttingen scheme of quantum mechanics and hoped to exploit the possibility of extending it to deal with more complex atoms. The latter had been treated earlier by employing action-angle variables in the theory of multiply periodic systems. However, the difficulty of being able to represent angles in matrix theory seemed to prevent all progress. Wentzel was a witness when, after certain not very promising attempts of Heisenberg and Pauli, the important new hint as to how to proceed in the case of aperiodic position variables was provided by the operator methods of Born and Wiener and, as a result, Heisenberg and Pauli were able to envisage an action-angle formalism in quantum mechanics. The programme of developing a unified scheme, in which

[338] Gregor Wentzel's background and his contributions to spectroscopy have been mentioned in Volume 1, Section III.4, and footnote 609.

both periodic and aperiodic motions could be treated, reminded Wentzel of an earlier work of his own on the same problem, a work in which he now discovered certain parallels to the new ideas of Born and Wiener and of Heisenberg and Pauli.

In a paper submitted to *Zeitschrift für Physik* in July 1924, Wentzel had attempted to calculate the properties of X-ray bremsstrahlung spectra, that is, the radiation emitted by electrons moving in hyperbolic orbits around the atomic core or nucleus in accordance with the Bohr–Sommerfeld theory of atomic structure (Wentzel, 1924d). In classical theory, the radiation from an accelerated atom could be obtained from the kinematics of the electron. Wentzel had proposed to take over the same formula into quantum theory, with the only change that the acceleration $\dot{\mathbf{v}}$ was replaced by a vector quantity $\mathbf{q}$.[339] He had found this quantity in the following way (see Wentzel, 1924d, appendix). For a system of f degrees of freedom a 'characteristic phase,' ϕ, is introduced as

$$\phi = -2\pi\left(\nu t + \sum_{k=1}^{f} j_k u_k\right). \tag{262}$$

In Eq. (262), ν is the frequency of the quantum transitions from the initial state with energy W_i to the final state with energy W_f; the u_k, $k = 1, \ldots, f$, denote the averages of the position coordinates of the system between the same states, and the j_k represent the changes in the canonically conjugate variables I_k. Then $\mathbf{q}$, the acceleration of the electron in quantum theory, is obtained from the equation

$$\mathbf{q} = \int \ldots \int \dot{\mathbf{v}}(u_k) \exp(-i\phi)\, dj_1 \ldots dj_f, du_1 \ldots du_f, \tag{263}$$

involving the phase ϕ, Eq. (262), and the classical acceleration $\dot{\mathbf{v}}$, taken as a function of the position coordinate u_k, $k = 1, \ldots, f$. All the $2f$ integrals in Eq. (263) extend from $-\infty$ to $+\infty$; that is, the right-hand side is a certain phase-space average of the classical acceleration vector $\dot{\mathbf{v}}$. Because of $\exp(-i\phi)$, the exponential weight function, $\mathbf{q}$ becomes identical with the values of $\dot{\mathbf{v}}$ at those coordinates u_k which satisfy the condition

$$\sum_{k=1}^{f} j_k u_k = -2\pi\nu t. \tag{264}$$

It can be seen that the quantities j_k determine the deviation from classical theory in the following sense: When the changes in the energy W and all j_k are small, $\mathbf{q}$ coincides with $\dot{\mathbf{v}}$, and the classical limit is achieved.[340] Hence for the

[339] The classical equation expressing the electric field strength $\mathbf{E}$ at the point $\mathbf{r}$ (where r is the distance from the electron's position to the point of observation) is $\mathbf{E} = (e/c^2 r)[\hat{\mathbf{n}} \times (\hat{\mathbf{n}} \times \mathbf{v})]$, where $\hat{\mathbf{n}}$ is the unit vector in direction of radiation.

[340] In this limit the integral in Eq. (263) yields, for arbitrary u_k, the vector function $\dot{\mathbf{v}}(u_k)$ because condition (264) is automatically satisfied. The reason is that ν may be replaced by $-(2\pi)^{-1}\sum_{k=1}^{f} \partial W/\partial I_k$; hence the right-hand side of Eq. (264) becomes identical with the left-hand side on account of the equations of motion ($\dot{u}_k = \partial W/\partial I_k$, $k = 1, \ldots, f$) and the fact that the position variables u_k are linear functions of t.

calculation of the soft (that is, long-wavelength) part of the bremsstrahlung spectra the classical equation may be used. For large values of j_k, however, Wentzel had found that this equation could not describe the observed data, but the agreement between theory and experiment was improved when $\dot{\mathbf{v}}$ was replaced by $\mathbf{q}$, Eq. (263). On the other hand, by specializing to strictly periodic systems of f degrees of freedom, Wentzel's I_k became the action variables, J_k, each of which was quantized; hence the j_k assumed integral values given by $\tau_k = 0, \pm 1, \pm 2, \ldots,$ and the position variables, u_k, were to be identified with the angle variables w_k. The energy, W, now depended only on the J_k, and conditions (264) now became

$$\sum_{k=1}^{f} \tau_k w_k = \frac{1}{h}\left[W(J_1, \ldots, J_f) - W\left(J_1 - \frac{\tau_1 h}{2\pi}, J_2 - \frac{\tau_2 h}{2\pi}, \ldots, J_f - \frac{\tau_f h}{2\pi}\right)\right] t. \tag{264'}$$

With the help of Eq. (264′), Wentzel's integral, Eq. (263), decomposed into an f-fold Fourier sum for $\mathbf{q}$, which could be written as

$$\mathbf{q} = \sum_{\tau_k} \dot{\mathbf{v}}(J_j) \exp(2\pi i \tau_k w_k), \tag{263'}$$

where the τ_k, $k = 1, \ldots, f$, ran to all integral values and the quantum frequencies occurred explicitly in the exponentials because of Eq. (264′).[341]

Though Eq. (263′) had been derived on the basis of correspondence arguments long before matrix mechanics and operator mechanics made their appearance, it was entirely compatible with the spirit of quantum mechanics. It represented nothing less than a generalization of Pauli's expression (258) to nonperiodic systems of f degrees of freedom. Because of this previous experience with aperiodic motions and his considerable skill in formal, complicated calculations, Wentzel was well prepared to enter into discussions with Pauli on action-angle variables in the beginning of 1926. He quickly mastered the new ideas of matrix and operator mechanics and began to build on Pauli's results on angle variables. He directed his attention primarily to those aspects of the theory which Heisenberg and Pauli had left out of their earlier discussions.[342] He developed a general theory of *multiply periodic* systems and answered, in particular, the question as to how to obtain the energy of these systems as a function depending on the action

[341] To calculate the bremsstrahlung intensities, Wentzel had studied the motion of an electron in an elliptical orbit, described by action and angle variables, J and w. He had then taken over the equations expressing J and w in terms of the energy W also for the hyperbolic motion, except that W was always replaced by $-W$. With these results he had evaluated Eq. (263).

Note that Wentzel used as angle variables the angles divided by 2π and action variables multiplied by the same factor.

[342] Since Pauli and Heisenberg had developed their action-angle formalism towards the end of January 1926, and since Pauli was always concerned that nobody should interfere with Heisenberg's plans for quantum mechanics, Wentzel did not start his work on multiply periodic systems before early in February.

variable alone. The results pleased Pauli very much, and he wrote to Bohr on 26 February 1926: 'By the way, Wentzel has performed a very beautiful investigation on the new quantum mechanics.' A month later *Zeitschrift für Physik* received Wentzel's paper, entitled '*Die mehrfach periodischen Systeme in der Quantenmechanik*' ('The Multiply Periodic Systems in Quantum Mechanics'), which was published in its issue of 22 May 1926 (Wentzel, 1926a).

The fundamental equations constituting the quantum theory of multiply periodic systems, described by f angle variables w_k and their canonically conjugate momenta, the f action variables J_k, $k = 1, \ldots, f$, are provided by the equations of motion

$$\dot{w}_k = \frac{2\pi i}{h}(Ww_k - w_k W) \tag{264a}$$

and

$$\dot{J}_k = \frac{2\pi i}{h}(WJ_k - J_k W), \tag{264b}$$

involving the energy W, and by the commutation relations between the dynamical variables, that is,

$$J_k w_l - w_l J_k = \frac{h}{2\pi i}\delta_{kl} \qquad \text{where } \delta_{kl} = 1 \text{ for } k = l, \text{ and } 0 \text{ otherwise,}$$

$$w_k w_l - w_l w_k = J_k J_l - J_l J_k = 0, \qquad k = 1, \ldots, f. \tag{265}$$

Provided the energy depends only on the action variables, J_l, $l = 1, \ldots, f$, every periodic variable—say, the Cartesian coordinates of the system or their conjugate momenta—can be expressed by means of an f-fold Fourier sum,

$$q = \sum_{\tau_k} q_{\tau_k}(J_l)\exp\left[2\pi i \sum_k \tau_k w_k\right], \tag{266}$$

where the subscript τ_k in the sum denotes the set of numbers $(\tau_1, \ldots, \tau_f)$. In the special case that q is the acceleration vector $\mathbf{q}$, Eq. (266) becomes identical with Wentzel's equation (263) of summer 1924. At the same time Eq. (266) generalizes Pauli's *Ansatz*, Eq. (258), to systems of several degrees of freedom; that is, Pauli's functions of energy and time have to be identified with Wentzel's functions of action and angle variables. With the help of the commutation relations, Eqs. (265), Wentzel derived the result that by taking the time derivative of q one reproduces a Fourier sum whose terms are identical with the terms on the right-hand side of Eq. (266)—i.e., the Fourier terms of q—multiplied by the factor $2\pi i \nu_{\tau_k} t$, where ν_{τ_k} are the quantum frequencies satisfying the equations

$$\nu_{\tau_k} = \frac{1}{h}\left[W(J_k) - W\left(J_k - \frac{\tau_k h}{2\pi}\right)\right], \qquad \tau_k = 0, \pm 1, \pm 2, \ldots . \tag{267}$$

Again Eq. (267) generalizes Pauli's result for periodic systems of one degree of freedom.[343] From it Wentzel concluded: 'If one requires, in addition, that the J's form a "closed set," it follows at once that they assume discrete values, separated from each other by integral multiples of $h/2\pi$' (Wentzel, 1926a, p. 82).[344] Finally, again under the assumption that the energy is a function of the action variables alone, i.e., $W = W(J_k)$, and with the help of the commutation relations, Eqs. (265), Wentzel showed that the variables q obey Heisenberg's multiplication rule for quantum variables.[345]

One question still remained open in this scheme of action-angle variables: Could general methods be established that would allow for the determination of energy W as a function of the action variables for a given multiply periodic system described by the variables q_k and p_k, $k = 1, \ldots, f$? The same method would then yield directly, via Eq. (267), the transition frequencies of the system. This question had not been answered thus far in quantum mechanics, not even by Heisenberg and Pauli. Hence Wentzel noted in the introduction of his paper: 'In Heisenberg's quantum mechanics [as also in the operator mechanics of Born and Wiener] there still is not available a general procedure for calculating the frequencies of spectral lines independently of their intensities' (Wentzel, 1926a, p. 80). However, after the introduction of action and angle variables the quantum-mechanical theory of multiply periodic systems showed a great similarity with the previous Bohr–Sommerfeld theory of atomic systems. It now occurred to Wentzel that: 'Since only the frequencies were determined in the previous quantum theory, one should ask the question: could somehow the old methods not be translated in the language of matrices so that the terms [i.e., energies] can be determined directly in the new theory as well?' (Wentzel, 1926a, p. 80).

In the classical theory of multiply periodic systems the problem of obtaining the energy, say of a system of f (periodic) degrees of freedom, as a function of the

[343] In order to prove the results, Wentzel started from the equation of motion for the variable q, i.e., $\dot{q} = (2\pi i/h)(Wq - qW)$, and took into account that $W = W(J_l)$. Then he proved the validity of the equation $\exp[2\pi i\tau_k w_k]W(J_1, \ldots, J_f) = W(J_1, \ldots, J_{k-1}, J_k - \tau_k h/2\pi, J_{k+1}, \ldots, J_f)$. By substituting it into the right-hand side of the equation of motion, he obtained the result $Wq - qW = (2\pi i/h)[W(J_k) - W(J_k - \tau_k h/2\pi)]q$. Since the left-hand side may be identified with $2\pi i\nu_{\tau_k}$, Eq. (267) follows.

[344] The postulate that all values, which an action variable J_k can assume, form a 'closed set' was, of course, a consequence of Born and Wiener's postulate stating that the set of functions, $\{x(t)\}$, associated with quantum-mechanical operators, was closed. This postulate then allowed one to determine the values of J_k up to a constant, i.e., $J_k = \tau_k h/2\pi + \text{const.}$, where the constant had to be determined from the properties of the system under investigation. (For instance, it could be calculated from the boundary conditions imposed on the system.)

Note that Wentzel, in his paper, took action variables which were 2π-times our variables J_k; thus, his discrete J_k were separated by integral multiples of h rather than of $h/2\pi$. As a result, Wentzel's angular variables, w_k, were 2π-times smaller than the ones used above. We have not followed his choice in order to keep the notation of action-angle variables uniform throughout this chapter.

[345] For example, the square of a variable q is represented by the f-fold Fourier sum, $q^2 = \sum_{\tau_k} q^2_{\tau_k}(J_l)\exp[2\pi i\sum\tau_k w_k]$, where the coefficient $q^2_{\tau_k}(J_l)$ is given by the expression

$$q^2_{\tau_k}(J_l) = \sum_{\sigma_k} q_{\sigma_k}(J_l)\, q_{\tau_k - \sigma_k}\left(J_1 - \frac{\sigma_1 h}{2\pi}, J_2 - \frac{\sigma_2 h}{2\pi}, \ldots, J_f - \frac{\sigma_f h}{2\pi}\right).$$

action variable J_k alone, is solved with the help of the Hamiltonian–Jacobi partial differential equation for the action function S. (See Eq. (129) for systems having one degree of freedom.) This function depends on the original position variables q_k and the transformed momentum variables, J_k, $k = 1, \ldots, f$, and may be defined by the equation

$$dS = \sum_{k=1}^{f} p_k dq_k + \sum_{k=1}^{f} w_k dJ_k , \tag{268}$$

expressing the total differential of S by sums of terms, $p_k\,dq_k$ and $w_k\,dJ_k$, where the p_k are considered to be functions of the q_k and J_k. The action function S provides a canonical transformation from the original set of dynamical variables, p_k and q_k, to the action-angle variables, J_k and w_k. Expressed in terms of the action-angle variables, the Hamiltonian will depend only on the J_k, which are constants of motion; hence the energy of the system will be given in terms of the values of the action variables. Thus, the energy of a classical system can be obtained, and it was this procedure which Wentzel wanted to reformulate in quantum theory.

In quantum theory all canonical transformations from one complete set of dynamical variables of a given system into another complete set—according to the investigations of Born, Heisenberg and Jordan (1926, Chapter 1, Section 4)—were unitary transformations. Since the angle variables did not admit of a matrix representation, Wentzel reformulated the classical action function not as a unitary matrix but as a more general object, say of the type of a unitary operator in the language of Born and Wiener (i.e., an operator which when multiplied by its Hermitian conjugate yields the unit operator), or of Pauli's functions of energy and time, or as the q-number objects of Paul Dirac.[346] By referring to the analogy argument, according to which the differential expressions of classical mechanics are always replaced by the differences of products in quantum mechanics, Wentzel suggested that the quantum-mechanical action function satisfies the following two sets of equations:

$$SJ_k - J_kS = \frac{1}{2}\sum_{l=1}^{f}\left[p_l(q_lJ_k - J_kq_l) + (q_lJ_k - J_kq_l)p_l\right] \tag{269a}$$

and

$$Sq_k - q_kS = \frac{1}{2}\sum_{l=1}^{f}\left[w_l(J_lq_k - q_kJ_l) + (J_lq_k - q_kJ_l)w_l\right], \tag{269b}$$

[346] Wentzel, in his published paper (Wentzel, 1926a), did not at all refer to the article of Born and Wiener (1926) from which Pauli's considerations on action-angle variable had taken their point of departure. Rather, he quoted the work of Dirac (1926a), which dealt with action-angle variables in the q-number scheme. Dirac's methods were indeed closer to Wentzel's own thinking than the operator methods of Born and Wiener. It should be emphasized, however, that Wentzel had obtained his results concerning the calculation of the energy states of multiply periodic systems already at the end of February (see Pauli to Bohr, 26 February 1926), i.e., before he had seen Dirac's paper.

where $k = 1, \ldots, f$. Equation (269a) helps to determine the dependence of the quantum-mechanical transformation function S on the variables q_k and J_k; Equation (269b) could be used to define the angle variable w_k.[347] With the help of Eqs. (269a) Wentzel calculated ΔS, the change which the action function S undergoes when the periodic variables go through a full period in a given angle, say w_n. Since w_n changes by 2π, while all other angles (i.e., w_k with $k \neq n$) are kept constant, and all periodic variables q_k, $k = 1, \ldots, f$, remain unaltered, Wentzel obtained the result

$$\Delta S = 2\pi J_n . \tag{270}$$

Hence, just as in classical theory, the action variables could be identified, up to a factor 2π, with the 'periodicity moduli of the action function' (Wentzel, 1926a, p. 83).[348]

These results enabled Wentzel to reformulate in quantum mechanics the methods which Sommerfeld had used many years previously to calculate the energy states of multiply periodic atomic systems subject to quantum conditions (Sommerfeld, 1916d). These systems had the property of being separable, that is, each of their momentum variables, p_k, depended only on the conjugate position coordinates, q_k. As a result S, the action function of a separable system of f degrees of freedom, decomposed into a sum of individual terms S_k, each of which was given—in agreement with Eq. (268)—by an integral $\oint p_k \, dq_k$ over a closed path of the variable q_k. Then ΔS_k, the 'periodicity modulus' of the specific action function S_k, was obtained from this 'phase-integral.' In Sommerfeld's theory a special role was played by the phase-integral ΔS_r, that is,

$$\Delta S_r = \oint p_r \, dr, \tag{271}$$

involving p_r, the radial momentum of an atomic system, and an integration over a closed path in the variable r, the radial coordinate of the system, from $r = r_{\min}$ to $r = r_{\max}$ and back to $r_{\min}$ during a full period of the system. In the separation

[347] Equation (269a) may be considered as the quantum-theoretical reformulation of Eq. (268), while Eq. (269b) corresponds to the classical equation, $w_k = \partial S / \partial J_k$, $k = 1, \ldots, f$. One should remember that by Eq. (268) the classical action function is determined only up to a constant additive term which might be a function of the action variables. This indeterminacy leads to the result that the angle variables are determined only up to constant phases (i.e., $w_k = \omega_k . t + \text{const.}$, $k = 1, \ldots, f$) which represent the positions of the electrons at time $t = 0$. Now an analogous indeterminacy occurs in quantum mechanics, where Eq. (269a) fixes the function S only up to a constant. This fact, however, should not be taken too seriously because in atomic systems the positions of the electrons at time $t = 0$ cannot be observed. Nevertheless, Wentzel proposed to make the definition of S unique by adding a further set of equations (see Wentzel, 1926a, p. 83, Eqs. (6)).

[348] Equation (270) follows from Eq. (269b) and the equation $(S + \Delta S)q_k - q_k(S + \Delta S) = \frac{1}{2}[(w_n + 2\pi)(J_n q_k - q_k J_n) + (J_n q_k - q_k J_n)(w_n + 2\pi)] + \ldots$, where the dots represent terms multiplied by the unchanged angle variables. By subtracting the above equation from Eq. (269b), Wentzel found the result $\Delta S q_k - q_k \Delta S = 2\pi(J_n q_k - q_k J_n)$ from which he concluded Eq. (270).

In order to complete the introduction of the quantum-mechanical action function, Wentzel had to prove that the variables J_k and w_k, entering in Eqs. (269a) and (269b), satisfied the commutation relations, Eqs. (265). (See Wentzel, 1926a, pp. 83–84.)

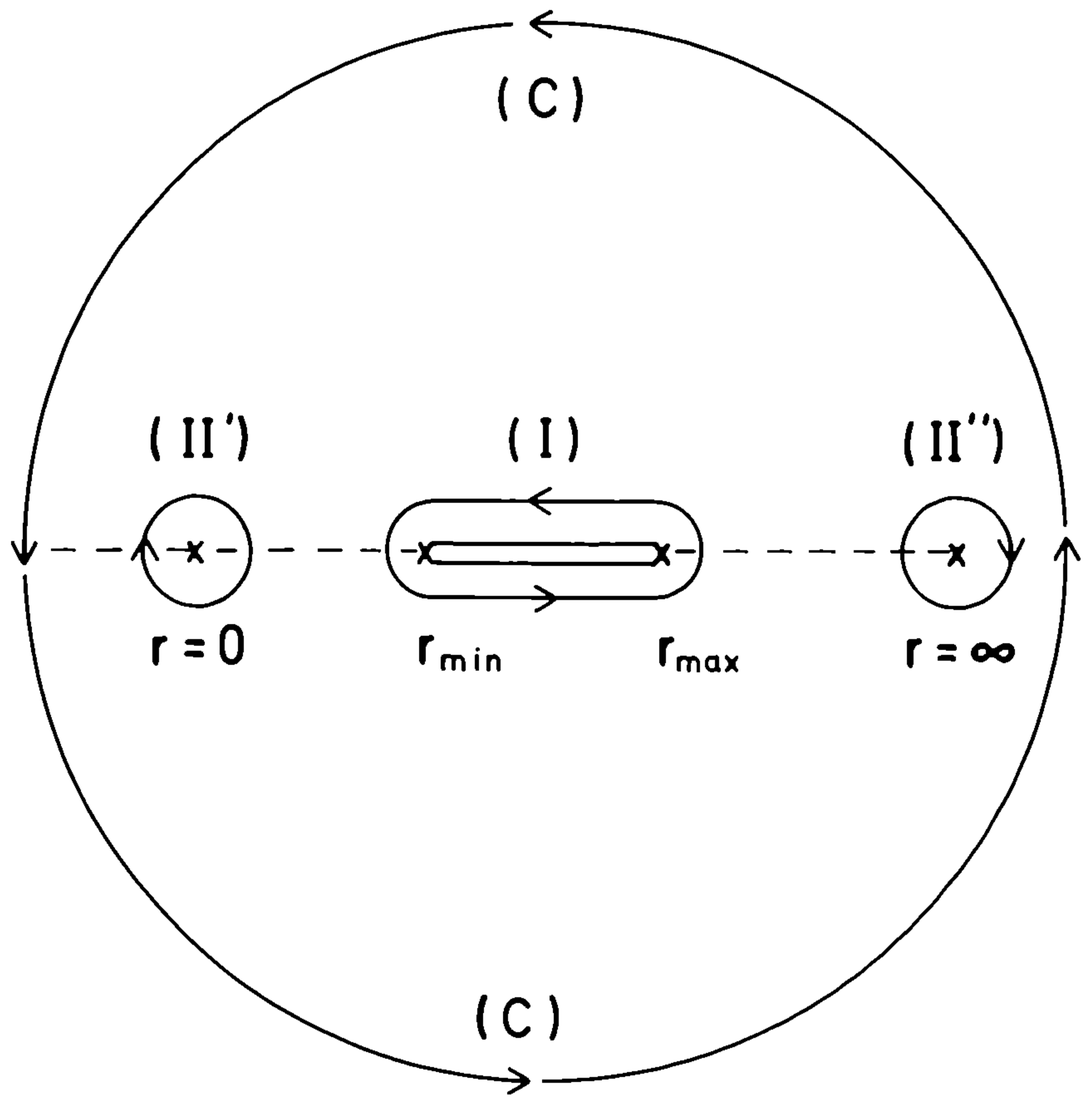


Figure 1 Sommerfeld's Complex Phase-Integral Evaluation.

of the degrees of freedom of the atomic system under investigation, the radial momentum p_r always assumed the form

$$p_r = \left(A + \frac{2B}{r} + \frac{C}{r^2} \right)^{1/2}, \tag{272}$$

with A, B and C the constants of integration, depending on the angular momentum and the energy. The functions (272) were not analytic on the real axis, where they possessed, for instance, a cut extending from $r_{\min}$ to $r_{\max}$. Therefore, the closed path in the integral on the right-hand side of Eq. (271) could not be carried out on the real axis, but had to pass from the first branching point, $r_{\min}$, into the complex plane of r on one Riemannian sheet, that is, a little below the real axis, until the second branching point, $r_{\max}$, was reached; then the path returned from $r_{\max}$ on the other Riemannian sheet of the square-root function, that is, a little above the real axis to $r_{\min}$. This path is denoted by (I) in Fig. 1. Sommerfeld had proposed to evaluate the integral (271) according to the rules of complex integration as follows. He turned the original path (I) into a large circle (C), within which all singularities were contained. Then the integral along the path (C) contributed nothing to the right-hand side of Eq. (271). But this was not the entire result of the integration; there still remained integrations to be carried out along the paths (II′) and (II″), which encircled the point singularities of the integral p_r, Eq. (272), lying at $r = 0$ and $r = \infty$, in the opposite sense of the original path (I). Their contribution was obtained from the residue formula (i.e., according to Cauchy's integral theorem) as

$$\Delta S_r = -2\pi i \left[\mathrm{Res}(0) + \mathrm{Res}(\infty) \right], \tag{273}$$

where Res(0) denotes the residue of p_r at $r = 0$ and Res(∞) the residue at $r = \infty$, which are the coefficients multiplying the first inverse power in the Laurent series for p_r around the points $r = 0$ and $r = \infty$, respectively.

Wentzel recognized that a similar situation arose in his theory of multiply periodic systems (Wentzel, 1926a, §3). When the quantum system is separable in the quantum variables, p_k and q_k, such that each p_k becomes a function of the conjugate variable q_k only, then the quantum-mechanical action function also becomes a sum of individual terms, S_k, $k = 1, \ldots, f$. Every S_k satisfies the f equations,

$$S_k J_l - J_l S_k = \tfrac{1}{2}\left[p_k(q_k J_l - J_l q_k) + (q_k J_l - J_l q_k)p_k \right], \qquad l = 1, \ldots, f, \quad (274)$$

which follow from Eq. (269a) and the commutation relations between p_k and q_k. Since the dependence of the momentum p_k on the coordinate q_k of a separable system in quantum mechanics is the same as in the Bohr–Sommerfeld theory, p_k may be expressed as a square root of a function of q_k similar to Eq. (272). Hence one has to take into account the cut in the complex plane associated with the quantum variable q_k when evaluating the periodicity modulus ΔS, Eq. (270). Wentzel assumed that in the complex plane associated with the quantum variable q_k a theorem similar to Cauchy's integral theorem would be valid; that is, the function S_k, defined by Eq. (274), returns to its initial value when the q_k-path is a closed curve in the complex plane within whose interior there is no singularity of the function $p_k(q_k)$, but it will change its value when the closed q_k-path encircles a pole or a cut of the same function. In particular, Wentzel expected that: 'One can, therefore, obtain the periodicity modulus of S, defined above, in the following way. Instead of taking the path around the branch-cut one goes around the poles of p in one of the Riemannian sheets, exactly as in Sommerfeld's procedure of evaluating the phase-integrals $\int p\,dq$' (Wentzel, 1926a, p. 84).

With this idea in mind, Wentzel first calculated the energy eigenvalues in two cases in which the results were well known from previous matrix calculations: the harmonic oscillator (Heisenberg, 1925c; Born and Jordan, 1925b) and the nonrelativistic hydrogen atom (Pauli, 1926a). Both calculations provided a welcome check upon his methods. However, Wentzel was particularly interested in treating the problem of the relativistic hydrogen atom, in which Pauli had not gotten anywhere with matrix methods. Wentzel succeeded brilliantly in demonstrating in this example the power of the action-angle scheme and, especially, of the reformulated method of complex integration. We shall now turn to it and present the principal steps of his solution.

In the relativistic hydrogen atom an electron, of rest mass m_e and charge $-e$, moves in a planar orbit (say, in the xy-plane) around the proton, whose charge is $+e$ and whose mass is so much larger than m_e that its motion can be neglected. The Hamiltonian, H, of the system is given as a function of the Cartesian coordinates, x and y, and their conjugate momenta, p_x and p_y, by means of the equation

$$p_x^2 + p_y^2 = 2m_e\left(H + \frac{e^2}{r}\right) + \frac{1}{c^2}\left(H + \frac{e^2}{r}\right)^2, \quad (275)$$

where $r = (x^2 + y^2)^{1/2}$ and c is the velocity of light *in vacuo*.[349] The separation of the variables had been achieved previously by transforming the Cartesian coordinates and momenta into spherical coordinates, r and ϕ, and momenta, p_r and p_ϕ. In quantum mechanics the Hamiltonian and the Cartesian dynamical variables are related formally in the same way as in classical theory; hence Eq. (275) is also valid for the corresponding quantum variables, x, y, p_x, p_y and H. Again, one can introduce spherical variables—the angle ϕ and the angular momentum p_ϕ, defined by Eqs. (212) and (252), respectively, and r and p_r—but one has to make sure that in doing so the commutation relations are satisfied. After carrying out the transformation to spherical variables one observes two facts: First, the transformed Hamiltonian does not depend on the angle ϕ, as in classical theory; hence p_ϕ *is a constant of motion*. Second, the quantum-mechanical relativistic hydrogen atom becomes a separable system in the new variables, in which p_r, the radial momentum, is given as a function of the radial coordinate r by the equation

$$p_r^2 = A + r^{-1}B + Br^{-1} + \left(C + \frac{1}{4}\,\frac{h^2}{4\pi^2}\right)r^{-2}, \tag{272'}$$

where the coefficients, A, B and C, depend on W, the energy of the electron (which replaces H, the Hamiltonian in Eq. (275), and p_ϕ, the angular momentum) as[350]

$$A = 2m_e W + \frac{W^2}{c^2}\,, \tag{272'a}$$

$$B = m_e e^2\left(1 + \frac{W}{m_e c^2}\right), \tag{272'b}$$

$$C = -p_\phi^2 + \frac{e^4}{c^2}\,. \tag{272'c}$$

Then, with the help of Eqs. (274), Wentzel proceeded to calculate the two terms, S_ϕ and S_r, of the action function, or, more accurately, the periodicity moduli of S_ϕ and S_r. With the help of Eq. (270), the expressions for J_ϕ and J_r, the action variables, are then obtained, which, via the coefficients A and B—Eqs.

[349] This formula can be found, for instance, in Sommerfeld's *Atombau und Spektrallinien* (Sommerfeld, 1919, p. 520). In the nonrelativistic limit the second term in Eq. (275) becomes negligible and the Hamiltonian assumes the well-known form, $H = (1/2m_e)(p_x^2 + p_y^2) - e^2/r$.

[350] Equation (272′) is obtained from Eq. (275) in the following way. One first observes that the left-hand side of Eq. (275), upon introducing the momenta p_r and p_ϕ (conjugate to the spherical coordinates), can be rewritten as $p_r^2 + p_\phi^2 - \frac{1}{4}(h^2/4\pi^2)/r^2$, which is identical with the classical expression (see Sommerfeld, 1919, p. 520), up to the term $-\frac{1}{4}(h^2/4\pi^2)r^{-2}$. The right-hand side can be left unchanged, except that the energy function W replaces the Hamiltonian H. In this equation ϕ does not occur, hence p_ϕ, as in classical theory, is a constant of motion. If one organizes the expression following from this transformed Eq. (275) for p_r in terms of inverse powers of the radial variable r and symmetrizes the products containing r^{-1}—because r does not commute with the energy term in B, Eq. (272′b)—one arrives at Eq. (272′).

(272′a) and (272′b)—involve the energy W. From these expressions, finally, W can be calculated as a function of the action variables, and the problem is solved.

The first part of this programme, the evaluation of the expression for J_ϕ, presented no difficulty: Wentzel simply found that J_ϕ was equal to the angular momentum p_ϕ.[351] The action function S_r, on the other hand, has to be determined from the equation[352]

$$S_r J_r - J_r S_r = \tfrac{1}{2}\left[\, p_r(rJ_r - J_r r) + (rJ_r - J_r r)p_r \right]. \tag{274a}$$

To solve this equation Wentzel used the analytic properties of the momentum p_r as a function of r, following from Eq. (272′). Evidently, this function exhibits a cut starting from one branch point, $r = r_{\min}$, to the other, $r = r_{\max}$, where $r_{\min}$ and $r_{\max}$ are the extremal values assumed by the radial variable r for a given W and a given angular momentum p_ϕ. The situation corresponds exactly to the one in the Bohr–Sommerfeld theory of atomic structure, where $\bar{p}_r$ is expressed by the right-hand side of Eq. (272). And, just as in the classical case, the quantum-mechanical function p_r, Eq. (272′), possesses poles for the values 0 and ∞ of the variable r. Hence, in the evaluation of the quantum-mechanical function only the singularities at $r = 0$ and $r = \infty$ will play a role.

The nature of these singularities is recognized from an expansion of p_r in a 'quantum mechanical Laurent series,' that is, a series of terms $P_n r^n$ or a series of terms $r^n P'_n$, where n assumes positive and negative integral values, and the coefficients P_n and P'_n will depend on expressions A, B and C, Eqs. (272′a, b, c). By substituting the Laurent series for p_r in Eq. (274a), S_r is obtained as a function of the radial variable r, the action variable J_r, and the P_n and P'_n, which may contain the action variable J_ϕ $(= p_\phi)$ via the coefficient C. This S_r can also be expanded in a Laurent series; however, apart from the power terms in the radial variable, terms proportional to $\ln r$ have to be added in all those cases in which the coefficients P_{-1} and P'_{-1}, multiplying the first inverse power of r in the Laurent series for p_r, are not zero. Hence S_r will assume the form

$$S_r = \tfrac{1}{2}(P_{-1} \ln r + \ln rP'_{-1} + \ldots), \tag{276}$$

where the dots denote the power terms in r.[353] Now the latter do not contribute

[351] This result was obtained as follows. Since p_ϕ, the angular momentum of the relativistic hydrogen atom, is a constant of motion, it commutes with the action variables J_r and J_ϕ. By applying Eq. (274), then, S_ϕ is given by the expression $S_\phi = \frac{1}{2}(p_\phi\phi + \phi p_\phi) + \text{const}$. Hence the periodicity modulus, due to Eq. (270), will be equal to $2\pi p_\phi$ because the angle ϕ changes by 2π.

[352] From the other equation for $S_r J_\phi - J_\phi S_r$ no information can be derived, as both sides are zero, due to the fact that J_ϕ is a constant of motion.

[353] Equation (276) is obtained as follows. One substitutes the Laurent series for p_r in Eq. (274a) and finds $S_r J_r - J_r S_r = \frac{1}{2}[P_{-1}(J_r - r^{-1}J_r r) + (rJ_r r^{-1} - J_r)P'_{-1} + \ldots]$, where the dots indicate contributions coming from terms with powers $n \neq -1$ in the Laurent series. Wentzel then proved the validity of the identities $\ln rJ_r - J_r \ln r = (J_r - r^{-1}J_r r) + \cdots = (rJ_r r^{-1} - J_r) + \cdots$ with the dots indicating terms with other powers of r (see Wentzel, 1926a, p. 85, Eq. (12)). Comparing the above equations with the equation for S_r, one recognizes that the $\ln r$-terms in S_r corresponds indeed to the r^{-1}-terms in the Laurent series for p_r.

to the change of the action function S_r when the variable r goes through a full period, while the logarithmic terms give ΔS_r, that is,

$$\Delta S_r = \pm \pi i (P_{-1} + P'_{-1}). \tag{273'}$$

Equation (273'), when applied to all poles of the function p_r, Eq. (272'), represents the quantum-mechanical analogue of the classical Eq. (273); it states that the periodicity modulus, ΔS_r, can be obtained with the help of the same ideas that were contained in Sommerfeld's evaluation by the complex integration method of the phase integral (271) with the momentum p_r, Eq. (272). The sign of ΔS_r has to be chosen in agreement with the classical limit, as will be seen below.

In the case of the relativistic hydrogen atom there are two singularities or poles of the function $p_r(r)$, at $r = 0$ and $r = \infty$, and the terms $n = -1$ (the 'residue' terms) in the quantum-mechanical Laurent series are found to be

$$P_{-1} = \sqrt{C} - \frac{1}{2}\,\frac{h}{2\pi i}, \qquad P'_{-1} = \sqrt{C} + \frac{1}{2}\,\frac{h}{2\pi i} \tag{277a}$$

for the pole at $r = 0$, and

$$P_{-1} = P'_{-1} = + \frac{B}{\sqrt{A}} \tag{277b}$$

for the pole at $r = \infty$. Hence the periodicity modulus of S_r is given by the equation[354]

$$\Delta S_r = 2\pi J_r = 2\pi\left(\frac{B}{\sqrt{-A}}\right) - (\sqrt{-C}). \tag{273'a}$$

Wentzel noted that: 'This, together with Eqs. [(272'a, b, c)], is exactly the formula from which Sommerfeld derived the fine structure of the Balmer series. It is true that we have been able to substantiate this formula here only in the first relativistic approximation; the higher approximations should, according to this method, be calculable only with great computational work' (Wentzel, 1926a, p. 94).[355] All one had to do to obtain the fine-structure formula was, of course, to

[354] For $r = 0$ the expansion for p_r is derived with the help of the identities $p_r = (1/2r)(p_r r + r p_r - h/2\pi i) = \frac{1}{2}(p_r r + r p_r + h/2\pi i)1/r$, where $\frac{1}{2}(p_r r + r p_r)$ can be replaced by the expression $(C + Br + rB + rAr)^{1/2}$; this can be expanded as the sum of $\sqrt{C}$ plus terms proportional to positive powers of r. Hence the 'residue,' $\pm \pi i(P_{-1} + P'_{-1})$, becomes $\pm 2\pi\sqrt{-C}$. Since in the classical analogue, the pole at $r = 0$ is encircled in the opposite direction as compared to the original path (I), the negative sign has to be chosen (see Fig. 1).

At $r = -\infty$ Wentzel used the expansion of p_r in terms of inverse powers of r, obtaining $P_{-1} = P'_{-1} = -BA^{-1/2}$. In the 'residue', $\pm 2\pi i(-BA^{-1/2})$, again the negative sign has to be taken. Hence the first term on the right-hand side of Eq. (273'a) is obtained.

[355] In 1915 Sommerfeld had obtained the result that $\oint p_r\, dr = 2\pi(B/\sqrt{-A}\,X - \sqrt{-C})$, where A, B and C are the expressions given by Eqs. (272'a, b, c) and $X = 1 - e^4/c^2 p_\phi^2$ (Sommerfeld, 1915c, p. 466). If the higher-order relativistic correction, $-e^4/c^2 p_\phi^2$, were neglected, Sommerfeld's result agreed formally with Wentzel's formula (273'a).

substitute the expressions for the coefficients, A, B and C, from Eqs. (272′a, b, c) into the right-hand side of Eq. (273′a); then one would be able to express the energy W as a function of the two action variables J_r and J_ϕ. Wentzel, in his paper, did not carry out this last, straightforward step.

The significance of Wentzel's theory of multiply periodic systems may be seen from two facts. First, it provided an effective method of dealing with complex atomic systems, which Born, Heisenberg and Jordan had failed to obtain although they spoke, in connection with the method of principal-axes transformation, 'about finding methods for direct calculations of the energy without explicitly solving the problem of motion: methods which correspond to Sommerfeld's method of complex integration' (BHJ, 1926, p. 581). What had been expressed as a wishful thought in the paper of Born, Heisenberg and Jordan, had now been accomplished by Wentzel by doing the most straightforward thing: translating Sommerfeld's method word by word into quantum mechanics. Wentzel had thus opened the way for treating, say, the problem of the helium atom. Although this aspect encouraged those who were interested in the helium problem, Wentzel's method of complex integration was not used in later calculations; in fact it was never applied beyond what Wentzel did himself, because the Schrödinger equation proved to be much more effective for obtaining the energy states of complex atomic systems.

The other aspect arising from Wentzel's work had, however, an historic importance. Wentzel had re-established the connection between quantum mechanics and the Bohr–Sommerfeld semiclassical theory of multiply periodic systems, a connection which seems to have been lost completely in the matrix approach. This connection expressed itself by the fact that the classical and the quantum-mechanical descriptions approach one another in a way similar to the one which Wentzel had suggested in summer 1924 with his formula, Eq. (263), for the quantum-theoretical acceleration. That is, there occurs a smooth transition from the quantum to the classical theory whenever the space-time description enters into the quantum-mechanical formalism. Wentzel would clarify the connection between the quantum and the classical theory, which was still somewhat hidden in his treatment of multiply periodic systems, in his later contributions to quantum mechanics, in particular in his discussion of the problem of determining eigenvalues in wave mechanics by means of a classical approximation (Wentzel, 1926c). Wentzel's work thus contributed to the physical interpretation of quantum mechanics by showing its connection with the classical theory.

As for the agreement between Wentzel's formula, Eq. (273′a), for the periodicity modulus of S_r and Sommerfeld's classical result, $\oint p_r\,dr$, it was not complete. The reason was that Sommerfeld had assumed, of course, that the action variable p_ϕ takes on values that are integral multiples of $h/2\pi$; in Wentzel's formalism, however, the action variables were only determined up to a constant which had to be obtained from further considerations. In a footnote at the end of his paper, Wentzel added: 'According to a personal communication, Mr. Pauli has obtained Eq. [(273′a)] by a perturbation calculation and thus found that the angular

momentum J_ϕ must be normalized as half-integral, which, as is well known, contradicts the original Sommerfeld interpretation of the hydrogen fine structure. The agreement with experience is re-established by the fact that one ascribes with [Uhlenbeck and] Goudsmit a magnetic moment to the electron (compare Heisenberg and Jordan, in press [Heisenberg and Jordan, 1926]' (Wentzel, 1926a, p. 94, footnote 1).

V.5 The Electron Spin, Fine Structure and Anomalous Zeeman Effects

In fall 1925 Uhlenbeck and Goudsmit had proposed to associate an intrinsic magnetic moment with each electron in order to describe the observed complex structure of atomic spectra consistently. This magnetic moment $\boldsymbol{\mu}$ should be related to $\mathbf{s}$, the angular momentum vector of its 'proper' rotation, by the equation

$$\boldsymbol{\mu} = \frac{e}{m_e c}\,\mathbf{s}. \tag{278}$$

The ratio $\boldsymbol{\mu}/\mathbf{s}$, i.e., the value $e/m_e c$ (where m_e and $-e$ denote the mass and charge of the electron, respectively, and c the velocity of light *in vacuo*), was assumed to be twice as large as the one obtained from the orbital motion of a point-like electron. As a consequence of the hypothesis of 'proper' rotation or 'spin' of the electron, a new model of atomic structure had emerged. Especially, the motion of an electron in the Coulomb field of a nucleus (having an atomic number Z and a mass much larger than m_e) and the action of an external magnetic field would be determined by a Hamiltonian H, which could be expressed as a sum of four terms, that is,

$$H = H_0 + H_1 + H_2 + H_3\,, \tag{279}$$

where the first, dominant term, H_0, accounts for the nonrelativistic motion of the *spinless* electron around the nucleus in the absence of the external field. The other three terms represent perturbations and are interpreted as follows. The interaction energy of the electron, having an orbital angular momentum $\mathbf{k}$ and spin $\mathbf{s}$, with the external magnetic field of strength $\mathbf{H}$, is described by the term H_1, where

$$H_1 = \frac{e}{2m_e c}\,\mathbf{H}\cdot(\mathbf{k} + 2\mathbf{s}). \tag{279a}$$

If the electron moves with a finite angular momentum $\mathbf{k}$ in the Coulomb field of the nucleus, an internal magnetic field, $\mathbf{H}_i$, will be induced in the atom, which acts on the magnetic moment $\boldsymbol{\mu}$ of the electron, producing a further magnetic perturbation term, H_2, in the Hamiltonian of the system. To evaluate the field

$\mathbf{H}_i$, Heisenberg had assumed (see Heisenberg to Pauli, 21 and 24 November 1925) that the electron is *at rest* and the nucleus moves with velocity **v**. Then, according to classical electrodynamics, $\mathbf{H}_i$ is given by the equation

$$\mathbf{H}_i = \frac{Ze(\mathbf{v} \times \mathbf{r})}{cr^3} = \frac{Ze\mathbf{k}}{m_e c}\left(\frac{1}{r^3}\right) \tag{280}$$

at the position of the electron. (In Eq. (280), **r** is the vector denoting the relative position of the electron with respect to the nucleus and r its magnitude; further, we have replaced the vector product $(\mathbf{v} \times \mathbf{r})$, multiplied by m_e, by **k**, the angular momentum of the electron, which is $m_e(\mathbf{r} \times \mathbf{v}_{el})$ or $-m_e(\mathbf{r} \times \mathbf{v}) = m_e(\mathbf{v} \times \mathbf{r})$.) Hence the internal field contributes a perturbation term H_2,

$$H_2 = (\mathbf{H}_i \cdot \boldsymbol{\mu}) = \frac{Ze^2}{m_e^2 c^2}\left(\frac{\bar{1}}{r^3}\right)(\mathbf{k} \cdot \mathbf{s}), \tag{279b}$$

to the energy of the atom, where the factor $\overline{r^{-3}}$ denotes the average over a full period of the orbiting electron (in the unperturbed atom). Finally, the relativistic corrections to the energy have to be considered as an additional perturbation term, H_3, which Sommerfeld (1915c, p. 470, Eq. (18)) had obtained in classical theory and could be written as

$$H_3 = -\frac{1}{2m_e c^2}\left[H_0^2 + 2e^2 Z H_0 \cdot \left(\frac{\bar{1}}{r}\right) + e^4 Z^2 \left(\frac{\bar{1}}{r^2}\right)\right], \tag{279c}$$

where H_0 is the energy of the unperturbed nonrelativistic atom.

These equations were supposed to constitute the basis of the magnetic and relativistic term splitting of the atoms. In spite of his reservations concerning the electron spin hypothesis, which we have mentioned earlier (Section V.1), Heisenberg had started already in November 1925 to put the entire model 'through the matrix mill' (Heisenberg to Pauli, 24 November 1925). Towards the end of the year he had reported to Pauli the preliminary results he had obtained in collaboration with Jordan. Although these appeared to be rather encouraging, the question had still remained open whether electron spin fitted into the quantum-mechanical scheme or even made any sense at all. Bohr, for example, was tempted to accept Uhlenbeck and Goudsmit's hypothesis as denoting important progress in the theory of atomic structure, while Pauli raised any number of objections. In between the two protagonists was Heisenberg, who saw the advantages of using 'Goudsmit's theory'—as he called it (Heisenberg to Pauli, 24 December 1925)—in order to explain the complex structure and the Zeeman effects of atomic spectra. But he was very bothered by the apparent failure of the theory to describe the splitting quantitatively, especially the factor 2 in Eq. (211) which spoiled the agreement between the observed and calculated fine-structure results. On the other hand, Heisenberg had also noted that, before one could reach a final verdict against the spin hypothesis, one had first to

evaluate consistently (in quantum mechanics) the averages, $\overline{r^{-2}}$ and $\overline{r^{-3}}$, occurring in Eqs. (279b) and (279c). This task involved an extension of the matrix scheme to include also a treatment of aperiodic variables, such as the angle variables. During January 1926 Heisenberg and Pauli, following the hints from Born and Wiener's operator scheme, had been able to develop the quantum-mechanical action-angle theory far enough that they could apply it for the purpose of calculating the desired average values. In particular, for a one-electron atom (with atomic number Z) whose electron had the angular momentum p_ϕ, they had obtained the results

$$\left(\frac{\bar{1}}{r^2}\right) = m_e \frac{\bar{\dot{\phi}}}{p_\phi}, \qquad \text{and} \qquad \left(\frac{\bar{1}}{r^3}\right) = \left(\frac{\bar{1}}{r^2}\right) \frac{e^2 Z m_e}{p_\phi^2 - \frac{1}{4}(h/2\pi)^2}, \tag{281}$$

where $\bar{\dot{\phi}}$ denoted the average over a full period (of the unperturbed motion of the electron) of the time derivative of the angle ϕ (the canonically conjugate variable to the angular momentum p_ϕ). The value of $\bar{\dot{\phi}}$ was given by the time derivative of the variable w, the canonically conjugate variable to the action variable J $(= nh)$ of the unperturbed atom, that is,

$$\bar{\dot{\phi}} = \dot{w} = -\frac{\partial H_0}{\partial J} = +\frac{4\pi R Z^2}{n^3}, \tag{282}$$

where $R = 2\pi^2 e^4 m_e / h^3$ is Rydberg's constant (see Heisenberg to Pauli, 27 January 1926). Since $\bar{\dot{\phi}}$ was thus found to be independent of the angular momentum **k** of the atom, a difficulty arose with the magnetic perturbation term, Eq. (279b), in the lowest state of the atoms; namely, in the *s*-state the denominator $p_\phi^2 - \frac{1}{4}(h/2\pi)^2$—which is equal to $k(k+1)(h/2\pi)^2$, with k the quantum number of the angular momentum **k** of the electron—in the expression for $\overline{r^{-3}}$ becomes zero and H_2 consequently infinite. 'For the [Uhlenbeck–] Goudsmit theory, however, this result is very sad,' Heisenberg commented on it in a letter to Pauli. 'If you have an idea, do please write to me; I don't know what to do with the Zeeman effects without Goudsmit [i.e., without the spin hypothesis]' (Heisenberg to Pauli, 27 January 1926).

Pauli agreed with Heisenberg's calculations and the conclusions he had drawn from them. 'Therefrom it follows,' he wrote back, 'first, that the numbers for the sum $E_{\text{rel}} + E_{\text{mag}}$ [Pauli's E_{rel} and E_{mag} were the quantum-mechanical terms H_2 and H_3, Eqs. (279b) and (279c), respectively] following from our equations are certainly completely inconsistent with the observations . . . Hence it appears that on the basis of quantum mechanics the fine structure does not come out correctly *either* with the [Uhlenbeck–] Goudsmit electrons *or* with point electrons [i.e., the electrons without "proper" rotation and intrinsic magnetic moment]!' (Pauli to Heisenberg, 31 January 1926).

A few days later, in a letter to Niels Bohr, Pauli reported 'of the great catastrophe, to which the new quantum mechanics leads concerning the fine

structure of the hydrogen spectrum' (Pauli to Bohr, 5 February 1926).[356] He listed in detail the various difficulties which arose in explaining the fine structure in quantum mechanics, both with a point-like electron and an Uhlenbeck–Goudsmit (i.e., spinning) electron, and claimed that neither model accounted for the empirical data; while the point-electron led, in connection with the observed selection rules, to a theoretical inconsistency, the electron with spin yielded quantitatively wrong results. 'The "moral of the story,"' he concluded, 'seems to be in the following: *It is true that one can try to associate with the electron a preferred axis with angular momentum and magnetism; the interaction energy of such an electron with a nucleus, as well as the interaction energy of two such electrons—so far as it deviates from that of point electrons and depends on the orientation of the electron axes—cannot, however, be simply and only represented by magnetic dipoles*' (Pauli to Bohr, 5 February 1926, Pauli's emphasis). In Pauli's view these difficulties—which also included the problems of explaining the singlet–triplet structure of the spectra of helium and alkaline-earth elements quantitatively, as well as the exclusion of several equivalent electrons in atoms—had to do with a problem, noted already by Uhlenbeck and Goudsmit, namely, whether an intrinsic angular momentum of the electron of the order of $h/4\pi$ and the small mass of the electron were mutually compatible at all (Uhlenbeck and Goudsmit, 1925, p. 954, second column, footnote 2). Hence he expected that only a treatment of the electron theory on a deeper level would give a satisfactory answer to these questions.

And yet, a more fundamental level of atomic and electron theory, which Pauli envisaged in February 1926, was really not necessary in order to remove the difficulties in the calculation of the fine structure of hydrogen spectral lines. Remarkably enough, Pauli had already indicated the solution of the problem in his letter to Heisenberg at the end of January, when he made what he called 'a totally stupid remark' ('*eine ganz dumme Bemerkung*'): namely, if one assumed the magnetic term arising from the action of the internal field, H_2, to have only one-half of the value given in Eq. (279b), then the diagonalization of the matrix replacing the sum $H_2 + H_3$, where H_3 is taken from Eq. (279c), would exactly reproduce Sommerfeld's fine-structure formula. But where would one possibly find the necessary factor of $\frac{1}{2}$? 'Therefore: "A half, a half, a kingdom for the factor $\frac{1}{2}$,"' Pauli exclaimed, paraphrasing King Richard's desperate appeal ('*Also: "Ein halb, ein halb, ein Königreich für den Faktor* $\frac{1}{2}$,"' Pauli to Heisenberg, 31 January 1926).

[356] As we have mentioned in Section V.1, Pauli had discussed the spin problem very often with Bohr, both in letters and conversations. Bohr had tried to convince him of the correctness of Uhlenbeck and Goudsmit's hypothesis and also suggested certain guesses for the quantum-theoretical values of H_3 and H_2, notably (see Bohr to Pauli, 24 December 1925) $E_{\text{mag}} = C[\pm\frac{1}{2}/k(k+1)]$ and $E_{\text{rel}} = C(k+\frac{1}{2})/k(k+1)$, with $C = 8\pi^2 Z^2 e^4 m_e / n^3 h^4 c^2$. Pauli, in his reply, had rejected these formulae because of several reasons (Pauli to Bohr, 30 December 1925). First, he argued that the angular momentum quantum number k was zero in an s-state and, therefore, the fine-structure splitting would be infinite; second, he insisted on a derivation of the terms H_2 and H_3 from a quantum-mechanical calculation.

Pauli saw no possibility of getting the factor $\frac{1}{2}$ from the theory, but Llewellyn Hilleth Thomas, a young visitor to Bohr's Institute from Cambridge, England, soon found it.[357] By a systematic relativistic discussion of the motion of an electron around the atomic nucleus, Thomas noticed that in calculating the internal magnetic field $\mathbf{H}_i$ one was not allowed to substitute, as Heisenberg had done in deriving Eq. (280), simply the nucleus–electron relative velocity $\mathbf{v}$. Rather, the rotational character of the electron's motion had to be taken into account. Thus, Thomas found that the rest frames of the electron (orbiting with velocity $\mathbf{v}$ around the nucleus) at two infinitesimally neighbouring times, t and $t + dt$, could be related to each other by an infinitesimal Lorentz transformation with velocity $d\mathbf{v} = \dot{\mathbf{v}}\,dt$ plus an infinitesimal rotation described by the axial vector $d\mathbf{w} = (1/2c)(\dot{\mathbf{v}} \times \mathbf{v})dt$ (Thomas, 1926; see also Frenkel, 1926). The field, acting on $\boldsymbol{\mu}$, the intrinsic magnetic moment of the electron, was, therefore, not given by the expression on the right-hand side of Eq. (280), but the term $(m_e c/e)(d\mathbf{w}/dt)$ had to be subtracted from it; that is, the effective internal field was only one-half of the field given by Eq. (280). Consequently, Eq. (279b) for the perturbation term H_2 had to be replaced by the equation

$$H_2 = \frac{1}{2}\,\frac{Ze^2}{m_e^2 c^2}\left(\frac{\overline{1}}{r^3}\right)(\mathbf{k}\cdot\mathbf{s}), \tag{279b′}$$

and the factor $\frac{1}{2}$ that Pauli had searched for in vain was provided.[358]

Niels Bohr, the advocate of the electron spin hypothesis for several months, was extremely happy with Thomas' result and he immediately informed Pauli about it. Thus, on the same day that Thomas completed his note and sent it to *Nature*, Bohr wrote to Pauli:

> Now, however, comes the surprise for the learned relativity theoretician and the scientist laden-with-responsibility [in Danish: "*laerde Relativitetsteoretikere og ansvarstyngede Videnskabsmaend,*" that is, Pauli]. A young Englishman [Llewellyn Hilleth] Thomas, who spent the past half year here, has recently found that the entire problem of the unfortunate factor 2 rests probably exclusively on an error in the calculation of the relativistic motion of the electron and of the nucleus. He has tried to give an elementary presentation of his consideration in a note, which he intends to send to *Nature*, and I enclose a copy of it with the request for severe criticism. (Bohr to Pauli, 20 February 1926)

Pauli studied Thomas' note, but he was not satisfied at all. He especially criticized two facts: that Thomas had not calculated the expression for the energy

[357] L. H. Thomas, born on 21 October 1903 in London, became a student at Cambridge in 1921, obtaining there his B.A. in 1924 and his doctorate in 1927. During winter 1925–1926 he spent several months in Copenhagen as Isaac Newton Student; there he found the general relativistic argument in favour of electron spin. In 1929 he joined Ohio State University as an assistant professor; he was promoted to associate professor in 1930 and made a full professor in 1936. In 1946 he joined the Watson Scientific Computing Center in New York City, and in 1950 became a professor of physics at Columbia University. In his later scientific career, Thomas worked especially on electron distribution in atoms and molecules and on the use of computing machines in solving scientific problems.

[358] Uhlenbeck and Goudsmit has also come across the difficulty concerning the factor $\frac{1}{2}$ in connection with the quantitative application of Eq. (279b). (See Bohr to Pauli, 20 February 1926.)

term arising from the interaction of the electron's magnetic moment with the internal field explicitly, but only the change of the angular momentum of the electron; and that he had discussed the situation only in frames moving with the electron and not in the frame where the nucleus is at rest. He came to the judgment: 'I hold Thomas' conclusion that, "The difficulty of explaining at the same time the spectral structures *therefore* disappears," to be wrong or at least not to be sufficiently justified . . . In any case, I regard the publication of the present note by Thomas in *Nature* as a mistake, and would be glad if you [i.e., Bohr] would prevent it or achieve a substantial change in the text of the note' (Pauli to Bohr, 26 February 1926).

The reason for Pauli's strong verdict was that he did not believe that any kinematical investigation could resolve the difficulty. He emphasized this belief very clearly in his letter to Bohr by saying: 'But I have the strong feeling that in this case absolutely nothing will be achieved by formal learning ("*formaler Gelehrsamkeit*") and that, if the assumption of the magnetic electron really strikes at the truth, a more detailed investigation of the structure of the electron and especially a theory of the electron-mass (eventually by taking into account gravitational fields) will be necessary' (Pauli to Bohr, 26 February 1926). Pauli admitted at the same time that he did not yet have the courage to undertake a study of these profound problems, though he did not consider it an '*a priori* hopeless undertaking' ('*ein von vornherein hoffnungsloses Unternehmen*,' Pauli to Bohr, 26 February 1926). Evidently, Pauli divided the problem connected with the role of the electron's properties in the explanation of atomic spectra into two classes: the *fundamental* problems included the theory of the electron's mass, the basic idea for the calculation of the separation of singlet and triplet terms in the helium spectra, and the physical reasons for the existence of the exclusion principle; a more *modest* task seemed to be to derive the fine structure of the hydrogen atom and the anomalous Zeeman effects from the same unified assumptions concerning the properties of the electron. (See Pauli to Bohr, 5 March 1926.)

Niels Bohr, though he sympathized with Pauli's opinions, nevertheless tried to convince him that Thomas' calculations were correct and had solved the problem of the factor $\frac{1}{2}$ satisfactorily. In particular, he argued that the use of the coordinate reference frames moving with the electron was justified (Bohr to Pauli, 3 March 1926). Still Pauli did not give in; in his letter to Bohr, dated 5 March 1926, he again repeated his objections against Thomas' kinematic treatment, claiming that the transformation of the results from the moving frames of reference to the coordinate system in which the nucleus was at rest might lead back to the old difficulties about the factor $\frac{1}{2}$. A few days later, on 8 March, Samuel Goudsmit, who had spent several weeks in Copenhagen, passed through Hamburg; he brought news about the details of Thomas' calculations for the purpose of finally converting the nonbelieving Pauli. But his efforts failed completely, and Pauli immediately sent a letter to Copenhagen, addressing it this time to Kramers. 'Today Mr. Goudsmit was here and I was very happy to see him,' Pauli wrote. 'What he told me about Thomas' calculations and Thomas'

manuscripts, which he had brought along, have strengthened my opinion most strongly that I am perfectly right in the arguments, which I have tried to formulate in my last two letters to Professor Bohr' (Pauli to Kramers, 8 March 1926). Even in greater detail than he had written previously to Bohr, he discussed the question of the different coordinate systems, appealing to Kramers' thorough knowledge of relativity theory and requesting him to check the details. 'This time it cannot, I believe,' he wrote, 'be difficult at all to reach an agreement, because we are not dealing with a religious or political matter, indeed not even with a quantum-theoretical question, but rather with a problem of classical physics' (Pauli to Kramers, 8 March 1926).

The last remark was correct: the decision whether Thomas' calculation gave the correct answer or not was based on a classical argument. Pauli also did not have to wait until Kramers replied. The answer was already contained in a letter which Bohr had written soon after receiving Pauli's letter of 5 March. In particular, Bohr had drawn attention to one fact: namely, that the goal of the calculation had been to obtain the average value of the internal magnetic field H_i or the associated perturbation energy; hence it should not matter what specific frame of reference one used, 'for the only thing, about which one can talk by the very nature of the problem, is the change of the orientation of the electron's [angular momentum] vector after one revolution in the electron orbit, where the velocity relative to the nucleus is the same as before [the revolution]' (Bohr to Pauli, 9 March 1926). Then, finally, Pauli gave in:

> Many thanks for your letter of 9 [March] and the enclosed remarks of Thomas [he replied]. Now all that remains for me is to *capitulate completely*! . . . In any case, I must now answer both questions of your letter in the sense you desired: 1. Thomas' calculation of the secular change of the electron's axis [i.e., of the electron's angular momentum] is correct in the approximation he considered. 2. Of course, I always had the opinion that the magnitude of this change is decisive for the fine structure. The corresponding energy change of the atom basically states nothing else physically; it must, I believe, always result formally automatically, if one describes the associated secular perturbation by equations of motion of the canonical type. (Pauli to Bohr, 12 March 1926)

The same day Pauli wrote a letter to Kramers, and on 13 March a postcard to Goudsmit, in which he withdrew his reservations against the electron spin hypothesis and acknowledged the correctness of the factor $\frac{1}{2}$ in Eq. (279b′).

With Pauli's capitulation the path was clear to solve what Bohr and Pauli had called the 'modest' part of the electron problem; namely, to perform quantum-mechanical calculations of the anomalous Zeeman effects and the fine structure of the hydrogen lines. In fact, Heisenberg had reported about these calculations, which he had carried out together with Pascual Jordan, on 15 February 1926 at the meeting of the Lower Saxony Section (*Gauverein Niedersachsen*) of the German Physical Society in Braunschweig.[359] All one had to do now was to bring

[359] See the announcement of Heisenberg's talk, entitled '*Über die Anwendung der Quantenmechanik auf das Problem der anomalen Zeemaneffekte*' ('On the Application of Quantum Mechanics to the Problem of the Anomalous Zeeman Effects'), in *Verh. d. Deutsch. Phys. Ges.* (3) **7** (1926), p. 13.

in the factor $\frac{1}{2}$ obtained by Thomas. This was done by Jordan, for Heisenberg had left Göttingen after the end of the winter semester for an extended vacation, and a few days after 12 March, the date of Pauli's capitulation, he (Jordan) sent the joint paper on the anomalous Zeeman effects—which also included the fine structure of hydrogen lines—to *Zeitschrift für Physik*, where it was published three months later (Heisenberg and Jordan, 1926).[360]

The main theoretical task, which Heisenberg wanted to achieve in the paper with Jordan, was to calculate the eigenvalues of the quantum-theoretical formulation of the Hamiltonian H, Eq. (279), with the perturbation terms given by Eqs. (279a), (279b′) and (279c), and to obtain from them the anomalous Zeeman effects and fine structure of spectral lines. To achieve this goal, the authors proceeded in several steps. First, they studied the question of whether the well-known selection and intensity rules, which had been obtained from the analysis of the anomalous Zeeman effect data, could be derived from the matrix mechanical treatment of the problem. The calculation of the amount of magnetic splitting turned out to be strenuous work, involving as it did the exact evaluation of certain inverse powers of the matrix of the radial variable for realistic three-dimensional atoms. With the extension of the matrix scheme to include angle variables, such as the one developed by Heisenberg and Pauli in January 1926, the authors succeeded in obtaining the desired results, which they summarized in the abstract of their paper in the following words: 'For explaining the anomalous Zeeman effect, Uhlenbeck and Goudsmit have applied Compton's hypothesis of the rotating electron. In the present paper [we] investigate the quantum mechanical behaviour of the atomic model characterized by this hypothesis. The result is that the Zeeman effects and the fine structures of the doublet spectra can be explained completely by the said hypothesis' (Heisenberg and Jordan, 1926, p. 263).[361]

The actual calculations of Heisenberg and Jordan were straightforward, for they followed exactly the pattern of matrix methods which had been developed

[360] Since Heisenberg was on vacation, he did not immediately receive the news about developments concerning the problem of electron spin. Thus, he did not witness Pauli's withdrawal of objections against Thomas' calculation. On 4 April 1926 Heisenberg wrote to Jordan from Munich: 'Since the day before yesterday I am again back in Germany [probably from a skiing trip to Austria], [and I am] filled with the best will to begin a new life and to understand again something of physics. Could you be so kind as to write to me what you have done with our unhappy work [on the anomalous Zeeman effects], whether you believe in the factor 2 [in the denominator of Eq. (279b′)], and whether you have meanwhile brought into order the radiation theory and all quantum theory of the world?' Jordan sent him the desired information immediately, especially a copy of the finished manuscript on the anomalous Zeeman effects (which he had already submitted for publication three weeks earlier), and Heisenberg wrote back: 'I admire your industry in rewriting our clumsy piece of work' (Heisenberg to Jordan, 9 April 1926). Thus, he agreed with Jordan's final version of the joint paper. Still, he wanted to know more details about the story of the factor 2, for he added: 'That you do not write to me why Thomas is right with the damned factor 2 is a shrieking scandal' (Heisenberg to Jordan, 9 April 1926).

[361] Heisenberg and Jordan referred here to the fact that already Arthur H. Compton had introduced the idea of a rotating electron in August 1921 (Compton, 1921d). Uhlenbeck and Goudsmit took note of this fact in their later papers (Goudsmit and Uhlenbeck, 1926, p. 625, footnote 2; Uhlenbeck and Goudsmit, 1926, p. 264).

in the three-man paper and in Pauli's treatment of the hydrogen atom. In order to diagonalize the quantum-mechanical Hamiltonian corresponding to Eqs. (279), (279a), (279b′) and (279c) with the help of perturbation theory, the authors had first to determine the degree of degeneracy of the unperturbed Hamiltonian H_0.[362] Although H_0 is degenerate with respect to **k**, the angular momentum of the electron, that is, the same energy eigenvalues W_n arise for different values of the quantum number k (which runs from $0, 1, 2, \ldots$ to $k_{\max}$, with $k_{\max} = n - 1$, where n is the principal quantum number—see Pauli's treatment of the hydrogen atom in Section IV.5), for the purpose of calculating the anomalous Zeeman effect one may assume that it is a diagonal matrix in H_0. The reason why this assumption can be made is that H_1 and H_2, the perturbation terms given by Eqs. (279a) and (279b′), respectively, depend only on **k** but not on its conjugate (angle) variable; hence already infinitesimally small perturbations of the type H_1 or H_2 will remove the degeneracy of H_0 with respect to the values of k without changing the other properties of the system. H_0 is also degenerate with respect to both the values of the spin **s** and the values of the total angular momentum **M** (composed of **k** and **s**). Both of these degeneracies are removed if an infinitely small external magnetic field **H** is applied in the z-direction.

There is, however, still one coordinate of the system with respect to which H_0 remains degenerate, as can be seen from Eq. (279b′) for the perturbation term due to the interaction of the spinning electron with the internal field. The expression for H_2 contains both **k** and **s** but not their conjugate angle variables. The only angle which occurs is the one between **k** and **s**, due to the scalar product of these vectors; hence one of the angle variables will not be determined, and H_0 must be considered degenerate with respect to the degree of freedom described by it. Heisenberg and Jordan chose this angle variable in such a way that H_0 became degenerate with respect to the values of s_z, the z-component of the spin **s**. Moreover, they observed that since the spin did not occur in the perturbation term H_3, Eq. (279c), the perturbation calculation had to be carried out only with the terms H_1 and H_2, Eqs. (279a) and (279b′). The relativity term H_3 thus contributed merely an additive constant to the final result, depending, of course, on the values of n (the so-called principal quantum number) and k which the state under investigation assumed.

Heisenberg and Jordan now turned to the evaluation of $H_1 + H_2$, the sum of the perturbation terms, Eqs. (279a) and(279b′), writing it as

$$H_1 + H_2 = \mu \frac{2\pi}{h}(k_z + 2s_z) + \lambda\left(\frac{2\pi}{h}\right)^2\left[k_z s_z + \frac{1}{2}(k_x + ik_y)(s_x - is_y) + \frac{1}{2}(k_x - ik_y)(s_x + is_y)\right], \tag{283}$$

[362] In agreement with Heisenberg and Jordan we drop the explicit matrix notation in the present section. Any quantum-theoretical variable which can be represented by a matrix is then assumed to be a matrix without being explicitly identified as such. On the other hand, the quantum-mechanical variables which are not periodic, such as angles, etc., will be written as they normally are; they just do not have proper matrix elements. (See Sections V.3 and V.4.)

where

$$\mu = \frac{he|\mathbf{H}|}{4\pi m_e c} \quad \text{and} \quad \lambda = \frac{Ze^2h^2}{8\pi^2 m_e^2 c^2}\left(\frac{\bar{1}}{r^3}\right). \tag{283a}$$

That is, they assumed the magnetic field $\mathbf{H}$ (with the absolute value $|\mathbf{H}|$) to be in z-direction and took into account only those transitions which occurred for constant values of the orbital angular momentum $\mathbf{k}$, the spin momentum $\mathbf{s}$ and M_z, the z-component of the total angular momentum $\mathbf{M}$ $(= \mathbf{k} + \mathbf{s})$, respectively.[363] The latter values were denoted by k (obtained from $\mathbf{k}^2 = (h/2\pi)^2 k(k + 1)$, by s (obtained from $\mathbf{s}^2 = (h/2\pi)^2 s(s + 1)$) and by j (obtained from $\mathbf{M}^2 = (h/2\pi)^2 j(j + 1)$. Thus, the matrix λ, which depends on n and k, i.e., $\lambda = \lambda(n, k)$, could be replaced by a constant in the following calculations.

According to the angular momentum formalism, the only matrix elements which do not vanish are $k_z(k, m_k; k, m_k)$, $(k_x + ik_y)(k, m_k; k, m_k + 1)$ and $(k_x - ik_y)(k, m_k; k, m_k - 1)$, and $s_z(s, m_s; s, m_s)$, $(s_x + is_y)(s, m_s; s, m_s + 1)$ and $(s_x - is_y)(s, m_s; s, m_s - 1)$. Here m_k and m_s denote, respectively, the quantum numbers of z-components of the orbital angular momentum (i.e., $k_z = (h/2\pi)m$) and of the spin angular momentum (i.e., $s_z = (h/2\pi)m_s$); moreover, one can add m_k and m_s to give m $(= m_k + m_s)$, the quantum number of the total angular momentum. For constant values of k, s, j and m, one obtains the following elements of the perturbation matrix:

$$(H_1 + H_2)(m_s, m_s) = \mu(m + m_s) + \lambda m_s(m - m_s), \tag{284a}$$

$$(H_1 + H_2)(m_s, m_s - 1) = (H_1 + H_2)(m_s - 1, m_s)$$

$$= \tfrac{1}{2}\lambda\left\{\left[s(s + 1) - m_s(m_s - 1)\right]\cdot\right.$$

$$\left.\left[k(k + 1) - (m - m_s)(m - m_s + 1)\right]\right\}^{1/2}. \tag{284b}$$

For a given set of quantum numbers k, s and m, the z-component of the spin, s_z, assumes values $m_s(h/2\pi)$ such that $k + m \geq m_s \geq -k + m$. For instance, if $s = \frac{1}{2}$ (which applies only when one electron is considered) and $m = k + \frac{1}{2}$, with k arbitrary, m_s will assume only the value $\frac{1}{2}$ (because the condition, $-s \leq s_z \leq +s$, also holds), and no degeneracy will arise in this case; the same is true for $m = -k - \frac{1}{2}$. However, when m lies between $-k - \frac{1}{2}$ and $k + \frac{1}{2}$, m_s will always have exactly two values, $+\frac{1}{2}$ and $-\frac{1}{2}$.

In order to calculate the energy contributed to a given state by the interaction of the spinning electron with the magnetic fields, external and internal, Heisenberg and Jordan made use of the perturbation theory of degenerate systems, especially of Eq. (154) derived in the three-man paper. The submatrix of

[363] Note that we have, in agreement with the authors, introduced the notation $\mathbf{M}$ for the total angular momentum.

$H_1 + H_2$, which has to be substituted in Eq. (154), is a (2×2)-matrix in the case of $s = \frac{1}{2}$, i.e., for doublet spectra. Hence, with Eqs. (284a) and (284b), one obtains the secular or characteristic equation,

$$\det\begin{vmatrix} W - \mu\left(m - \frac{1}{2}\right) + \frac{\lambda}{2}\left(m + \frac{1}{2}\right), & -\frac{\lambda}{2}\left[k(k+1) - \left(m + \frac{1}{2}\right)\left(m - \frac{1}{2}\right)\right]^{1/2} \\ -\frac{\lambda}{2}\left[k(k+1) - \left(m + \frac{1}{2}\right)\left(m - \frac{1}{2}\right)\right]^{1/2}, & W - \mu\left(m + \frac{1}{2}\right) - \frac{\lambda}{2}\left(m - \frac{1}{2}\right) \end{vmatrix}$$

$$= 0, \tag{285}$$

which has the two solutions,

$$W_m = \mu\left[m - \frac{v}{4\left(k + \frac{1}{2}\right)} \pm \frac{1}{2}\left(1 + 2\frac{m}{k + \frac{1}{2}}v + v^2\right)^{1/2}\right], \tag{286}$$

where

$$v = \frac{\lambda}{\mu}\left(k + \frac{1}{2}\right), \tag{286a}$$

provided m does not assume the extremum values $\pm(k + \frac{1}{2})$. If it does, the degeneracy is automatically removed, and the submatrix of $(H_1 + H_2)$ for the states characterized by the quantum numbers $k, s = \frac{1}{2}$ and $m = \pm(k + \frac{1}{2})$ contains only one (diagonal) element given by

$$W_{m = \pm(k+1/2)} = \mu\left[\pm(k+1) + \frac{v}{2}\left(\frac{1}{2} - \frac{1}{(k + \frac{1}{2})}\right)\right]. \tag{287}$$

These formulae were identical with the ones obtained by Sommerfeld when he re-interpreted Woldmar Voigt's equations for the anomalous Zeeman effects in accordance with the Bohr–Sommerfeld theory of atomic structure (Sommerfeld, 1922a; see also Heisenberg, 1922a, p. 284, Eq. (18)).[364] The theoretical understanding of the anomalous Zeeman effects in the old approach (which we discussed in Volume 1, Chapter IV) was, however, fundamentally different from the present calculation based on quantum mechanics; the coincidence of the results obtained by so different means was, therefore, deemed fortunate. Heisenberg was really 'excited and enthusiastic about the fact that the old formulae of Voigt actually did come out of quantum mechanics' (Heisenberg, Conversations).

The difference between the methods used by Sommerfeld, on the one hand, and by Heisenberg and Jordan, on the other, showed up in particular in the calculation of the intensities of the Zeeman lines. Voigt had discussed the

[364] In the paper of Heisenberg and Jordan (1926), the first equation (p. 270, Eq. (23)) contained a few errors; for instance, in the denominator of the second term in the bracket one had to insert a factor 4, and the square had to be multiplied by a factor $\frac{1}{2}$.

question of line intensities on the basis of the classical electromagnetic theory of emission of light by bound charges, and Sommerfeld had not changed this aspect in his treatment of 1922. In Heisenberg and Jordan's calculation the intensities had to be obtained from the matrix elements of **q**, the position vector matrix of the orbiting electron. The amplitudes of the radiation emitted from a rotating electron in an unperturbed atom, when a state having the angular momentum $kh/2\pi$ passes into a state of angular momentum $(k-1)h/2\pi$, were proportional to the quantities

$$|q_z^0(k,m,m_s;k-1,m,m_s)|^2, \qquad |(q_x^0+iq_y^0)(k,m-1,m_s;k-1,m,m_s)|^2$$

and

$$|(q_x^0-iq_y^0)(k,m,m_s;k-1,m-1,m_s)|^2,$$

respectively. These quantities had been calculated in the three-man paper and could be taken over now (see BHJ, 1926, p. 604, Eqs. (31)–(33)).[365] The perturbations due to the magnetic fields transform the position matrices from $\mathbf{q}^0$ to **q**, and also give rise to transitions in which m_s changes by $0, \pm 1$. The matrix S, which generates this transformation of the position matrices, especially the particular submatrices of S belonging to a given set of values, k, s and m, can be calculated from the perturbation theory of degenerate systems.[366] Thus, for the transition $k=1$ to $k=0$, $m=\frac{1}{2}$ to $m=\frac{1}{2}$, $m_s=\pm\frac{1}{2}$ to $m_s=\frac{1}{2}$, Heisenberg and Jordan obtained the result (Heisenberg and Jordan, 1926, p. 272, Eq. (30))

$$\left|q_z\left(1,\frac{1}{2},\pm\frac{1}{2};0,\frac{1}{2},\frac{1}{2}\right)\right|^2=\frac{1}{2}|A(1)|^2\left[1\pm\frac{\mu^2+\frac{1}{2}\lambda}{\left(\mu^2+\lambda\mu+\frac{9}{4}\lambda^2\right)^{1/2}}\right], \quad (288)$$

[365] Note that m in the three-man paper (BHJ, 1926, p. 604, Eqs. (31)–(33)) has to be replaced by $m_k=m-m_s$.

[366] For example, Heisenberg and Jordan calculated the submatrix of the transformation matrix S, by means of which the submatrix of (H_1+H_2), associated with a given set of quantum numbers k, m and $s=\frac{1}{2}$, was diagonalized. They then obtained for the matrix elements the result (see Heisenberg and Jordan, 1926, p. 271, Eqs. (26))

$$S_{-1/2,\,-1/2}=C_{-1/2}\frac{\lambda}{2}\left[k(k+1)-\left(m^2-\frac{1}{4}\right)\right]^{1/4},$$

$$S_{-1/2,\,+1/2}=C_{+1/2}\frac{\lambda}{2}\left[k(k+1)-\left(m^2-\frac{1}{4}\right)\right]^{1/2},$$

$$S_{+1/2,\,-1/2}=\frac{1}{2}C_{-1/2}\left\{\mu+\lambda m-\left[\mu^2+2\lambda\mu m+\lambda^2\left(k+\frac{1}{2}\right)^2\right]^{1/2}\right\},$$

$$S_{+1/2,\,+1/2}=\frac{1}{2}C_{+1/2}\left\{\mu+\lambda m+\left[\mu^2+2\lambda\mu m+\lambda^2\left(k+\frac{1}{2}\right)^2\right]^{1/2}\right\},$$

where the coefficients $C_{-1/2}$ and $C_{+1/2}$ depend only on k and m. (Evidently, the suffixes $\pm\frac{1}{2}$ denote the matrix indices associated with the values of the spin quantum numbers m_s, i.e., the submatrix S is written as $(S_{m_s,\,m_s})$.)

where $A(1)$ is a factor characteristic of the state with angular momentum $k = 1$. The intensity of the corresponding transition line (parallel to the external field **H**) is proportional to expression (288), which again agreed completely with Sommerfeld's earlier result (Sommerfeld, 1922a, p. 266).

Heisenberg and Jordan's matrix calculation, in the case of $s = \frac{1}{2}$, accounted perfectly for the observations of the anomalous Zeeman effect in alkali doublets, and all the expressions obtained were equivalent to Sommerfeld's successful formulae of 1922. The new theory described equally well the situation for the spectra of alkaline earths. The latter have series of singlet and triplet spectral lines when no external magnetic field is present, which may be readily explained by assuming that the spin angular momenta of the two outer electrons (responsible for the radiation) either subtract from each other to give $s = 0$ or add up to give $s = 1$.[367] It is not difficult to show that the application of the matrix mechanical perturbation scheme to the triplet system leads, in general, to a cubic characteristic equation for the eigenvalues of W_m, the magnetic perturbation energy; that is, each set of states denoted by the quantum numbers, k, m and s (where s assumes the value 1), is normally split by the perturbation into three states according to the three different values of m_s $(= -1, 0, +1)$. Similar considerations apply to the higher multiplets. In the case of the singlet, however, one may disregard the spin altogether.

Heisenberg was also happy to derive another result, a 'summation principle' ('*Summationsprinzip*'), from the equations of perturbation theory. Let m_1 be the smallest and m_2 the largest value of m_s admitted for a state described by the set of quantum numbers, k, s and m, then it follows from the corresponding characteristic equation that the sum of the energy terms is given by the equation[368]

$$\sum_{m_s = m_1}^{m_2} W_{m_s} = \sum_{m_s = m_1}^{m_2} \left[\mu(m_s + m_1) + \lambda m_s(m - m_s) \right]. \tag{289}$$

The right-hand side of this equation depends linearly on **H**, the field strength of the external field. Heisenberg had already arrived at this 'summation principle' in the special case of doublets and triplets in his very first treatment of the anomalous Zeeman effects (Heisenberg, 1922a).

Heisenberg and Jordan drew further conclusions applicable to all multiplets, that is, independent of the specific system under investigation, by considering perturbation Hamiltonians, Eq. (283), in cases when either $\lambda \ll \mu$ or $\lambda \gg \mu$. In the

[367] Evidently, the model assumed to describe the empirical situation is one in which the two spin vectors, $\mathbf{s}_1$ and $\mathbf{s}_2$, add up to $\mathbf{s}_1 + \mathbf{s}_2$, and the two orbital angular momenta, $\mathbf{k}_1$ and $\mathbf{k}_2$ add up to $\mathbf{k}_1 + \mathbf{k}_2$; thus, the total angular momentum **M** is obtained by adding $(\mathbf{s}_1 + \mathbf{s}_2)$ to $(\mathbf{k}_1 + \mathbf{k}_2)$. The case of the so-called $(j - j)$-coupling (where spin and angular momentum of each electron contribute to the momenta, i.e., $\mathbf{M}_1 = \mathbf{s}_1 + \mathbf{k}_1$ and $\mathbf{M}_2 = \mathbf{s}_2 + \mathbf{k}_2$, and $\mathbf{M}_1$ and $\mathbf{M}_2$ have finally to be added up in order to give the total angular momentum of the two-electron system) has to treated differently.

[368] We have corrected an obvious error in Eq. (17) of Heisenberg and Jordan (1926, p. 269); in the second term on the right-hand side of this equation m_1 must be replaced by m_s.

first case, the external magnetic field exerts a much stronger influence on the spectral lines than the internal field and the perturbation matrix $(H_1 + H_2)$ becomes diagonal, up to the order λ^2, in the quantum number m_s. Since, in this case, there are no transitions with $\Delta m_s \neq 0$, each spectral line will only split into a (normal) Zeeman triplet under the action of a magnetic field because of the change $\Delta m_k = 0, \pm 1$ of the z-component of the orbital angular momentum, in agreement with the Paschen–Back effect.

The other extreme situation, i.e., the case $\lambda \gg \mu$, occurs when the splitting due to the external magnetic field **H** is small compared with the natural line splitting (i.e., the distances between the components of a multiplet). When **H** is zero, the perturbation is given only by the term H_2 in Eq. (283), which may be rewritten as $\frac{1}{2}\lambda(2\pi/h)^2[\mathbf{M}^2 - \mathbf{k}^2 - \mathbf{s}^2]$, where $\mathbf{M} = (\mathbf{k} + \mathbf{s})$. Hence H_2 becomes a diagonal matrix in j, the quantum number of the total angular momentum **M**, having the elements

$$H_2(j, k, s; j, k, s) = \tfrac{1}{2}\lambda\big[\, j(j+1) - k(k+1) - s(s+1)\big]. \tag{290a}$$

In the language of a classical mechanical model, the direction of the total angular momentum **M** serves as the axis of precession. When the magnetic field is switched on, it gives rise to separate precessions of the orbital angular momentum **k** and the spin angular momentum **s**; however, their components perpendicular to the axis of precession (the direction of **M**) do not contribute to $\overline{H}_1$, the average of the (classical) perturbation energy over a complete orbit. This argument could be taken over into quantum theory, as the matrices in question commute in the approximation $\mu \ll \lambda$. Hence $\overline{H}_1$ was also given there by the components parallel to **M**, that is,

$$\overline{H}_1 = \left(\frac{2\pi}{h}\right)\mu\frac{(\mathbf{H}\cdot\mathbf{M})}{|\mathbf{H}|}\left[1 + \frac{(\mathbf{M}\cdot\mathbf{s})}{\mathbf{M}^2}\right]. \tag{290b}$$

The matrix elements of the total perturbation energy can be obtained by adding H_2 and $\overline{H}_1$ as

$$\big(\overline{H}_1 + H_2\big)(j,k,s;\, j,k,s) = \mu m\left[1 + \frac{j(j+1) - k(k+1) + s(s+1)}{2j(j+1)}\right]$$
$$+ \frac{1}{2}\lambda\big[\, j(j+1) - k(k+1) - s(s+1)\big] \tag{291}$$

for $\mu \ll \lambda = \lambda(k)$. The factor within square brackets in the first term on the right-hand side of Eq. (291) is identical with the well-known g-factor, while the second term expresses the 'interval rule,' both discovered by Alfred Landé in 1923 from an analysis of Back's data on the anomalous Zeeman effects (Landé, 1923a).

The final section (Section 5) of Heisenberg and Jordan's paper was devoted to a calculation of the fine structure in the case of the one-electron atom, where $s = \frac{1}{2}$. The authors evaluated the average values of the inverse powers of the radial matrices r carefully and found for $H_2 + H_3$, the sum of the internal field and the fine-structure field terms, the result

$$(H_2 + H_3)\,(n, j; n, j) = \frac{2R^2h^2Z^4}{m_e c^2 n^3}\left[\frac{3}{4n} - \frac{1}{j + \frac{1}{2}}\right], \tag{292}$$

for an atom with atomic number Z in a state characterized by the principal quantum number n and the total angular momentum quantum number j (also R = Rydberg constant $= 2\pi^2 e^4 m_e / h^3$).[369] For hydrogen, one has only to put $Z = 1$ and Sommerfeld's well-known fine-structure formula follows.

From Eq. (292) it also becomes evident that a difficulty, which had been noted by Heisenberg and Pauli during the early stages of their discussion on fine structure, disappeared: this was the fact that the expression for the magnetic energy seemed to become infinite because of the factor $\overline{r^{-3}}$. (See the second Eq. (281) with the denominator $p_\phi^2 - \frac{1}{4}(h/2\pi)^2$ being equal to $k(k+1)$.) The reason is that in this case the other factors in H_2 cancel this infinity; for zero angular momentum there is no magnetic energy at all.

Equation (292), though derived strictly for only one-electron atoms, could be applied approximately also to non-hydrogen-like atoms, for example to alkali atoms, where the spectral lines are emitted by the valence electron in the outermost shell. Then, of course, the atomic number Z had to be replaced by an effective number Z_{eff}, which is smaller than the atomic number but larger than unity.[370] Therefore, the separation of the alkali doublets is much larger than the splitting of the hydrogen lines. Furthermore, Eq. (292) accounted for the X-ray doublets, for, as Heisenberg and Jordan noted: 'In particular, it follows from the absence of k in Eq. [(292)] that the "screening doublets" can be explained by the hypothesis of Compton, Uhlenbeck and Goudsmit [i.e., the hypothesis of rotating

[369] By inserting the expression for $\overline{r^{-1}}$ $(= 1/e^2 \cdot (2RhZ/n^2))$ (from Pauli's theory of hydrogen atom and for $\overline{r^{-2}}$ $(= (m((h/2\pi)(k+\frac{1}{2}))^{-1}) \cdot (4\pi RZ^2/n^3))$ from Eq. (281) into Eq. (279c), and adding it to Eq. (279b′), the result followed that

$$(H_2 + H_3)(j, k, s;\, j, k, s) = \frac{2Rh^2Z^4}{m_e c^2 n^3}\left[\frac{j(j+1) - k(k+1) - s(s+1)}{2k(k+\frac{1}{2})(k+1)} - \frac{1}{k+\frac{1}{2}} + \frac{3}{4n}\right].$$

For doublets having $s = \frac{1}{2}$ and $j = k \pm \frac{1}{2}$ this reduced to Eq. (292). In this equation we have replaced the variables, j, k and s, on the left-hand side by n and j, because s is fixed and k does not enter on the right-hand side anymore.

[370] The actual calculation of Z_{eff} is rather complicated. However, from a classical consideration, one can see that Z_{eff} is larger than unity. Assume that the orbits of the valence electron move partly outside the core of the alkali atom and partly inside. Then one has to replace the factor $(\overline{Z/r^3})$ in Eq. (279b′) by a sum, i.e., $(\overline{Z/r^3})_{\text{ext}} + (\overline{Z/r^3})_{\text{int}}$ where the subscripts 'ext' and 'int' refer to the orbits outside and within the core, respectively. Evidently, the internal term is much larger than the external term, in which $Z = 1$; hence Z_{eff} is larger than 1.

electrons]' (Heisenberg and Jordan, 1926, p. 277). In spite of Gregor Wentzel's arguments against this conclusion (on the basis of a classical calculation), Heisenberg remained convinced that the atomic model including spin would allow one to treat the problems of complex structure consistently.[371]

The methods of matrix mechanics thus turned out to be very successful in dealing with the multiplet structure and the anomalous Zeeman effects. Especially, Landé's empirical rules followed automatically. The question remained, however, whether also certain violations of Landé's rules, which had been observed in connection with the so-called partial Paschen–Back effect (see Back, 1925b), could be explained in the new theory. The partial Paschen–Back effect occurred if the magnetic field strength was large enough so that the condition $\mu \gg \lambda = \lambda(k)$ was satisfied for one state (the one with the higher angular momentum quantum number k) and not for the other state involved in the transition. In that case certain Zeeman lines of the magnetic triplet arising from the combination of P and D terms had not been found. Lucie Mensing then took up the problem. By calculating the transformation matrix for the triplets according to the procedure of Heisenberg and Jordan, she obtained the result that the element for the transition from the state with quantum numbers $j = k$ and $m = 0$ to the state with quantum numbers $m_s = 0$ and $m = 0, \pm 1$ was zero (Mensing, 1926b). Hence, according to matrix mechanics, the corresponding transition lines would not occur, as was confirmed by Back's data. For doublet terms, on the other

[371] In his analysis of X-ray doublets, Wentzel had previously separated two types of doublets: the 'relativistic' doublets whose separation was of relativistic origin but proportional to Z^4 and screening doublets whose separation was also of relativistic origin but proportional to Z_{eff}^4, due to the 'screening' of the charge of the nucleus by inner electrons (see Sommerfeld and Wentzel, 1921). In spring 1926 Wentzel doubted that an expression of the type, $(4\pi^2 e^3 R/h^2 c^2 n^3 k(k-1))Z_{\text{eff}}^4$, would describe the separation of screening doublets, which could be phenomenologically described by the formula

$$\Delta\nu = R\left[\frac{1}{n^2 k(k-1)} \cdot \alpha(Z - s_0)^4 + \Gamma_1(n,k)\alpha^2(Z - s_1)^6 + \Gamma_2(n,k)\alpha^3(Z - s_2)^8 + \ldots\right],$$

where $s_0, s_1, s_2 \ldots$ were the screening constants and $\Gamma_1(n,k), \Gamma_2(n,k)$, etc., the coefficients depending on the quantum numbers n and k, and $\alpha\,(= 2\pi e^2/hc)$ denoted the fine-structure constant. In March 1926 he submitted a paper in which he proposed to calculate s_0 in the first term in the classical manner, obtaining values which were considerably larger than the observed ones (Wentzel, 1926b). From his calculations he concluded: 'It is difficult to find a plausible reason to explain the discrepancy. Perhaps the potential law for the spherical shells [which he had used to describe the screening effect to the order α] is not suitable in quantum mechanics to describe the interaction between the electron and the shell in first approximation? Or perhaps the model of the top-like electron [spinning electron] is still not sufficient to represent all the fine details of spectroscopic phenomena?' (Wentzel, 1926b, p. 914).

Wentzel apparently favoured the second possibility. Heisenberg, on the other hand, did not trust Wentzel's calculations, and he wrote to Jordan: 'I find Wentzel's paper on screening horrible; it would have been a wonder if anything correct had resulted from it. I do not at all believe in physical difficulties, say with the S-terms [in which Wentzel had found the largest discrepancies between the theoretical and the empirical screening constants s_0]. On the other hand, the mathematical calculation is not easy' (Heisenberg to Jordan, 28 July 1926). Pauli also shared Heisenberg's point of view; he argued, in particular, that Wentzel's occasional use of half-integral quantum numbers for the angular momentum could not be justified on the basis of quantum mechanics including the spin hypothesis. (See, Pauli to Wentzel, 5 July 1926.)

hand, Mensing's calculation showed that the transition elements were always nonzero, and the experimental data did not indeed exhibit any anomalies in this case. Moreover, with respect to the nonzero intensities of the lines, experiments and the quantum-mechanical results agreed completely. Lucie Mensing, the young lady who proved herself capable of performing the most delicate theoretical calculations of anomalous Zeeman effects, was offered a post-doctoral research position at Tübingen where Back and Landé were the leading lights of atomic spectroscopy.[372]

V.6 Key to the Helium Problem

It might seem rather surprising that Pauli, who had been concerned so much with the programme of extending the matrix methods to include action-angle variables and with the problem of explaining the fine structure of hydrogen, was not at least co-author of one of the two papers, i.e., Wentzel's paper on the multiply periodic systems (Wentzel, 1926a) or Heisenberg and Jordan's paper on the complex structure and anomalous Zeeman effects (Heisenberg and Jordan, 1926). Being unhappy with the course of events, he just stayed away from any publication. Indeed, in March 1926 Pauli had several reasons to be dissatisfied with the progress of his scientific work. During the previous few months he had wasted much time on arguments against the electron spin hypothesis until he finally withdrew, in his letter to Niels Bohr (dated 12 March 1926), his objections against the relativistic calculations of L. H. Thomas. Thus, an important conceptual aspect of the theory of atomic structure, to which Pauli had devoted himself since fall 1922, had been established by a graduate student from Cambridge, England, in spite of his authoritarian criticism. At the same time, Paul Dirac, another Cambridge research student, had developed the quantum-mechanical scheme of action-angle variables much faster than Pauli, although a number of the fundamental ideas from which both of them departed had been the same. The explanation of this failure of such a seasoned theoretical physicist as Pauli to compete with practically unknown graduate students like Thomas and Dirac may be sought to a large extent in Pauli's attitude towards atomic physics in general and the priorities he had set up for himself in approaching specific problems.

Early in 1926 Pauli, more than anybody else working on quantum mechanics, still believed the theory to be only in a preliminary stage of development. Therefore, he still adhered to the cautious procedure he had learned from Niels Bohr several years earlier when the previous atomic theory had run into serious contradictions with certain empirical facts: namely, to look with suspicion at every result obtained from theory and not accept it until it had been fully established by agreement with experiment. It was in this spirit that he had

[372] Concerning this invitation Pauli wrote to Landé: 'I heard from Miss Mensing that you intend to get her as an assistant. I am very glad about it; it will be very good for her' (Pauli to Landé, 2 June 1926).

attacked, in October 1925, the problem of calculating the energy states of the hydrogen atom with the help of matrix mechanics. Afterwards, since he was convinced of the correctness of Sommerfeld's fine-structure formula of 1916, he gave priority to extending matrix methods in such a way that the energy states of the relativistic hydrogen atom could be obtained.[373] When the puzzle with the factor $\frac{1}{2}$ had been solved in favour of Thomas' calculation, the quantum-mechanical derivation of the fine-structure formula became trivial and Pauli left it to Heisenberg and Jordan to publish it in their paper on the anomalous Zeeman effects (Heisenberg and Jordan, 1926). Another theoretical problem of this kind—i.e., one in which the results could be tested by comparison with the empirical data—was to calculate the intensities of the hydrogen spectral lines, and Pauli had thought about it in particular since December 1925.[374] In the beginning of February 1926 this problem still concerned him, for he wrote to Niels Bohr: 'At the moment I am occupied with carrying further the solutions of the matrix equations for the hydrogen atom in order also to arrive at the final formulae for the intensities' (Pauli to Bohr, 5 February 1926). During March and April he continued his attempt but, even though he made use of the method of action-angle variables, he did not achieve his goal.[375]

Pauli had been anxious to exploit matrix methods as far as possible, rather than make use of a new formalism which might be more adequate to deal with the intensity problem. Thus, he had not really examined the possibilities that were offered in this respect by the operator scheme of Born and Wiener. Erwin Schrödinger's first paper on wave mechanics (Schrödinger, 1926c) took him, therefore, by surprise, for he had not expected that the operator methods would lead to anything useful in practice. The set of functions, which Born and Wiener had associated with the quantum-mechanical system under consideration (Born and Wiener, 1926, §2), and which Pauli had scrupulously avoided in his formalism dealing with angle variables, now reappeared triumphantly as Schrödinger's eigenfunctions. However, Pauli made use of the new approach very soon; after proving the equivalence of the wave and matrix methods, he was 'occupied with the problem of computing from the eigenfunctions—calculated by Schrödinger—the transition probabilities for hydrogen' (Pauli to Jordan, 12 April 1926). Though he obtained certain results on the intensities of Balmer lines, he did not publish them immediately, but rather waited for an experimental verification, in

[373] How important Pauli considered the correct explanation of the hydrogen fine structure may be gleaned from remarks in a letter to Sommerfeld, dated 9 February 1926. There he remarked: 'In any case, I completely share the point of view which you adopted at the end of your paper [Sommerfeld and Unsöld, 1926], namely, no matter whether one tries to interpret the energy in the relativistic, or in the magnetic, or in any other model-dependent sense, one must retain your 1916 formula for the fine-structure levels quantitatively' (Pauli to Sommerfeld, 9 February 1926).

[374] That Pauli had begun to work seriously on the problem of calculating the intensities of spectral lines becomes evident from a letter of Heisenberg's. In it Heisenberg wished Pauli 'a beautiful Christmas and a good New Year, filled with hydrogen transition probabilities, theory of helium, etc.' (Heisenberg to Pauli, 24 December 1925).

[375] In March 1926 Pauli wrote again to Bohr; 'Meanwhile I have further laboured on the intensities of Balmer lines, and still hope to be able to master this problem' (Pauli to Bohr, 5 March 1926).

spite of the fact that he knew that Schrödinger was calculating similar things and he had to hurry.[376] He did not have to worry long about the question of publication. Schrödinger had included some preliminary results of his own calculation on the intensities of Balmer lines in a communication which he submitted to *Annalen der Physik* in the beginning of May (Schrödinger, 1926f, p. 476); Pauli informed Schrödinger about his results, and Schrödinger, in a postscript added in proof to his paper, gave the full intensity formulae of the hydrogen lines obtained by Pauli (Schrödinger, 1926f, pp. 489–490).

The continuing concern with the intensities of the hydrogen lines prevented Pauli from working on other important problems of the day. After the question of electron spin had been settled, and the formulae describing the fine structure of the hydrogen lines and the empirically observed anomalous Zeeman effects had been derived from a systematic application of the quantum-mechanical scheme, one of the few remaining questions which concerned the complex structure of spectral lines was the problem of explaining the helium spectra. Pauli had hoped that Bohr and Goudsmit, who had been collaborating together in Copenhagen earlier in 1926, would find a solution.[377] However, no progress came out of these efforts. As Goudsmit recalled later on about his discussions with Bohr:

> That was a complete fiasco. Bohr knew what the next problem was, [but] I did not understand it; it was not [up] my alley. By looking at the helium spectrum Bohr understood right away that [if] you turned over one of the two spins [of the electrons in the helium atom] the energy suddenly would be entirely different. We got together every day in the hope that his talking to me would put some idea in my mind which would solve the problem. It didn't. So after five or six weeks he gave me a first class ticket back to Holland on a sleeper . . . with the admonition that I had to get off at Hamburg and tell Pauli about the Thomas factor 2. (Goudsmit, AHQP Interview, 7 December 1963, p. 25)

[376] Thus, Pauli wrote to Landé; 'Recently I have derived—based on the connection between Schrödinger's [work] and that of Göttingen—formulae for the absolute strengths of the Balmer lines. Ladenburg shall check them with the help of dispersion measurements. Since, however, Schrödinger is calculating similar things, I do not yet know where, how and when I shall publish the results' (Pauli to Landé, 2 June 1926).

Since 1925 Rudolf Ladenburg had studied, together with his collaborators Hans Kopfermann and Agathe Carst at the *Physikalische Reichsanstalt* in Berlin, the anomalous dispersion formula. (See Ladenburg's discussion remark in Meissner, 1925b, p. 689.) He had reported about his results on various occasions, first on 14 February 1926 at the Lower Saxony Section (*Gauverein Niedersachsen*) of the German Physical Society in Braunschweig, where Pauli had been present (see *Verh. d. Deutsch. Phys. Ges.* (3) **7**, 1926, p. 13). Later in 1926 he obtained preliminary results on the intensity ratio of the two strongest lines in the Balmer spectrum; this value seemed to confirm the predictions of Schrödinger and Pauli (Ladenburg, 1926b, p. 1213). However, the experimental errors were still considerable, and it took another year until more accurate data were obtained.

[377] Samuel Goudsmit went to Copenhagen in February 1926 to work with Bohr on spectroscopic questions arising from the incorporation of electron spin in atomic theory. 'Goudsmit was here during the past 14 days,' Bohr wrote to Pauli on 20 February 1926, 'and he and I have discussed in detail the question of atomic structure and of spectral laws . . . Goudsmit and I hope to write something together in the near future on our views concerning these matters, and I shall report to you in detail about it' (Bohr to Pauli, 20 February 1926).

In Hamburg, Goudsmit did not succeed in persuading Pauli of the correctness of Thomas' calculation.[378] Moreover, Pauli was 'not very pleased to hear from Goudsmit that he had not made the slightest progress in the two most important problems [of spectroscopy], which are still unsolved; namely, the calculation of the separation of singlet and triplet terms (e.g., ortho- and parhelium) and the physical foundation of the formal exclusion rules for equivalent orbits' (Pauli to Kramers, 8 March 1926). As Goudsmit recalled, he was out of his depth in dealing with these questions:

> I invented all kinds of magnetic interactions [to explain the large energy differences between the corresponding singlet and triplet terms of the helium atom, i.e., terms having identical quantum numbers n and k but different spin quantum number s], and I have a letter from Bohr in which he tells me that I am all wet. But still I wrote [to Bohr] about it later; I kept on thinking about it. Then he called [upon] Heisenberg who indeed found the solution—the antisymmetric wave functions and so on. That was *way beyond* me. I would never have been able to guess anything like it. (Goudsmit, AHQP Interview, 7 December 1963, p. 25)[379]

From mid-November to mid-March 1926 Heisenberg did not submit any paper for publication that dealt with new methods in or applications of matrix mechanics. This, for Heisenberg, was an unusually long time, and he later tried to explain the reason for his silence as follows: 'It was partly due to the absence of Born [who was in America], partly due to the teaching which I had to do, and partly also a kind of exhaustion. I felt, "Well, now let's wait for a time"' (Heisenberg, Conversations, p. 302).[380] Though Heisenberg's work with Jordan on the treatment of anomalous Zeeman effects had been in progress for some time, he did not wait around for the question of spin to be settled in favour of Uhlenbeck and Goudsmit by Thomas' calculation; as soon as his teaching obligations for the winter semester 1925–1926 were over, Heisenberg left Göt-

[378] Goudsmit passed through Hamburg on 8 March 1926. (See Pauli to Kramers, 8 March 1926, quoted in Section V.5.)

[379] In his letter to Bohr, dated 12 April 1926, Goudsmit proposed a strong mutual coupling of the spins of the two electrons arising from the quantization of the sum of the two spin vectors, $\mathbf{s}_1$ and $\mathbf{s}_2$ (which Goudsmit called R_1 and R_2). Bohr, in his reply, stated: 'I must point out, however, that it is very difficult to see in your considerations a sufficient mechanical justification' (Bohr to Goudsmit, 17 April 1926). He claimed especially that there had to be a correspondence between classical and quantum-theoretical forces; hence a strong spin–spin force had somehow to be understood on the classical level. Bohr added that Kramers and Pauli, who were both in Copenhagen at the time, shared his opinion.

[380] The three-man paper (BHJ, 1926) was received by *Zeitschrift für Physik* on 16 November 1925. After that, on 21 December 1925, Heisenberg submitted a review article to *Mathematische Annalen*, dealing with the foundations of quantum mechanics (which he had presented in lectures in Hilbert's mathematical seminar) (Heisenberg, 1926a). In the winter semester 1925–1926 he gave the main lecture course on the kinetic theory of matter, replacing Max Born who was (in America) on leave of absence. During that time the debate on the question of electron spin delayed the publication of Heisenberg and Jordan's paper on the anomalous Zeeman effects (Heisenberg and Jordan, 1926). Pascual Jordan also did not publish any paper during the winter semester because he was completely absorbed in writing the book (with James Franck) on the excitation of atoms by collisions (Franck and Jordan, 1926).

tingen for a vacation extending to March and most of April 1926.[381] He needed some relaxation, for he was looking forward to going to Copenhagen afterwards to replace Hendrik Kramers as a member of the Institute of Theoretical Atomic Physics—Bohr's Institute—and lecturer at the University. Having turned down the offer of an associate professorship [or *Extraordinariat*] at the University of Leipzig, he arrived in Copenhagen towards the end of April, refreshed in body and spirit and full of new ideas and ambition.[382]

In Copenhagen, Heisenberg plunged immediately into one of the outstanding problems being discussed there: the calculation of the energy states of the helium atom. No progress had been made in it since about the middle of March when Pauli had withdrawn his serious objections against the electron spin hypothesis and written to Bohr: 'Now I only hope that my gibes about the "new heresy" ("*neue Irrlehre*") [i.e., the electron spin hypothesis] would have the result that your zeal would be so incited that you would be able to finish off soon the remaining points of the programme (helium spectrum, separation of the singlet–triplet system, equivalent orbits)!' (Pauli to Bohr, 12 March 1926).[383] Heisenberg took his assignment very seriously and, within a few days of his arrival in Copenhagen, reported to Pauli about his initial progress:

> Then I want to write to you that we have found a rather decisive argument that your exclusion of equivalent orbits [i.e., of the electron orbits in helium denoted by the same values of the four quantum numbers n, j, k and m_s] is connected with the singlet–triplet separation. Consider the energy written as a function of the transition probabilities. Then a large difference results, if one—at the energy of H-atoms—has transitions to $1S$, or if, according to your ban [exclusion principle], one puts them equal to zero. That is, par- and ortho-[helium] do have different energies, independently of the interaction between the magnets. (Heisenberg to Pauli, 5 May 1926)

[381] During the first part of his semester vacation Heisenberg did not occupy himself with physics at all. On returning to Munich, he wrote to Jordan on 4 April that he was starting to do physics again. He then learned from Jordan that the spin question had been decided favourably in the meantime, as Thomas' calculation had been found to be correct. Bohr, in a letter dated 20 February 1926, had already informed Heisenberg about Thomas' calculation; Heisenberg, however, had not been able to understand Thomas' note (Thomas, 1926), a copy of which Bohr had sent him. (See Heisenberg to Bohr, 26 February 1926: 'Naturally, it will be a great success if one manages to conquer the factor 2. I must confess, however, that—with all good will—I have not succeeded until now in understanding the note.')

[382] In a postcard to Jordan, dated 9 April 1926, Heisenberg mentioned that he would arrive in Göttingen on the evening of 20 April, and leave on 24 or 25 April for Copenhagen. Before he could do so, however, he received the offer from Leipzig and travelled, therefore, from Göttingen to Leipzig and Dresden in order to discuss matters with the Leipzig faculty and the Saxonian Ministry of Education on 25 and 26 April. (See the letter of Richard Courant to Niels Bohr, 24 April 1926.) The position at Leipzig, which had been occupied by George Jaffé, was later offered to Pauli and finally to Gregor Wentzel who accepted.

[383] After returning from Copenhagen to Hamburg, Pauli wrote to Wentzel in the beginning of May about the state of affairs: 'In the case of spectra of the alkaline-earth type, there are still unexplained difficulties: one does not see how the separation between the singlet and triplet terms, which is many times greater than the triplet separation, arises [in the mechanical model]; moreover, one does not yet have an interpretation for my "housing office for equivalent orbits" ("*Wohnungsamt für äquivalente Bahnen*" [i.e., the exclusion principle]). (Both problems are perhaps connected.)' (Pauli to Wentzel, 8 May 1925).

And, just three weeks later, Heisenberg could write to Max Born who had meanwhile returned from the United States: 'Since coming here I have very seriously worked on [the problem of] the helium spectrum, and I believe that now all the essential things are all right; however, I have not yet completed the quantitative calculation' (Heisenberg to Born, 26 May 1926).

There were several reasons for Heisenberg's rapid progress in dealing with the problem of the helium atom in quantum mechanics. First, he was evidently more qualified to work on quantum mechanics than both Goudsmit and Bohr who had tackled this problem earlier. Goudsmit's abilities, in any case, lay in the analysis of empirical data, while Bohr did not have enough experience in applying the complicated formalism of matrix mechanics, without which no progress could be made. Second, Heisenberg had had a long personal association with the helium problem ever since he had attempted, in fall 1922, to account for the observed energy states of helium by using half-quantum numbers in the theory of multiply periodic systems. While the half-quantum numbers had been discarded in 1923 as being in contradiction with the principles of the then quantum theory (see, for instance, Born and Heisenberg, 1923b, p. 243), they had returned in fall 1925 in the guise of electron spin. After Heisenberg had accepted the spin hypothesis he could also reasonably expect that a repetition of his old, unpublished calculations within the scheme of matrix mechanics would yield results in agreement with the observed data.[384]

However, before being able to calculate the energy states of helium, Heisenberg had to handle two difficulties which he stated in a letter to Max Born: '1. Because of the equivalent orbits and Pauli's ban [exclusion principle], quantum mechanics did not seem to be applicable to the problem of several-electron atoms. 2. The large separation between the singlet and triplet systems could not be explained by a magnetic interaction of the rotating electrons; one had therefore to assume with Hund a large force of unknown origin between the magnets' (Heisenberg to Born, 26 May 1926).[385]

[384] Heisenberg had reported about his calculation of the helium atom with half-integral quantum numbers in letters to his teacher Arnold Sommerfeld during winter 1922–1923 (Heisenberg to Sommerfeld, 28 October, 4 December 1922, 4 and 15 January 1923). At that time he had tried to obtain the ionization potential of helium, and by assuming the angular momentum values of the electron orbits in the ground state to be $\frac{1}{2}h/2\pi$ and $-\frac{1}{2}h/2\pi$, respectively, he had obtained a result very close to the empirical value of 24.5 V. (See Volume 2, Section II.3.)

[385] A calculation of the usual (electromagnetic) interaction *between* the magnetic moments of the two electrons yielded a splitting which was smaller than the splitting *within* a given multiplet due to the interaction of the magnetic moment of one electron with the internal field $\mathbf{H}_i$, Eq. (279b′). (See, e.g., Hund, 1927, pp. 123–124.) This result had been noticed by Heisenberg soon after he learned about the electron spin hypothesis. 'I myself have very great reservations against Goudsmit's attempt to establish zoology [i.e., the spectroscopic terms] in such a simple way,' he had written to Pauli, and added: 'For the alkaline-earth atoms the energy difference between singlet and triplet terms with *equal k* must result only from the different interaction of the electron magnets upon each other. This yields a completely wrong order of magnitude' (Heisenberg to Pauli, 24 November 1925). Therefore, if one wanted to keep up the magnetic origin of the singlet–triplet splitting, one had to assume a strong, so-far-unknown interaction between the magnetic moments of the electrons in the helium atom.

Heisenberg had known both of these difficulties from before and had discussed them with Pauli since fall 1925. He attributed great importance to the first question, namely, how to express the exclusion principle in the quantum-theoretical formalism in which, at first sight, there seemed to be no place for it, as it seemed primarily to be concerned with the explanation of the anomalous Zeeman effects of many-electron atoms. Thus, on 16 November 1925, he had written to Pauli: 'I have also thought a little about the "stress" ("*Zwang*") and I still have the hope that it might arise when two (or several) electrons are coupled,' adding in parenthesis the remark, 'and in fact just on account of your exclusion of equivalent orbits' (Heisenberg to Pauli, 16 November 1925). The later calculations on the anomalous Zeeman effects (Heisenberg and Jordan, 1926) had shown, however, that the electron spin hypothesis accounted very well for the observed phenomena and there was no need to postulate an extra 'stress' ('*Zwang*'). However, the problem of explaining quantitatively the differences between corresponding terms of the singlet and triplet series in the spectra of helium and alkaline-earth atoms had remained. On the basis of a thorough analysis of the empirical data Heisenberg decided to attack simultaneously the two difficulties about which he had written to Born.

Heisenberg's discussions in Copenhagen, especially with Bohr, were first of all concerned with what could possibly be learned from the known empirical facts and whether these facts offered any hint for their theoretical treatment. 'The normally existing grouping of alkaline-earth terms in a singlet system and a triplet system,' Pauli had noted in his *Handbuch* article, ' . . . contradicts, however, the purely relativistic interpretation of the complex structure. In this grouping the separation between both term systems is large compared to the triplet splitting' (Pauli, 1926b, p. 254). The value of the energy difference between the components of the triplet had been found to be of the order of the corrections due to fine structure (Millikan and Bowen, 1925a); hence it was accounted for in matrix mechanics by the combined effect of the magnetic coupling of the electron spin and the relativistic correction, as had already been shown by Heisenberg and Jordan (1926). The singlet–triplet energy difference seemed to suggest, on the other hand, that it might arise from the interaction energy of the two electrons, for as Pauli had remarked earlier: 'Although the dependence of the separation between singlet and triplet systems on the atomic number has not yet been observed, nevertheless it seems to be natural to assume that this separation is caused by the dependence of the electric interaction energy of both valence electrons on the orientation of their orbits and will, therefore, first behave like a screening doublet' (Pauli, 1926b, p. 254). That is, the singlet–triplet energy difference had the order of magnitude of the correction due to 'screening' in Rydberg's empirical term formulae.

Another remarkable fact about the intensities of lines had been observed in the spectra of alkaline-earth atoms: whereas the transitions between terms of the same sequence, say triplets, were of strengths similar to those of the lines in the neighbouring alkali atoms, the intercombination lines were 'in general weak and

irregular in the case of real alkaline-earth elements Mg, Ca, Sr, Ba, in which the distance between singlet and triplet terms is large compared to the distance of triplet terms from each other' (Pauli, 1926b, p. 218). This situation showed up most prominently in the spectrum of helium, the purest two-electron system in nature: consistent with the fact that the ratio of the component splitting in ortho-helium to the average distance of the ortho-helium multiplet from the corresponding parhelium singlet is much smaller than in the alkaline-earth atoms, the intercombination lines in helium were found to be particularly weak.[386]

All these empirical facts suggested strongly that, in order to understand two-electron systems, one had first to establish a proper theory of helium spectra, and the most important task of this theory was to determine the energy difference between the ortho-helium and parhelium terms *from the Coulomb repulsion of the electrons*. That is, from the Coulomb repulsion of the electrons a new strong force had to arise that would explain the large singlet-triplet splitting. At this point Heisenberg was able to exploit his earlier acquaintance with the helium problem. Especially, he could refer to an idea which had been emphasized by Born and himself several years previously (Born and Heisenberg, 1923a). They had then discovered that, in atomic systems having two or more electrons, the phases between the motions of the electrons assumed fixed values when the perturbation—i.e., the electron–electron interaction—was switched on; for these systems, which they had called 'accidentally degenerate,' they had developed a proper perturbation scheme.[387] The phase relations gave rise to a certain resonance effect which had been discussed in Göttingen quite frequently in connection with a mechanical example. As Heisenberg recalled: 'There is the old classical experiment which is shown in every lecture course: two pendula being coupled by a piece of rubber. It was natural to study this system in quantum theory, and we certainly discussed it already very early in the period before quantum mechanics, especially in connection with dispersion theory' (Heisenberg, Conversations, p. 362).

In the dispersion approach, which had been used especially since Kramers' work in 1924, the motions of electrons in atoms were represented by harmonic

[386] At the time when Pauli wrote the *Handbuch* article on quantum theory—i.e., in summer 1925—only the strongest intercombination lines between the ortho-helium and parhelium states had been observed (Lyman, 1924).

It should be mentioned at this point that the triplet structure of ortho-helium lines had not yet been (experimentally) settled in 1926 when Heisenberg developed the theory of the helium spectrum. Thus, in one of the best known tables in those days the ortho-helium lines were listed as 'doublets'; however, it was also indicated that the Zeeman effect of the 'helium' doublet lines had a character different from those of the alkali-doublets (see Paschen and Götze, 1922, p. 28, especially footnote 3). This empirical situation did not prevent Pauli from writing in his *Handbuch* article: 'Nevertheless it is obvious to compare, apart from the complex structure, both term systems of helium with the two term systems of alkaline earths, since, at any rate, one is dealing with a two-electron system in the case of helium as well.' (Pauli, 1926b, p. 218)

[387] Actually, Max Born had discovered the existence of such phase relations earlier in a discussion of models of the hydrogen atom (Born, 1922).

oscillations of given frequencies. Thus, the coupling of the valence-electrons in an atom could be described by the following simplified model: two identical oscillators, each having the same mass m and frequency ν $(=\omega/2\pi)$, were coupled by an (in general small) interaction term, $m\lambda q_1 q_2$, where q_1 and q_2 denote the position variables of the oscillators 1 and 2. In classical mechanics, the system will develop a resonance and a beat frequency will interfere with the original frequency giving rise to two new, shifted frequencies. This effect may be demonstrated easily by performing a canonical transformation of the variables, the momenta p_1 and p_2 and the position coordinates q_1 and q_2. The Hamiltonian, H, of the system—i.e., $H = (1/2m)(p_1^2 + m^2\omega^2 q_1^2) + (1/2m)(p_2^2 + m^2\omega^2 q_2^2) + m\lambda q_1 q_2$—assumes the form

$$H = \frac{1}{2m}\left(p_1'^2 + m^2\omega_1'^2 q_1'^2\right) + \frac{1}{2m}\left(p_2'^2 + m^2\omega_2'^2 q_2'^2\right), \tag{293}$$

where the new position variables are given by the equations

$$q_1' = \frac{1}{\sqrt{2}}(q_1 + q_2) \quad \text{and} \quad q_2' = \frac{1}{\sqrt{2}}(q_1 - q_2), \tag{293a}$$

and the new frequencies, ν_1' and ν_2', by the equations

$$2\pi\nu_1' = \omega_1' = \sqrt{\omega^2 + \lambda} \quad \text{and} \quad 2\pi\nu_2' = \omega_2' = \sqrt{\omega^2 - \lambda}\ . \tag{293b}$$

Hence, there arise two specific motions, one with frequency ν_1 in which the two oscillators vibrate in phase, and the other with lower frequency ν_2 in which they vibrate with an opposite phase.

Heisenberg quickly recognized that the reformulation of the classical example of coupled oscillators in quantum mechanics would yield an analogous result. That is, if one considered two identical oscillators which were coupled by a small perturbation term, $m\lambda q_1 q_2$, where q_1 and q_2 are the position variables of the individual oscillators, then exactly the same coordinate transformation as in the classical theory, Eq. (293a), could be made. Again, the transformation leads to a Hamiltonian having the form given by Eq. (293), and the eigenvalues of this Hamiltonian are

$$W_{n_1' n_2'} = \left(n_1' + \tfrac{1}{2}\right)h\nu_1' + \left(n_2' + \tfrac{1}{2}\right)h\nu_2', \tag{294}$$

where n_1' and n_2' assume integral values, 0, 1, 2, . . . , etc., and ν_1' and ν_2' denote the two frequencies given by Eq. (293b).

In the three-man paper, where a universal algebraic formalism of perturbation theory applicable to arbitrary quantum-mechanical systems had been developed, Born, Heisenberg and Jordan had mentioned explicitly the fact that now 'two originally identical, non-degenerate systems (where all the frequencies of the individual system should be different) coupled by arbitrary forces,' could be treated (BHJ, 1926, p. 589). The authors had, however, not treated any example,

although the two coupled harmonic oscillators did have exactly the properties which they mentioned.[388] The algebraic perturbation theory was first used by Heisenberg and Jordan in their calculations of the anomalous Zeeman effects, where they considered an atomic system whose Hamiltonian (including the terms H_1 and H_2, Eqs. (279a) and (279b′), due to magnetic interaction) was degenerate in the z-component of the electron spin (Heisenberg and Jordan, 1926, Section 3). The perturbation calculation then yielded multiplets which were separated (for small magnetic field strengths) by magnitudes of the same order as the fine-structure terms that arose both from magnetic interactions and relativistic corrections. It was the result of this calculation which convinced Heisenberg that the interaction arising from the addition of electron spin alone to the existing mechanical description of atoms would not be sufficient to explain the differences between the terms of different sequences in the spectrum of helium. Such a large difference, he was certain, would only be caused by a perturbation term containing the Coulomb replusion between the two electrons. That was the basic idea which Heisenberg followed and tried to formulate mathematically, and it indeed provided the key to the solution of the helium problem.

Between 5 May, when he wrote to Pauli about his initial thoughts, and 26 May, when he sketched the scheme in his letter to Born, Heisenberg developed the method for calculating the energy states of the helium atom. He began by considering a system consisting of two electrons moving around a (heavy) nucleus. In zeroth approximation, the influence of the electrons on each other and effects due to spin were neglected and the Hamiltonian, H_0, was assumed to be made up of two terms, H^a and H^b, each referring to the motion of one electron (denoted by the indices a and b, respectively). He assumed, in agreement with the earlier discussions of many-electron atoms, that in this approximation the second electron just altered the effective charge of the nucleus; that is, the effective charge took on values between e and $2e$, depending upon whether the second electron moved inside or outside the orbit of the other electron. Thus, the Hamiltonians of the two electrons, i.e., H^a and H^b, were identical, and, as a result, there existed a peculiar degeneracy of the Hamiltonian H^0. The reason for this was that any eigenvalue W^0_{nm} of H^0, in which the electrons a and b were not in equivalent orbits, that is,

$$W^0_{nm} = W^a_n + W^b_m \quad \text{with} \quad W^a_n \neq W^b_m, \tag{295}$$

[388] One may naturally ask the question why Born, Heisenberg and Jordan, in fall 1925, did not illustrate the use of their general perturbation scheme by treating a few, obvious examples. 'When we wrote that paper,' Heisenberg recalled, 'our general situation was this: We saw that we had entered a new field where one could pick up new things wherever one went. Therefore, we just took what came our way but without following it further, because we realized that to follow it further would take some time and we would have to do some work' (Heisenberg, Conversations, p. 363). Surely, the application of the abstract formalism of degenerate systems, as developed by Max Born in the three-man paper (BHJ, 1926, Chapter 3), to calculate something as practical as the states of helium did not seem to be the most pressing problem. 'But it may be that we discussed it among ourselves,' Heisenberg remarked later. 'We had already in mind that when we had finished the paper [i.e., BHJ, 1926] then, sooner or later, we must try the helium problem and the problem of two equal [i.e., identical] coupled systems' (Heisenberg, Conversations, p. 364).

was twofold degenerate, for an exchange of the motions of the two electrons, indicated here by an exchange of the two quantum numbers n and m of the partial systems, did not alter W^0_{nm}.[389] This degeneracy was removed by the Coulomb repulsion of the electrons.

Heisenberg informed Born: 'The line of thought is totally linked with your calculation for two originally identical systems in '*Quantum Mechanics II*' [i.e., in BHJ, 1926, Chapter 3]. Your result was that for two similar systems *both* the states, which emerge by exchanging the electrons between each other, split into two separate terms; only for originally *equivalent* orbits does there remains *one* state' (Heisenberg to Born, 26 May 1926). In the case of the two-electron system, described by a degenerate Hamiltonian H^0 and an interaction term H^1, one obtains—with the help of Born's scheme—two equations to determine W^1, the perturbation energy, namely,

$$\left[W^1 - H^1(nm,nm) \right] x_{nm} - H^1(nm,mn) x_{mn} = 0 \tag{296a}$$

and

$$- H^1(mn,nm) x_{nm} + \left[W^1 - H^1(mn,mn) \right] x_{mn} = 0, \tag{296b}$$

which have to be satisfied simultaneously for any pair (n, m) with $n \neq m$. The $H^1(nm, nm)$, etc., are the corresponding elements of $\overline{H^1}$, the time average of the perturbation matrix H^1; the matrix $\overline{H^1}$ contains all those elements of H^1 whose transition frequencies in the unperturbed system (described by H^0) are zero. The x_{nm} and x_{mn} form the columns of the 2×2-submatrix of a transformation matrix which is supposed to diagonalize the perturbation Hamiltonian H^1.[390] Equations (296) yield, similar to the classical problem of two coupled oscillators, two solutions, which are denoted by $\bullet$ and $\times$, respectively, that is

$$\begin{aligned} W^{1\bullet}_{nm} &= H^1(nm,nm) + H^1(nm,mn), \\ x^{\bullet}_{nm} &= S_{nm,\,nm} = \frac{1}{\sqrt{2}}, \qquad x^{\bullet}_{mn} = S_{mn,\,nm} = \frac{1}{\sqrt{2}}, \end{aligned} \tag{297a}$$

and

$$\begin{aligned} W^{1\times}_{mn} &= H^1(nm,nm) - H^1(nm,mn), \\ x^{\times}_{nm} &= S_{nm,\,mn} = \frac{1}{\sqrt{2}}, \qquad x^{\times}_{mn} = S_{mn,\,mn} = -\frac{1}{\sqrt{2}}, \end{aligned} \tag{297b}$$

[389] Heisenberg had already used this kind of exchange energy in a letter to Sommerfeld, dated 28 October 1922. There he had treated a system of two elliptic orbits, each occupied by an electron (a or b), whose planes made an angle of 60°, and had argued that the exchange of two orbits (which he described as 'slipping through' ('*Hindurchschlüpfen*') of the electrons) accounted for the relativistic corrections to the helium ground state.

[390] The corresponding quantum-theoretical equation is $W^1 = S^{-1}\overline{H^1}S$.

where use has been made of the symmetry of the matrix $\overline{H^1}$ with respect to the electrons (i.e., $H^1(nm, nm) = H^1(mn, mn)$ and $H^1(nm, mn) = H^1(mn, nm)$, respectively).

Heisenberg did not write down the above equations in his letter to Born because he assumed that Born would immediately make the simple calculation himself, but he pointed out their consequence:[391] 'For all following considerations the decisive result now is this: The term system can be divided into two separate parts (● and ×), such that transitions occur only within the ● system or within the × system, but never from the ● system to the × system.' And he drew a figure to illustrate this (see Fig. 2). Heisenberg then explained: 'The transitions from [the states] ● to [the states] × vanish not only in the dipole radiation, but also in the quadrupole radiation and in all higher approximations, provided only that the two [unperturbed] systems are really totally equivalent, as for instance two electrons revolving around *one* nucleus. One can also never induce transitions by collisions from one system to the other. You will be able to verify this easily' (Heisenberg to Born, 26 May 1926).

Heisenberg found the results of these simple considerations so important that, about two weeks after he had written the letter to Born, he submitted a paper entitled. *'Mehrkörperproblem und Resonanz in der Quantenmechanik'* ('Many-Body Problem and Resonance in Quantum Mechanics') to *Zeitschrift für Physik* (Heisenberg, 1926b), in which he gave a full discussion of the points he had mentioned to Born. Especially, he proved there the statement that no transitions occur from the terms denoted by ● to terms denoted by × in the following way. The radiation amplitudes, $f_{n_1m_1, n_2m_2}$, of the unperturbed system are symmetric with respect to an exchange of the electrons a and b (i.e., $f_{n_1m_1, n_2m_2} = f_{m_1n_1, m_2n_2}$). Heisenberg then computed the special amplitudes of the perturbed system with the help of the tranformation matrix S, whose elements were given by Eqs. (297a) and (297b).[392] Thus, for example, in the case of the specific transition between two states of the term sequence denoted by ● (as in Fig. 2), say (n_1m) and (n_2m) with $n_1, n_2 < m$, he obtained the transition amplitude,

$$f'_{n_1m, n_2m} = f_{n_1m, n_2m} + f_{n_1m, mn_2}. \tag{298a}$$

The right-hand side of Eq. (298a) is different from zero unless the two unperturbed amplitudes vanish. On the other hand, if he took transitions between states belonging to different term sequences, say (n_1m) and (mn_2) with $n_1, n_2 < m$, where the first one belongs to the sequence denoted by ● and the second to the sequence denoted by ×, he found the amplitude to be

$$f'_{n_1m, mn_2} = f_{n_1m, n_2m} - f_{mn_1, mn_2}, \tag{298b}$$

[391] Equations (296a, b) and (297a, b) can be found, however, in Heisenberg's paper on the many-body problem and resonance in quantum mechanics, which grew out of the general considerations on coupled systems. (See Heisenberg, 1926b, p. 418, Eqs. (10), (11) and (12).)

[392] The elements, other than the ones given by Eqs. (297a) and (297b), are zero.

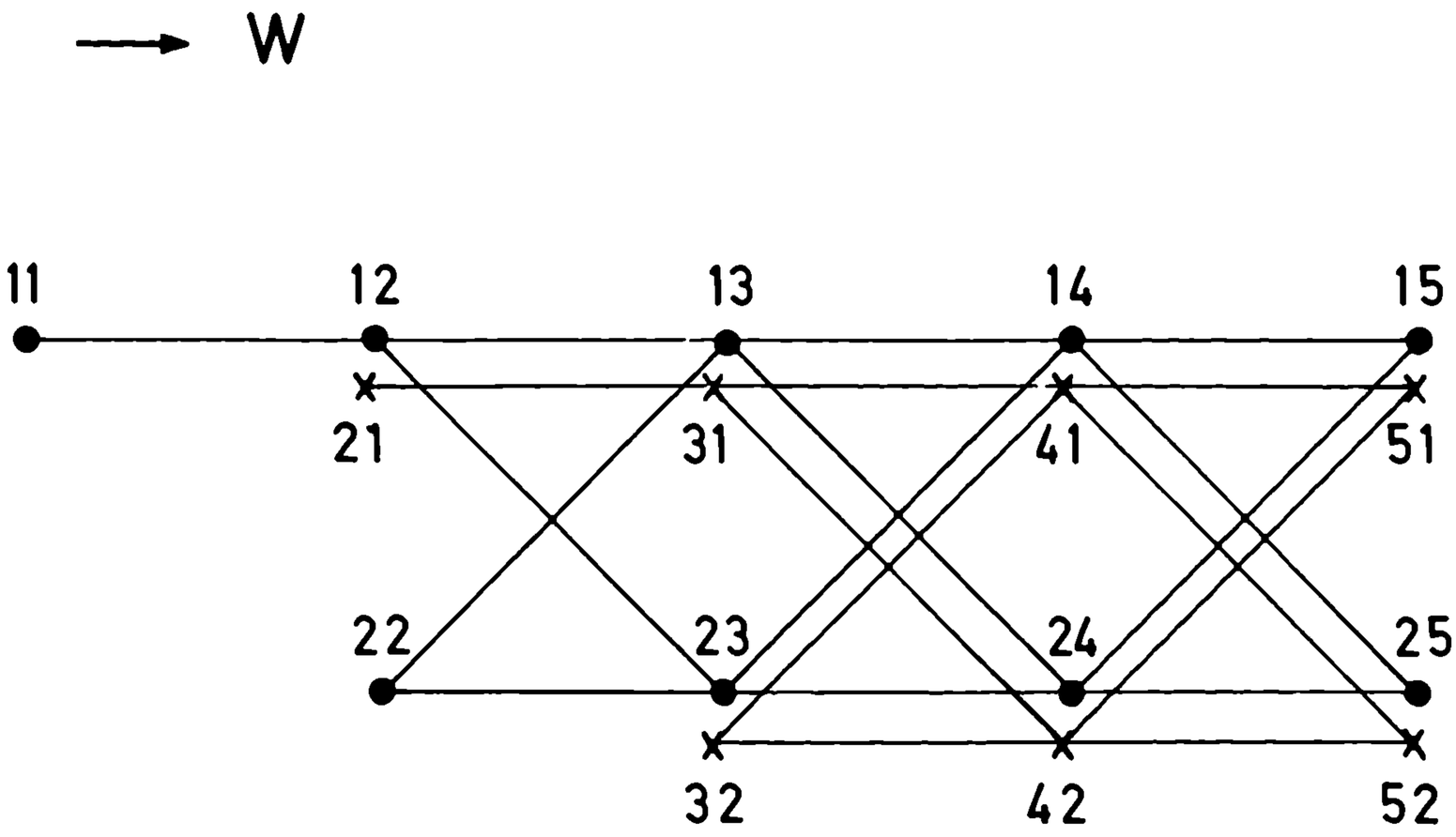


Figure 2 Splitting of the Terms in a Two-Electron System. The Terms of the System are Organized According to Increasing W. The Notation of the Terms by Dots (●) and Crosses (×) is in Agreement with the One used in Eqs. (297a) and (297b). The Numbers at Each Term (i.e., 11, 12, etc.) Indicate the Value of the Quantum Numbers n and m Associated with the Electrons a and b, Respectively (When the System is Unperturbed).

which is zero. Hence two states from different sequences do not combine. Finally, Heisenberg showed that the so-called 'equivalent states,' that is, states in which the electrons of the unperturbed states assume equivalent orbits, say of quantum number m, do not combine with the terms of the sequence denoted by crosses (×); thus, for instance, the amplitude $f'_{mm_1,\,mm}$, with $m_1 < m$, is found to be given by the equation

$$f'_{mm_1,\,mm} = \frac{1}{\sqrt{2}}\left(f_{m_1m,\,mm} - f_{mm_1,\,mm}\right), \tag{298c}$$

which again is zero because of the symmetry properties of the amplitudes of the unperturbed system. Similarly, for the combination of terms of the sequence denoted by dots (●) with the 'equivalent states,' one obtains the transition amplitude

$$f'_{n_1m,\,mm} = \frac{1}{\sqrt{2}}\left(f_{n_1m,\,mm} + f_{mn_1,\,mm}\right) = \sqrt{2}\, f_{n_1m,\,mm}. \tag{298d}$$

This result implies that the 'equivalent states' have indeed to be counted as members of the dot terms (●), as indicated by Heisenberg already in Fig. 2. [393]

[393] The results contained in Eqs. (298a–d) were consistent with the theorem of 'spectroscopic stability' stated in the three-man paper (BHJ, 1926, p. 590; see Section III.5 in this volume). Heisenberg noted explicitly in his paper:

> The line intensities within the two term series are in first approximation the same as the intensities between the corresponding original terms of the single term system (a and b), since the amplitudes of the type $f_{n_1m,\,mn_2}$, which correspond to simultaneous transitions of both systems, disappear in first approximation. Only for combinations with those states, which correspond to equivalent states of the partial systems a and b, is the intensity in the one term series—due to Eqs. [(298c, d)]—twice as large as in the original system; the other term series contains no such states. Therefore, the "total intensity" is not changed by the resonance in the first approximation. (Heisenberg, 1926b, p. 419)

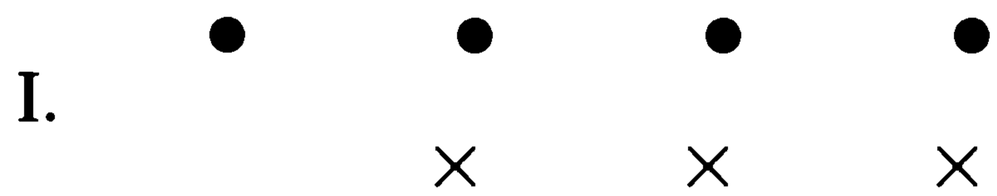


Figure 3 Term Spectrum of Helium without Spin. The Energy Increases from Left to Right.

The results contained in Eqs. (298a–d) encouraged Heisenberg to treat the helium atom as a system showing what he called the 'resonance effect,' i.e., a behaviour characterizing systems which can approximately be separated into two partial systems such that 'the absorption frequency of one system coincides with the emission frequency of the other system, and *vice versa*' (Heisenberg, 1926b, p. 416). The helium spectra indeed possessed many features similar to the one derived in the case of two coupled linear oscillators (of the same frequency) which Heisenberg had discussed in his letter to Born and, in more detail, in the above-mentioned paper on the many-body problem and resonance. Thus, he explained, for example, in his letter to Born:

> Now the application to helium is as follows: 1. We assume that the electrons have no magnets [i.e., are spinless]. Then the term spectrum evidently looks like [Fig. 3] and the two systems, in which the term spectrum separates, are parhelium [denoted by the dots (●), in agreement with Fig. [2]] and ortho-helium [denoted by the crosses (×)]. With this explanation one understands immediately, first, that the separation between par- and ortho-helium is of the order of magnitude of the Rydberg correction; second, that no combinations occur between the two systems in this approximation; and third, that the 1*s*-term [i.e., the term denoted by (11) in Fig. [2] or the first dot (●) in the first row of Fig. [3]] in the ortho-system is missing. (Heisenberg to Born, 26 May 1926)

Indeed, the singlet–triplet difference of corresponding terms—as, for instance, of (12) and (21) or of (24) and (42) in Fig. 2—is obtained from Eqs. (297a) and (297b) as

$$\Delta_{nm} = 2H^{1}(nm, mn), \tag{299}$$

where $H^{1}(nm, mn)$ may be interpreted physically as the difference of the perturbation energies which the Coulomb repulsion between the electrons adds to the energies, W_n^a and W_m^b, of the single electron orbits. The same Coulomb repulsion was also responsible for the so-called Rydberg correction in the empirical helium term formula; hence Eq. (299) provided a qualitative understanding of the helium spectra. And, as Heisenberg wrote to Born: 'Quantitatively I have computed the singlet–triplet separation for *p*-terms; this agrees rather well [with the data] for He and Li^{+}' (Heisenberg to Born, 26 May 1926).

This coarse model for treating the helium atom did not, of course, solve all the problems which Heisenberg had set for himself. It did not account for effects due to electron spin, nor did it make any use of the exclusion principle (whose role in atoms having several electrons had become well established). In order to approach these problems, Heisenberg developed modifications of his simplified

model, which he mentioned as items two and three in his letter to Born. He began by describing a second model as: '2. If one now puts arbitrary magnets on the electrons [i.e., the electrons are associated with a magnet having a magnetic moment $\boldsymbol{\mu}$ in a fixed direction, say along the z-axis], then the picture I of the terms [see Fig. 3] remains exactly the same, but now the electrons are in general not equivalent anymore, hence weak inter-combinations occur' (Heisenberg to Born, 26 May 1926). That is, if the interaction between the electrons' magnetic moments (in a fixed direction in space) and the internal magnetic field $\mathbf{H}_i$ (caused by the helium core at the position of the electrons—see Section V.5) is considered, then the energy of a given orbit (i.e., the average energy of the electron moving in an orbit, neglecting the Coulomb repulsion between the electrons) depends on the relative orientation of the magnetic moment $\boldsymbol{\mu}$ (and, therefore, also of the spin $\mathbf{s}$, according to Eq. (278)) with respect to the plane of the orbit. Although this magnetic energy (whose strength is equal to the scalar product $\boldsymbol{\mu} \cdot \mathbf{H}_i$) is comparatively small, it removes the initial degeneracy of the two-electron system. The perturbation calculation, including the Coulomb term and the magnetic term, then leads more or less to the same spectra as shown in Fig. 2 or Fig. 3. However, each term (nm) gets slightly shifted and transitions between the dotted (●) and crossed (×) sequences begin to take place; their intensities are determined by the magnetic interaction and are, therefore, small.

Though model 2 represented more features of the helium atom than the spinless model, Heisenberg was not yet satisfied. In particular, this model did not quite explain the structure of the helium spectra, for, with a fixed direction of the spin, both terms would be singlets. Moreover, it was difficult to understand how nature would create a preferred direction of electron spin. Heisenberg, therefore, proceeded to deal with a third model, which he described to Born as follows: '3. If one puts the right [Uhlenbeck–] Goudsmit magnet *without* a given direction on each electron, then the electrons are again equivalent. But the separation [after the Coulomb repulsion term H^1 is switched on] into the two systems is different from the previous one. To begin with, for each term of I, four positions of the magnet must occur, hence four new terms will result [instead of one in I]' (Heisenberg to Born, 26 May 1926). Indeed, Heisenberg knew from his earlier considerations of the magnetic interaction (e.g., from the discussion of the anomalous Zeeman effects) that the spins of two electrons might add up to produce either a singlet (total spin 0) or a triplet (total spin 1) having three separate states. This term scheme, to a good approximation, is represented by Fig. 4, which Heisenberg drew in his letter to Born, calling it system II: in it the spin singlet term and the middle term of the spin triplet (i.e., the term of the spin triplet with zero component of the spin in the direction of the internal magnetic field) coincided. Now it seemed that the division into two term sequences lost its meaning: on the one hand, each quadruplet could be divided into a singlet and a triplet system representing qualitatively both term spectra of helium; on the other, weak transitions between all terms appeared to be possible. However, the last conclusion was wrong, for there occurred a new, different separation of term sequences, which was denoted again by dots (●) and crosses (×) in Fig. 4.

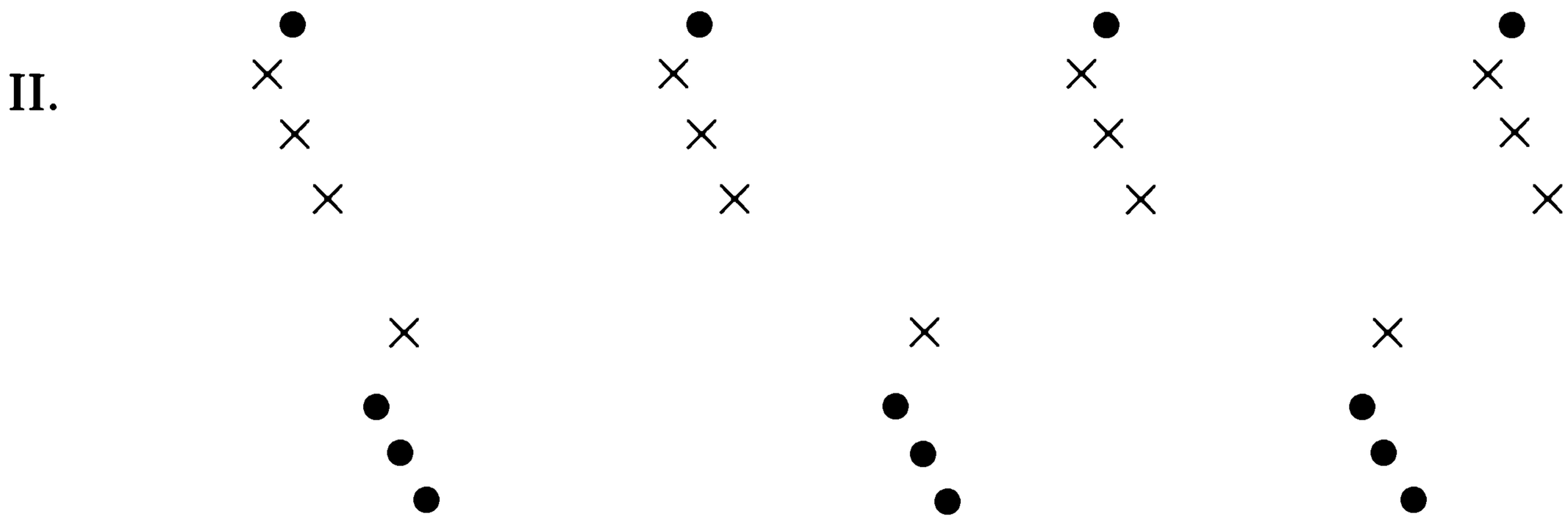


Figure 4 Term Scheme for the Helium Atom.

Heisenberg reported to Born: 'The one system (● in II) has the singlets in parhelium and the triplets in ortho-helium, while the other (×) has the triplets in parhelium and the singlets in ortho-helium. The two systems, which are thus characterized [that is, the terms of the two sequences ● and ×] *cannot* now combine at all; on the other hand, weak inter-combinations between parhelium and ortho-helium [states] are now possible within one system [● or ×]([they have the] order of magnitude of [the ratio of] magnetic interaction to singlet–triplet [energy difference])' (Heisenberg to Born, 26 May 1926).

The above claims rested on two observations which Heisenberg discussed in some detail in his published paper on the spectra of atoms with two electrons (Heisenberg, 1926c, Part II, § 1). First, it is evident from Fig. 4 that in each of the quadruplets of the first and second row the two states (having the same component, m_s, of the total spin in the direction of the inner field $\mathbf{H}_i$) are degenerate if one neglects the mutual interaction between $\boldsymbol{\mu}_a$ and $\boldsymbol{\mu}_b$, the magnetic moments of the two electrons a and b. However, taking into account the latter interaction—which again is considerably smaller than the interaction between the internal magnetic field of the atom and the electrons' magnetic moments—one may repeat a perturbation calculation by using a scheme similar to the one given in Eqs. (296), but with a different perturbation Hamiltonian. As a result the two degenerate states of each quadruplet, which have zero m_s-value, will split into a separate pair with an energy difference of the order of the interaction energy between the magnetic moments $\boldsymbol{\mu}_a$ and $\boldsymbol{\mu}_b$. The higher term of each pair of states with zero m_s may then be grouped together with the two remaining states of the quadruplet, to which both electrons contribute the total m_s-value, either $+1$ or -1; the remaining lower state of each pair, however, becomes an isolated state. Thus, each quadruplet is ordered into a singlet (with total spin momentum zero) and a triplet (with total spin momentum $h/2\pi$ and the components $m_s = +1, 0, -1$). Moreover, no singlet of the upper term sequence in Fig. 4 will combine with any triplet state of the same sequence; similarly there will be no singlet–triplet transitions among the terms of the lower sequence in Fig 4. Strong transitions occur between different states of singlets and between different states of triplets in the same term sequence.

The second important point to be noted is that the singlet terms of the lower sequence (denoted by ×) show a special symmetry when the two electrons of the

atom exchange their orbits *and* their spin components (i.e., m_s^a and m_s^b) in the direction of the internal field. As a result, nonvanishing transition amplitudes might be obtained between the triplet states of the upper sequence and the singlet states of the lower sequence (in Fig 4), whose magnitude is determined by the interaction between the two magnetic moments. Similarly, transitions are to be expected between the singlet states of the upper sequence and the triplet states of the lower sequence. The calculation of the transition amplitudes, which included the spin degrees of freedom, confirmed these possibilities. It also showed that transitions between the triplets of the upper series and the triplets of the lower series were forbidden, as were the transitions between the singlet states of the upper and the singlet states of the lower sequence.[394]

At one point the model 3 in Heisenberg's letter to Born was not correct; it turned out to be *too rich in so far as it provided two ortho- and two parhelium term sequences*. In order to reproduce the empirical facts which implied one term sequence for each (ortho- or par-) helium spectrum, Heisenberg proposed the following hypotheses: 'The system ● [in Fig. 4] represents the helium spectrum and [the system] × does not exist at all.' This hypothesis, he argued in his letter to Born, would not contradict the purely quantum-mechanical calculation 'since all the probabilities of transition to the other system vanish anyway; the term spectrum is "complete" ("*vollständig*") in your sense.'[395] The exclusion of one of

[394]The reason is that the transformation matrix S, whose elements enter into the matrix equation $f' = S^{-1}fS$ (where f describes the matrix of the amplitudes of the unperturbed system—which is degenerate both in the orbital and spin angular momentum quantum numbers), is now a product of two transformation matrices, S' and S''. S' refers to the transformation matrix which has been introduced in connection with the perturbation term H^1 stemming from the Coulomb repulsion of the electron. (This perturbation term removes the orbital degeneracy of the electrons.) S'' removes the degeneracy in m_s, i.e., it has been introduced in connection with the perturbation term stemming from the interaction of the two magnetic moments of the electrons. Now, for instance, if one calculates the amplitude of transition between a triplet state of the upper sequence and a singlet state of the lower sequence, all the elements of the matrix S^{-1}—which multiply f from the left-hand side—are identical and equal to $1/\sqrt{2}$. On the other hand, the elements of the matrix S on the right-hand side stem from the second columns of the matrices S' and S''. Both are antisymmetric with respect to the quantum number involved, as can be seen from Eq. (297b). However, if both the orbital *and* the spin quantum numbers are exchanged, the two minus signs in the matrix elements $S_{mn,\,mn}$ (in the second Eq. (297b)) cancel. Hence all the matrix elements under consideration again become identical and equal to $1/\sqrt{2}$, and finite transition amplitudes are obtained.

It should be mentioned that it is also possible to prove that transitions which are not allowed according to the above results can also not be excited by collisions of atoms. Heisenberg proved this by replacing the two-electron system by a system of coupled oscillators in which nonlinear terms (in the oscillator coordinates q_1 and q_2) are taken into account. Again he found that the separation between the two term sequences is strictly maintained. Since the effect of collisions between atoms can be described in terms of these nonlinear couplings, no new transitions emerge in two-electron systems even if collisions occur between them (Heisenberg, 1926b, pp. 415–416).

[395]In the joint paper with Norbert Wiener, Born had postulated the existence of a closed set of functions associated with each operator in quantum mechanics (Born and Wiener, 1926, p. 92); in particular, this set had been defined by the fact that any operator of a mechanical system applied to a function leads to another function of the set. The functions may be chosen to have the form $\exp(2\pi i\nu t)$, where the ν are the quantum-mechanical frequencies. Now if only transitions between states of one term sequence occur, no frequencies other than those given by the equation $h\nu = W_n - W_m$, where W_n and W_m are the terms of the sequence, arise. Hence this term sequence leads to a closed and *complete* set of functions, $\{\exp(2\pi i\nu t)\}$, as postulated by Born and Wiener.

the term sequences was not only possible, to Heisenberg it even seemed to be a necessary requirement because it 'agrees with the Bose statistics, which reduces the statistical weights by a factor of two and automatically yields Pauli's ban [exclusion of equivalent orbits]' (Heisenberg to Born, 26 May 1926).

Thus, Heisenberg finally brought in the exclusion principle which was known to play an important role in atomic systems containing several electrons. He had also demonstrated, as he had mentioned three weeks earlier to Pauli (Heisenberg to Pauli, 5 May 1926), how the exclusion of equivalent orbits, that is, orbits which have all of their quantum numbers (including m_s) identical, was connected with the singlet–triplet energy difference in helium. Moreover, he had made a move in the direction of realizing Pauli's programme of a deeper understanding of the exclusion principle, for, as he noted: 'If only one (●) of the two systems occurs in nature, then it yields, on one hand, a reduction of the statistical weights, just in the sense proposed by Bose; on the other hand, Pauli's exclusion of equivalent orbits is automatically satisfied for the correct choice of the system' (Heisenberg, 1926b, p. 422). By generalizing this consideration to the case of atomic systems containing n electrons, in which he proposed to retain only that one of the $n!$ noncombining sequences (arising from the perturbation calculations) which contained no equivalent orbits, Heisenberg concluded that 'Pauli's exclusion prescription and Bose's rule are the same, that they do not contradict quantum mechanics, and that one can calculate the energy values and all other quantities in quantum mechanics' (Heisenberg to Born, 26 May 1926).

In relating together Pauli's exclusion principle and the Bose statistics [for light-quanta] Heisenberg, who was hardly an expert in statistical mechanics, had been particularly misled by the observation that Bose's counting method consisted in reducing $n!$, the statistical weight of a state containing n identical particles (equal to 2 in the case of a two-electron system), to unity. Enrico Fermi had already discovered a little earlier that there was another counting method which, obeying Pauli's exclusion principle, would also reduce the statistical weight in the same way, but that it was different from the method proposed by Bose for light-quanta and applied by Einstein to a monatomic ideal gas (Fermi, 1926a). Although Fermi's work was also published in an issue of *Zeitschrift für Physik* before the end of May (Fermi, 1926b), Heisenberg did not sense the connection between Fermi's ideas and his own investigations on two-electron atoms.[396]

Pauli, who, in June 1926, received from Heisenberg a copy of the manuscript

[396] Paul Dirac, like Heisenberg, made use of Schrödinger's wave mechanical approach in treating quantum systems of several particles; without taking notice of Fermi's papers, he rediscovered Fermi (–Dirac) statistics a few months later (Dirac, 1926f). Even more remarkable, Pauli, who, based on his record at that time, was a greater expert on statistical mechanics than either Heisenberg or Dirac, did not at all recognize the connection established by Fermi and Dirac between the statistics of identical particles, the many-electron problem and the exclusion principle. At first, Pauli maintained that it (the Fermi–Dirac statistics) was all wrong, but several months later, after he had carefully studied the fluctuation term in gases, he wrote to Heisenberg: 'I now think much more kindly about the Fermi–Dirac statistics, and several arguments appear to me to speak in its favour' (Pauli to Heisenberg, 19 October 1926).

on the helium problem (see Heisenberg to Pauli, 8 June 1926), was not satisfied with Heisenberg's attempt at a deeper justification of the exclusion principle by applying 'unattractive additional constraints' ('*unschöne Zusatzverbote*,' Pauli to Wentzel, 5 July 1926) to the quantum-mechanical equations. He hoped for a better argument; but since, for the time being he was unable to find it, Heisenberg went ahead with his helium calculation and eventually submitted a paper on the spectra of two-electron systems towards the end of July (Heisenberg, 1926c). In this paper he simplified some of his previous arguments by applying a different method: Schrödinger's wave mechanics. Thus, for example, he easily obtained the result that in the case of electrons with spin (i.e., the model 3 in his letter to Born) the resulting term systems denoted by dots (●) and crosses (×), respectively, would not combine; the reason was that the crossed terms had to be described by Schrödinger functions, which were antisymmetric in both the space and spin coordinates of the electrons, while the dotted terms had to be described by totally symmetric Schrödinger functions, and no transitions occurred between any symmetric and antisymmetric states. The second, and for Heisenberg the more important, point, was a practical one: by the end of May he was well prepared to embark upon the detailed computation of helium states. In principle, the tools of matrix mechanics had already been extended—by Cornelius Lanczos, Born and Norbert Wiener, Pauli and Heisenberg himself, and by Gregor Wentzel and Paul Dirac—to the extent that the two-electron system could be treated. However, at the same time Erwin Schrödinger's method of the wave equation had also been published (Schrödinger, 1926c), and Heisenberg did not hesitate using it (rather than the clumsier matrix methods) to obtain the elements of perturbation matrices occurring in equations of the type of Eqs. (296) for the energy states of the helium atom.[397] Thus, a calculation which was considered as a climactic proof of the validity of matrix mechanics, provided an immediate triumph of the alternative scheme of wave mechanics.

However, the key to the helium problem—the resonance effect in coupled equivalent quantum systems—had been found within the Göttingen formulation of quantum mechanics. This resonance effect would play a crucial role in the solution of two outstanding problems of atomic theory: the problem of covalent binding of molecules, such as the hydrogen molecule, and the problem of understanding the origin of ferromagnetism. While the hydrogen molecule was first treated by Walter Heitler and Fritz London, who associated the covalent

[397] Heisenberg knew from Pauli that Schrödinger's scheme was equivalent to the extended matrix formalism and could be used, especially, to calculate matrix elements elegantly. In his paper on the two-electron atoms, Heisenberg stuck to matrix methods and considerations as far as principles were concerned; however, when it came to calculate the matrix elements of the perturbation terms in the Hamiltonian, he remarked simply: 'The calculation of the quantities is performed in the simplest way according to Schrödinger's method of eigenfunctions' (Heisenberg, 1926c, p. 503).

Let us add here in passing that Gregor Wentzel, who had discussed the Stark effect (including that of the second order in the electric field strength) at the joint meeting of the *Gauverein Württemberg, Baden und Bayern* of the German Physical Society on 16 May 1926 in Stuttgart, did not publish his results in contrast to his announcement (see, *Verh. d. Deutsch. Phys. Ges.* (3) **7**, p. 26, (1926)). The obvious reason was that meanwhile Schrödinger had come out with his calculations of the Stark effect by means of the wave mechanical methods (Schrödinger, 1926f).

binding with the exchange forces arising from the resonance effect (Heitler and London, 1927), Heisenberg himself attacked the problem of ferromagnetism beginning in fall 1926. There one had to cope with a similar difficulty as in the helium problem, namely, how only the weak spin–spin interactions (i.e., the interactions between the magnetic moments arising from the spin due to Eq. (278)) can create the 'strong molecular field' introduced by Pierre Weiss many years earlier (Weiss, 1907). Heisenberg discovered that in the phenomenon of ferromagnetism it was again the Coulomb forces between the electrons that gave rise to the molecular field, whereas the mutual interaction of electron spins only created the empirically observed regular alignments of the magnetic moments of the atoms in the solid state (Heisenberg, 1928).[398]

Thus, in spite of the fact that Pauli's question concerning the physical explanation of the exclusion principle could not yet be answered, the helium calculation opened the path towards an understanding of complex atomic systems.[399] In any case, Heisenberg was very optimistic. 'I am convinced,' he wrote to Jordan a few days after submitting his paper on helium spectra to *Zeitschrift für Physik*, 'that the spectra of all chemical elements can be obtained (apart from the satellite structure [i.e., those caused by a magnetic moment of the nucleus, as in Pauli's treatment of the hyperfine structure (Pauli, 1924c)]) from quantum mechanics in a unique manner without physics [i.e., additional physical assumptions] by bone-headed calculation' (Heisenberg to Jordan, 28 July 1926).

[398] Heisenberg first suggested his ideas about solving the problem of ferromagnetism in a letter to Pauli. He wrote:

> I have myself thought a little about the theory of ferromagnetism, superconductivity, and similar obscenities ("*Schweinerein*"). The idea is this: In order to use Langevin's theory of ferromagnetism, one has to assume a large coupling force between the spinning electrons (*only these* truly rotate). As in the case of helium, this force should be provided indirectly by the resonance. I believe that one can prove in general that the *smallest* energy results when the spin-vectors are parallel to each other. The energy differences in question are of the *electrical* order of magnitude, but they vanish *very* fast with increasing distance [between the atoms]. I have the feeling (without knowing the matter in the least) that this might suffice in principle for an interpretation of ferromagnetism. In order to decide the question why most materials are not ferromagnetic, but some are, one must of course calculate quantitatively; one could perhaps make it plausible that in the case of Fe, Co and Ni the situation [for ferromagnetism] is most favourable. It is the same for conductivity; here the resonance migration of electrons *à la* Hund may explain the situation. (Heisenberg to Pauli, 4 November 1926)

[399] In trying to approach the problem of explaining the exclusion principle, Pauli, in fall 1926, turned to fluctuation considerations. Since the fluctuation formula for particles obeying Fermi statistics (which corresponds to Einstein's formula, Eq. (181)) contains a wave-like term with a negative sign, Pauli speculated: 'Perhaps this behaviour might be interpreted by assuming phase relations between the different de Broglie eigenvibrations; this might give a hint for a future physical explanation of the fundamental basis of Fermi-statistics and therewith also of the "equivalence rule" [i.e., the exclusion principle]' (Pauli, 1927, pp. 95–96). Pauli continued to study the problems connected with spin and statistics over many years, returning to them time and again during the following three decades, before a more satisfactory 'explanation' of the exclusion principle from general quantum-theoretical considerations could be given; this was the so-called spin-statistics theorem. (See, e.g., Pauli, 1955.)

Interestingly enough, in fall 1926 Heisenberg also concerned himself with the treatment of fluctuations in quantum-mechanical systems (Heisenberg, 1926e). However, Heisenberg's work dealt with the connection between the concepts of matrix mechanics and wave mechanics, In pursuing this question futher, Heisenberg arrived at the uncertainty relations, which constituted a fundamental step in the physical interpretation of quantum mechanics (Heisenberg, 1927b).

References

Note: We have used the following abbreviations for journals and periodicals. Journals and periodicals not mentioned here are cited by their full titles. In the list of references we have given the dates, if available, of the particular issues in which the papers were published—except the *Sitzungsberichte* or *Comptes rendus* of the European academies. The references are arranged in alphabetical order according to the author, and in time-order according to the date of publication of the papers. Books or contributions to books (though not the yearbooks, which are treated as journals) have always been placed at the end of the annual list of a given author. The time-ordering letters (a, b, c, . . .) refer to all the papers of an author cited in Volumes 1 to 4.

American Journal of Physics:	*Amer. J. Phys.*
Annalen der Physik, fourth series:	*Ann. d. Phys. (4)*
Archives for the History of Quantum Physics (Interviews):	AHQP Interview
Comptes rendus hebdomadaires des séances de l'Académie des Sciences (Paris):	*Comptes rendus (Paris)*
Encyklopädie der mathematischen Wissenschaften mit Einschluß ihrer Anwendungen:	*Encykl. d. math. Wiss.*
Ergebnisse der exakten Naturwissenschaften:	*Ergeb. d. exakt. Naturwiss.*
Journal of the Franklin Institute:	*J. Franklin Inst.*
Journal of Mathematics and Physics of the M.I.T.:	*J. Math. & Phys. M.I.T.*
Journal de physique théorique et appliquée (Paris):	*J. phys. (Paris)*
Journal für die reine und angewandte Mathematik:	*J. reine & angew. Math.*
Mathematisk-Fysiske Meddelelser, Det Kgl. Danske Videnskabernes Selskab:	*Kgl. Danske Vid. Selsk. Math.-fys. Medd.*
Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-physikalische Klasse:	*Nachr. Ges. Wiss. Göttingen*
Naturwissenschaften:	*Naturwiss.*
Philosophical Magazine, third, fifth and sixth series:	*Phil. Mag. (3), (5), (6)*

Philosophical Transactions of the Royal Society of London: *Phil. Trans. Roy. Soc. (London)*

The Physical Review, second series: *Phys. Rev. (2)*

Physikalische Zeitschrift: *Phys. Zs.*

Proceedings of the Irish Academy: *Proc. Irish Acad.*

Proceedings of the National Academy of Sciences (U.S.A.): *Proc. Nat. Acad. Sci. (U.S.A.)*

Proceedings of the Royal Society of London, series A: *Proc. Roy. Soc. (London)* A

Rendiconti del Circulo Matematico di Palermo: *Rend. Circ. Mat. Palermo*

Rendiconti del Reale Accademia Lincei: *Rend. R. Accad. Lincei*

Reports of the British Association for the Advancement of Science: *Reports Brit. Assoc. Adv. Sci.*

Sitzungsberichte der Bayerischen Akademie der Wissenschaften (München), mathematisch-physikalische Klasse: *Sitz.ber. Bayer. Akad. Wiss. (München)*

Sitzungsberichte der Preussischen Akademie der Wissenschaften (Berlin), physikalisch-mathematische Klasse: *Sitz.ber. Preuss. Akad. Wiss. (Berlin)*

Skrifter, Det Kgl. Danske Videnskabernes Selskab: *Kgl. Danske Vid. Selsk. Skrifter*

Verhandlungen der Deutschen Physikalischen Gesellschaft, second series: *Verh. d. Deutsch. Phys. Ges. (2)*

Zeitschrift für Physik *Z. Phys.*

ALLEN, HERBERT STANLEY

1927 *Probleme der Atomdynamik. Erster Teil: Die Struktur des Atoms; Zweiter Teil: Die Gittertheorie des festen Zustandes. Dreißig Vorlesungen gehalten im Wintersemester, 1925–26, am Massachusetts Institute of Technology.* Von Prof. Max Born, *Nature* **119**, 77–79 (book review published in the issue of 15 January 1927).

APPELL, PAUL-ÉMILE

1885 Sur une méthode élémentaire pour l'obtenir les développements en série trigonométrique des fonctions elliptiques, *Bulletin de la Société Mathématique de France* **13**, 13–18 (presented at the meeting of 6 December 1884, published 1885).

AUTONNE, LÉON

1902 Sur l'Hermitien, *Rend. Circ. Mat. Palermo* **16**, 104–128 (dated 1 August 1901, published in issue No. 1/2 of 3 March 1902).

BACK, ERNST

1925b Regelwidrige Zeemaneffekte von Multipletts I. Stufe, *Z. Phys.* **33**, 579–600 (received 27 May 1925, published in No. 8 of 18 August 1925).

BELL, ERIC TEMPLE

1937 *Men of Mathematics,* London: V. Gollancz; paperback reprint, New York: Simon & Schuster, 1965.

BERNOULLI, DANIEL

1732 Observationes de seribus quae formantur ex additione vel subtractione quacumque terminorum se mutuo consequentium, *Commentarii Academiae Scientiarum Imperialis Petropolitanae* **3**, 92ff. (presented 1728, published 1732).

BLÜH, OTTO

1926 Neuere Ergebnisse auf dem Gebiete der Dielektrizitätskonstanten, *Phys. Zs.* **27**, 226–267 (received 21 December 1925, published in issue No. 8 of 15 April 1926).

BÔCHER, MAXIME

1907 *Introduction to Higher Algebra,* New York: Macmillan. German translation: *Einführung in die höhere Algebra,* Leipzig: B. G. Teubner, 1911.

BOHR, NIELS

1913b On the constitution of atoms and molecules (Part I), *Phil. Mag. (6)* **26**, 1–25 (dated 5 April 1913, published in issue No. 151 of July 1913); reprinted in: *On the Constitution of Atoms and Molecules: Papers of 1913 reprinted from the Philosophical Magazine with an Introduction by L. Rosenfeld*, Copenhagen: Munksgaard Ltd., New York: W. A. Benjamin, 1963, pp. 1–25.

1918a On the quantum theory of line-spectra. Part I. On the general theory, *Kgl. Danske Vid. Selsk. Skrifter*, 8. Raekke, **IV.1**, 1–36 (dated November 1917, published in April 1918); reprinted in N. Bohr: *Collected Works*, *Volume 3*, Amsterdam–New York–Oxford: North–Holland Publishing Company, 1976, pp. 67–102.

1918b On the quantum theory of line spectra. Part II. On the hydrogen spectrum, *Kgl. Danske Vid. Selsk. Skrifter*, 8. Raekke, **IV.1**, 37–100 (published in December 1918); reprinted in N. Bohr: *Collected Works*, *Volume 3*, 1976, pp. 103–166.

1925a Über die Wirkung von Atomen bei Stößen, *Z. Phys.* **34**, 142–157 (received 30 March 1925, with a postscript dated July 1925, published in issue No. 2/3 of 28 September 1925).

1925b Atomic theory and mechanics, *Nature* **116**, 845–852 (elaborated text of an address presented on 30 August 1925 at the sixth Scandinavian Mathematical Congress in Copenhagen, published in the Supplement to the issue of 5 December 1925). German translation: Atomtheorie und Mechanik, *Naturwiss.* **14**, 1–10 (published in the issue of 1 January 1926).

BOHR, NIELS, HENDRIK ANTHONY KRAMERS AND JOHN CLARKE SLATER

1924 The quantum theory of radiation, *Phil. Mag. (6)* **47**, 785–822 (dated January 1924, published in issue No. 281 of April 1924); reprinted in *Sources of Quantum Mechanics* (van der Waerden, 1967), pp. 159–176. German translation: Über die Quantentheorie der Strahlung, *Z. Phys.* **24**, 69–87 (received 22 February 1924, published in issue No. 2 of 22 May 1924); reprinted in *Collected Scientific Papers* (Kramers, 1956), pp. 271–289.

BOOLE, GEORGE

1844 On a general method in analysis, *Phil. Trans. Roy. Soc. (London)* **134**, 225–282 (communicated by S. Hunter Christie; received 12 January 1844, read 18 January 1844, published with a postscript of 31 August 1844).

BORN, MAX

1909b Die Theorie des starren Elektrons in der Kinematik des Relativitätsprinzips, *Ann. d. Phys. (4)* **30**, 1–56 (received 13 June 1909, published in issue No. 11 of 31 August 1909); reprinted in *Ausgewählte Abhandlungen 1*, pp. 105–160.

1914a Der Impuls-Energie-Satz in der Elektrodynamik von Gustav Mie, *Nachr. Ges. Wiss. Göttingen*, pp. 23–37 (presented by D. Hilbert at the meeting of 20 December 1913); reprinted in *Ausgewählte Abhandlungen 2*, pp. 470–483.

1921 Über elektrostatische Gitterpotentiale, *Z. Phys.* **7**, 124–140 (received 25 August 1921, published in issue No. 2 of 21 October 1921); reprinted in *Ausgewählte Abhandlungen 1*, pp. 434–450.

1922 Über das Modell der Wasserstoffmolekel, *Naturwiss.* **10**, 677–678 (letter dated 27 June 1922, published in the issue of 4 August 1922).

1923b Atomtheorie des festen Zustandes (Dynamik der Kristallgitter), *Encykl. d. Math. Wiss. V/3*, pp. 527–781 (dated 7 September 1923, published in issue No. 4 of 28 October 1923).

1924b Über Quantenmechanik, *Z. Phys.* **26**, 379–395 (received 13 June 1924, published in issue No. 6 of 20 August 1924); reprinted in *Ausgewählte Abhandlungen 2*, pp. 61–77, also in *Begründung der Matrizenmechanik* (Born, Heisenberg and Jordan, 1962), pp. 13–29. English translation: Quantum mechanics, in *Sources of Quantum Mechanics* (van der Waerden, 1967), pp. 181–198.

1925 *Vorlesungen über Atommechanik*, (Vol. 2 of the series *Struktur der Materie*, M. Born and J. Franck, eds.), Berlin: J. Springer Verlag. English translation (by J. W. Fisher): *Mechanics of the Atom*, London: Bell, 1927.

1926a Zur Quantenmechanik der Stoßvorgänge. (Vorläufige Mitteilung), *Z. Phys.* **37**, 863–867 (received 25 June 1926, published in issue No. 12 of 10 July 1926); reprinted in *Ausgewählte Abhandlungen 2*, pp. 228–232, and in *Zur statistischen Deutung der Quantentheorie* (Born, 1962a), pp. 48–52.

1926b Quantenmechanik der Stoßvorgänge, *Z. Phys.* **38**, 803–827 (received 21 July 1926, published in issue No. 11/12 of 14 September 1926); reprinted in *Ausgewählte Abhandlungen 2*, pp. 233–257, and in *Zur statistischen Deutung der Quantentheorie* (Born, 1962a), pp. 53–77.

1926c Das Adiabatenprinzip in der Quantenmechanik, *Z. Phys.* **40**, 167–192 (received 16 October 1926, published in issue No. 3/4 of 6 December 1926); reprinted in *Ausgewählte Abhandlungen 2*, pp. 258–283, and in *Zur statistischen Deutung der Quantentheorie* (Born, 1962a), pp. 93–118.

1926d *Problems in Atomic Dynamics* (lectures delivered at the M.I.T.), Cambridge (Mass.): M.I.T. Press; paperback edition, 1970.

1940 Prof. Otto Toeplitz, *Nature* **145**, 617 (published in the issue of 20 April 1940); reprinted in *Ausgewählte Abhandlungen 2*, pp. 612–613.

1962 *Ausgewählte Abhandlungen*, 2 volumes, Göttingen: Vandenhoeck & Ruprecht.

1962a *Zur statistischen Deutung der Quantentheorie* (No. 1 of the series *Dokumente der Naturwissenschaft, Abteilung Physik*, A. Hermann, ed.), Stuttgart: E. Battenberg Verlag.

1978 *My Life: Recollections of a Nobel Laureate*, London: Taylor & Francis, Ltd., New York: Charles Scribner's Sons.

BORN, MAX, AND WERNER HEISENBERG

1923a Über Phasenbeziehungen bei den Bohrschen Modellen von Atomen und Mole-

keln, *Z. Phys.* **14**, 44–55 (received 16 January 1923, published in issue No. 1 of 5 March 1923).

1923b Die Elektronenbahnen im angeregten Helium, *Z. Phys.* **16**, 229–243 (received 11 May 1923, published in No. 4 of 9 July 1923); reprinted in *Ausgewählte Abhandlungen 2* (Born, 1962), pp. 23–37.

BORN, MAX, WERNER HEISENBERG AND PASCUAL JORDAN

1926 Zur Quantenmechanik. II., *Z. Phys.* **35**, 557–615 (received 16 November 1925, published in issue No. 8/9 of 4 February 1926); reprinted *in Ausgewählte Abhandlungen 2* (Born, 1962), pp. 155–213, and in *Zur Begründung der Matrizenmechanik* (Born, Heisenberg and Jordan, 1962), pp. 77–135. English translation: On quantum mechanics II, in *Sources of Quantum Mechanics* (van der Waerden, 1967), pp. 321–385.

1962 *Zur Begründung der Matrizenmechanik* (No. 2 of the series *Dokumente der Naturwissenschaften, Abteilung Physik*, A. Hermann, ed.), Stuttgart: E. Battenberg Verlag.

BORN, MAX, AND PASCUAL JORDAN

1925a Zur Quantentheorie aperiodischer Vorgänge, *Z. Phys.* **33**, 479–505 (received 11 June 1925, published in issue No. 7 of 15 August 1925); reprinted in *Ausgewählte Abhandlungen 2* (Born, 1962), pp. 97–123.

1925b Zur Quantenmechanik, *Z. Phys.* **34**, 858–888 (received 27 September 1925, published in No. 11/12 of 28 November 1925); reprinted in *Ausgewählte Abhandlungen 2* (Born, 1962), pp. 124–154, and in *Zur Begründung der Matrizenmechanik* (Born, Heisenberg and Jordan, 1962), pp. 47–76. English translation (except for Chapter 4): On quantum mechanics, in *Sources of Quantum Mechanics* (van der Waerden, 1967), pp. 277–306.

1930 *Elementare Quantenmechanik (Zweiter Band der Vorlesungen über Atommechanik)*, Berlin: J. Springer Verlag.

BORN, MAX, AND THEODOR VON KÁRMÁN

1912 Über Schwingungen von Raumgittern, *Phys. Zs.* **13**, 297–309 (received 20 March 1912, published in issue No. 8 of 15 April 1912); reprinted in *Ausgewählte Abhandlungen 1* (Born, 1962), pp. 231–243.

1913b Über die Verteilung der Eigenschwingungen von Punktgittern, *Phys. Zs.* **14**, 65–71 (received 30 November 1912, published in issue No. 2 of 15 January 1913); reprinted in *Ausgewählte Abhandlungen 1* (Born, 1962), pp. 249–255.

BORN, MAX, AND WOLFGANG PAULI

1922 Über die Quantelung gestörter mechanischer Systeme, *Z. Phys.* **10**, 137–158 (received 29 May 1922, published in issue No. 3 of 7 July 1922); reprinted in *Ausgewählte Abhandlungen 2* (Born, 1962), pp. 1–22, and in *Collected Scientific Papers 2* (Pauli, 1964), pp. 48–69.

BORN, MAX, AND OTTO STERN

1919 Über die Oberflächenenergie der Kristalle und ihren Einfluß auf die Kristallgestalt, *Sitz.ber Preuss. Akad. Wiss. (Berlin)*, pp. 901–913 (communicated by A. Einstein to the meeting of 13 November 1919); reprinted in *Ausgewählte Abhandlungen 1* (Born, 1962), pp. 382–394.

BORN, MAX, AND NORBERT WIENER

1926a A new formulation of the laws of quantization of periodic and aperiodic phenomena, *J. Math. & Phys. M.I.T.* **5**, 84–98.

1926b Eine neue Formulierung der Quantengesetze für periodische und aperiodische Vorgänge, *Z. Phys.* **36**, 174–187 (received 5 January 1926, published in issue No. 3 of 12 March 1926); reprinted in *Ausgewählte Abhandlungen 2* (Born, 1962), pp. 214–227.

BOSE, SATYENDRA NATH

1924a Plancks Gesetz und Lichtquantenhypothese, *Z. Phys.* **26**, 178–181 (received 2 July 1924, published in issue No. 3 of 11 August 1924).

BOTHE, WALTHER

1923c Zur Quantentheorie des normalen Photoeffektes, *Z. Phys.* **17**, 137–151 (received 25 June 1923, published in No. 2 of 9 August 1923).

1923e Die räumliche Energieverteilung in der Hohlraumstrahlung, *Z. Phys.* **20**, 145–152 (received 23 October 1923, published in issue No. 2 of 17 December 1923).

1924a Über die Wechselwirkung zwischen Strahlung und freien Elektronen, *Z. Phys.* **23**, 214–224 (received 18 March 1924, published in issue No. 3/4 of 19 April 1924).

1924b Die Emissionsrichtung durch Röntgenstrahlen ausgelöster Photoelektronen, *Z. Phys.* **26**, 59–73 (received 6 June 1924, published in issue No. 1 of 5 August 1924).

BOTHE, WALTHER, AND HANS GEIGER

1925a Experimentelles zur Theorie von Bohr, Kramers und Slater, *Naturwiss.* **13**, 440–441 (letter dated 18 April 1925, published in the issue of 15 May 1925).

1925b Über das Wesen des Comptoneffekts; ein experimenteller Beitrag zur Theorie der Strahlung, *Z. Phys.* **32**, 639–663 (received 25 April 1925, published in issue No. 9 of 12 June 1925).

BOULIGAND, GEORGES

1924a Sur le problème de Dirichlet harmonique, *Comptes rendus (Paris)* **178**, 55–57 (received 26 December 1923, presented at the meeting of 2 January 1924).

1924b Domaines infinis et cas d'exception du problème de Dirichlet, *Comptes rendus (Paris)* **178**, 1054–1057 (presented by H. Lebesgue at the meeting of 24 March 1924).

BROGLIE, LOUIS DE

1924e *Recherche sur la théorie des quanta. Thèses presentées à la Faculté des Sciences de l'Université de Paris pour obtenir le grade de docteur ès sciences physiques par Louis de Broglie; première thèse: Recherche sur la théorie des quanta, deuxième thèse: Propositions données par la Faculté, soutenues le 25 Novembre devant la Commission d'examen: J. Perrin, Président, E. Cartan, Ch. Mauguin, P. Langevin, Examinateurs*, Paris: Masson et Cie.

1925 Recherche sur la théorie des quanta, *Annales de physique (10)* **3**, 22–128 (published in issue No. 1 of January/February 1925).

CARSON, JOHN RENSHAW

1925 Electric circuit theory and the operational calculus, *Bell System Technical Journal* **4**, 685–761 (published in issue No. 4 of October 1925).

CAUCHY, AUGUSTIN LOUIS

1815 Mémoire sur les fonctions que ne peuvent obtenir que deux valeurs—Égales et de signes contraires par suite de transpositions opérées entre les variables qu'elles renferment, *Journal de l'Ecole Polytechnique* **10**, 29–112 (read on 30 November 1812); reprinted in *Oeuvres Complètes 1*, pp. 91–169.

1826 *Leçons sur les applications du calcul infinitésimal à la géométrie*, Paris: Imprimerié Royale; reprinted in *Oeuvres Complètes*, IIe serie, Vol. 5, pp. 3–528, especially pp. 244–285 (15e leçon).

CAYLEY, ARTHUR

1855 Remarques sur la notation des fonctions algébriques, *J. reine & angew. Math.* **50**, 282–285 (Memoir No. 3, dated 24 May 1855, published in issue No. 4).

1858 A memoir on the theory of matrices, *Phil. Trans. Roy. Soc. (London)* **148**, 17–37 (received 10 December 1857, read 14 January 1858); reprinted in *Collected Mathematical Papers 2*, pp. 475–496.

1889– *Collected Mathematical Papers*, 13 volumes.

1898

COMPTON, ARTHUR HOLLY

1921d The magnetic electron, *J. Franklin Inst.* **192**, 145–155 (read on 27 December 1920 before Section B of the American Association for the Advancement of Science, published in issue No. 2 of August 1921).

COURANT, RICHARD, AND DAVID HILBERT

1924 *Methoden der mathematischen Physik. (Erster Band)*, Berlin: J. Springer Verlag.

CRAMER, GABRIEL

1750 *Introduction à l'analyse des lignes courbes algébriques*, Geneva.

DEBYE, PETER

1910 Der Wahrscheinlichkeitsbegriff in der Theorie der Strahlung, *Ann. d. Phys. (4)* **33**, 1427–1434 (received 12 October 1910, published in issue No. 16 of 20 December 1910).

1912a Einige Resultate einer kinetischen Theorie der Isolatoren, *Phys. Zs.* **13**, 97–100 (received 12 December 1911, published in issue No. 3 of 1 February 1912); Nachtrag zur Notiz über eine kinetische Theorie der Isolatoren, *Phys. Zs.* **13**, 295 (received 2 March 1912, published in issue No. 7 of 1 April 1912). English translation: Some results of a kinetic theory of insulators, in *Collected Papers of Peter J. W. Debye*, New York and London: Interscience Publishers, 1954, pp. 173–179.

1916a Quantenhypothese und Zeeman–Effekt, *Nachr. Ges. Wiss. Göttingen*, pp. 142–153 (presented at the meeting of 3 June 1916).

DIRAC, PAUL ADRIEN MAURICE

1925d The fundamental equations of quantum mechanics, *Proc. Roy. Soc. (London)* **A109**, 642–653 (communicated by R. H. Fowler; received 7 November 1925, published in issue No. A752 of 1 December 1925); reprinted in *Sources of Quantum Mechanics* (van der Waerden, 1967), pp. 307–320.

1926a Quantum mechanics and a preliminary investigation of the hydrogen atom, *Proc. Roy. Soc. (London)* **A110**, 561–579 (communicated by R. H. Fowler; received 22 January 1926, published in issue No. A755 of 1 March 1926); reprinted in *Sources of Quantum Mechanics* (van der Waerden, 1967), pp. 417–427 (except Sections 5–7).

1926f On the theory of quantum mechanics, *Proc. Roy. Soc. (London)* **A112**, 661–677 (communicated by R. H. Fowler; received 26 August 1926, published in issue No. A762 of 1 October 1926).

1927a The physical interpretation of the quantum dynamics, *Proc. Roy. Soc. (London)* **A113**, 621–641 (communicated by R. H. Fowler; received 2 December 1926, published in issue No. A765 of 1 January 1927).

1977 Recollections of an exciting era, in *History of Twentieth Century Physics*, (Varenna Summer School, 31 July–12 August 1972), New York: Academic Press Inc., pp. 109–146.

EDDINGTON, ARTHUR STANLEY

1927 *Stars and Atoms*, Oxford: Clarendon Press.

EHRENFEST, PAUL

1925a Energieschwankungen im Strahlungsfeld oder Kristallgitter bei Superposition quantisierter Eigenschwingungen, *Z. Phys.* **34**, 362–373 (received 29 August 1925, published in issue No. 5–7 of 20 October 1925); reprinted in *Collected Scientific Papers*, pp. 513–524.

1925b Bemerkungen betreffs zweier Publikationen über Energieschwankungen, *Z. Phys.* **35**, 316 (received 25 November 1925, published in issue No. 4 of 31 December 1925); reprinted in *Collected Scientific Papers*, p. 525.

1959 *Collected Scientific Papers* (M. J. Klein, ed.), Amsterdam: North-Holland Publishing Company.

EINSTEIN, ALBERT

1905b Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt, *Ann. d. Phys. (4)* **17**, 132–148 (received 18 March 1905, published in issue No. 6 of 9 June 1905); reprinted in Albert Einstein: *Die Hypothese der Lichtquanten*, (No. 7 of the series *Dokumente der Naturwissenschaften*, A. Hermann, ed.), Stuttgart: E. Battenberg, 1966, pp. 26–42. English translation: On a heuristic point of view about the creation and conversion of light, in D. ter Haar: *The Old Quantum Theory*, Oxford–London–Edinburgh–New York–Toronto–Sydney–Paris–Braunschweig: Pergamon Press, 1967, pp. 91–107.

1909a Zum gegenwärtigen Stand des Strahlungsproblems, *Phys. Zs.* **10**, 185–193 (received 23 January 1909, published in issue No. 6 of 15 March 1909).

1916d Strahlungs-Emission und -Absorption nach der Quantentheorie, *Verh. d. Deutsch. Phys. Ges. (2)* **18**, 318–323 (received 17 June 1916, presented at the meeting of 21 July and published in issue No. 13/14 of 30 July 1916).

1916e Quantentheorie der Strahlung, *Mitteilungen der Physikalischen Gesellschaft Zürich* **16**, 47–62 (published in issue No. 18).

1923d Bietet die Feldtheorie Möglichkeiten zur Lösung des Quantenproblems, *Sitz.ber. Preuss. Akad. Wiss. (Berlin)*, pp. 359–364 (presented at the meeting of 13 December 1923).

1924c Quantentheorie des einatomigen idealen Gases, *Sitz.ber. Preuss. Akad. Wiss. (Berlin)*, pp. 261–267 (presented at the meeting of 10 July 1924).

1925a Quantentheorie des einatomigen idealen Gases. 2. Abhandlung, *Sitz.ber. Preuss. Akad. Wiss. (Berlin)*, pp. 3–14 (presented at the meeting of 8 January 1925).

1925b Quantentheorie des idealen Gases, *Sitz.ber. Preuss. Akad. Wiss. (Berlin)*, pp. 18–25 (presented at the meeting of 29 January 1925).

1925c Bemerkung zu P. Jordans Abhandlung "Zur Theorie der Quantenstrahlung," *Z. Phys.* **31**, 784–785 (received 22 January 1925, published in issue No. 10 of 21 March 1925).

EINSTEIN, ALBERT, HEDWIG BORN AND MAX BORN

1969 *Briefwechsel 1916–1955, kommentiert von Max Born*, Munich: Nymphenburger Verlagshandlung. English translation (by I. Born): *The Born–Einstein Letters. Correspondence between Albert Einstein and Max and Hedwig Born from 1916 to 1955, with commentaries by Max Born*, London and Basingstoke: The Macmillan Press Ltd., 1971.

EINSTEIN, ALBERT, AND PAUL EHRENFEST

1923 Zur Quantentheorie des Strahlungsgleichgewichts, *Z. Phys.* **19**, 301–306 (received 16 October 1923, published in issue No. 5/6 of 8 December 1923); reprinted in *Collected Scientific Papers* (Ehrenfest, 1959), pp. 485–490.

ELSASSER, WALTER

1925 Bemerkungen zur Quantenmechanik freier Elektronen, Naturwiss. **13**, 711 (letter dated 18 July 1925, published in the issue of 14 August 1925).

EPSTEIN, PAUL SOPHUS

1916a Zur Theorie des Starkeffekts, *Phys. Zs.* **17**, 148–150 (received 29 March 1916, published in issue No. 8 of 15 April 1916).

1916c Zur Theorie des Starkeffektes, *Ann. d. Phys. (4)* **50**, 489–521 (received 9 May 1916, published in issue No. 13 of 25 July 1916).

EULER, LEONHARD

1748 *Introductio in Analysim Infinitorum*, Vol. II, Lausanne: M. M. Bousquet & Socios.

FERMI, ENRICO

1926a Sulla quantizzazione del gaz perfetto monoatomico, *Rend. R. Accad. Lincei* **3**, 145–149 (presented at the meeting of 7 February 1926); reprinted in *Collected Papers (Note e Memorie) I*, (E. Amaldi, E. Persico, F. Rasetti and E. Segré, eds.), Chicago: The University of Chicago Press, 1962, pp. 181–185.

1926b Zur Quantelung des idealen einatomigen Gases, *Z. Phys.* **36**, 902–912 (received 24 March 1926, published in issue No. 11/12 of 11 May 1926); reprinted in *Collected Papers I*, pp. 186–195.

FÖPPL, LUDWIG

1912 Stabile Anordnungen von Elektronen im Atom, *J. reine & angew. Math.* **141**, 251–302 (from the doctoral dissertation, University of Göttingen, 1 March 1912).

FOURIER, JOSEPH

1822 *Théorie analytique de la chaleur*, Paris.

FRANCK, JAMES, AND PASCUAL JORDAN

1926 *Anregung von Quantensprüngen durch Stöße* (Vol. III of the series *Struktur der Materie in Einzeldarstellungen*, M. Born and J. Franck, eds.), Berlin: J. Springer Verlag.

FREDHOLM, IVAR

1900 Sur une nouvelle méthode pour la resolution du problème de Dirichlet, *Öfversight af Kung. Vetenskabs Akademie Förhandlingar (Stockholm)* **57**, 39–46 (published in the issue of 10 January 1900).

1903 Sur une classe d'équations fonctionelles, *Acta Mathematica* **27**, 365–390 (published in the issue of 30 March 1903).

FRENKEL, JAKOW

1926 Die Elektrodynamik des rotierenden Elektrons, *Z. Phys.* **37**, 243–262 (received 2 May 1926, published in issue No. 4/5 of 5 June 1926).

FROBENIUS, FERDINAND GEORG

1878 Über lineare Substitutionen und bilineare Formen, *J. reine & angew. Math.* **84**, 1–63 (dated May, 1877, published in issue No. 1); reprinted in *Gesammelte Abhandlungen 1*, pp. 343–405.

1879 Theorie der linearen Formen mit ganzen Koeffizienten, *J. reine & angew. Math.* **86**, 146–208 (dated April 1878, published in issue No. 2/3); reprinted in *Gesammelte Abhandlungen 1*, pp. 482–544.

GOLDSTEIN, HERBERT

1975 Prehistory of the "Runge-Lenz" vector, *Amer. J. Phys.* **43**, 737–738 (received 16 January 1975, published in issue No. 8 of August 1975).

1976 More on the prehistory of the Laplace or Runge-Lenz vector, *Amer. J. Phys.* **44**, 1123–1124 (received 16 December 1975, published in issue No. 11 of November 1976).

GOUDSMIT, SAMUEL, AND RALPH DE LAER KRONIG

1925 Die Intensität der Zeemankomponenten, *Naturwiss.* **12**, 90 (letter dated 17 December 1924, published in the issue of 30 January 1925).

GOUDSMIT, SAMUEL, AND GEORGE EUGENE UHLENBECK

1926 Die Kopplungsmöglichkeiten der Quantenvektoren im Atom, *Z. Phys.* **35**, 618–625 (received 27 November 1925, published in issue No. 8/9 of 4 February 1926).

GRAEFFE, KARL HEINRICH

1837 *Die Auflösung der höheren numerischen Gleichungen*, Zurich.

HAMILTON, WILLIAM ROWAN

1844 A new species of imaginary quantities connected with the theory of quaternions, *Proc. Roy. Irish Acad.* **2**, 424–434.

1853 *Lectures on Quaternions*, Dublin: Hodges and Smith.

HEAVISIDE, OLIVER

1893a On operators in physical mathematics. Part I, *Proc. Roy. Soc. (London)* **52**, 504–529 (received 15 December 1892, read at the meeting of 2 February 1893 and published in issue No. 320 of February 1893).

1893b On operators in physical mathematics. Part II, *Proc. Roy. Soc. (London)* **54**, 105–143 (received 8 June 1893, read at the meeting of 15 June 1893 and published in issue No. 326 of June 1893).

HEISENBERG, WERNER

1922a Zur Quantentheorie der Linienstruktur und der anomalen Zeemaneffekte, *Z. Phys.* **8**, 273–297 (received 17 December 1921, published in issue No. 5 of 15 February 1922).

1925a Über eine Anwendung des Korrespondenzprinzips auf die Frage nach der Polarisation des Fluoreszenzlichtes, *Z. Phys.* **31**, 617–626 (received 30 November 1924, published in issue No. 7/8 of 7 March 1925).

1925b Zur Quantentheorie der Multiplettstruktur und der anomalen Zeemaneffekte, *Z. Phys.* **32**, 841–860 (received 10 April 1925, published in issue No. 11/12 of 30 June 1925).

1925c Über die quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen, *Z. Phys.* **33**, 879–893 (received 29 July 1925, published in issue No. 12 of 18 September 1925); reprinted in *Begründung der Matrizenmechanik* (Born, Heisenberg and Jordan, 1962), pp. 31–46. English translation: Quantum-theoretical re-interpretation of kinematic and mechanical relations, in *Sources of Quantum Mechanics* (van der Waerden, 1967), pp. 261–276.

1926a Über quantenmechanische Kinematik und Mechanik, *Mathematische Annalen* **95**, 683–705 (received 21 December 1925).

1926b Mehrkörperproblem und Resonanz in der Quantenmechanik, *Z. Phys.* **38**, 411–426 (received 11 June 1926, published in issue No. 6/7 of 10 August 1926).

1926c Über die Spektra von Atomsystemen mit zwei Elektronen, *Z. Phys.* **39**, 499–518 (received 24 July 1926, published in issue No. 7/8 of 26 October 1926).

1926e Schwankungserscheinungen und Quantenmechanik, *Z. Phys.* **40**, 501–506 (received 6 November 1926, published in issue No. 7 of 20 December 1926).

1927b Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik, *Z. Phys.* **43**, 172–198 (received 23 March 1927, published in issue No. 3/4 of 31 May 1927).

1928 Zur Theorie des Ferromagnetismus, *Z. Phys.* **49**, 619–636 (received 20 May 1928, published in issue No. 9/10 of 16 July 1928).

1943a Die "beobachtbaren Größen" in der Theorie der Elementarteilchen, *Z. Phys.* **120**, 513–538 (received 8 September 1942, published in issue No. 7–10 of 25 March 1943).

1943b Die beobachtbaren Größen in der Theorie der Elementarteilchen. II., *Z. Phys.* **120**, 673–702 (received 30 October 1942, published in issue No. 11/12 of 6 April 1943).

1944 Die beobachtbaren Größen in der Theorie der Elementarteilchen. III., *Z. Phys.* **123**, 93–112 (received 12 May 1944, published in issue No. 1/2 of 10 October 1944).

HEISENBERG, WERNER, AND PASCUAL JORDAN

1926 Anwendung der Quantenmechanik auf das Problem der anomalen Zeemaneffekte, *Z. Phys.* **37**, 263–277 (received 16 March 1926, published in issue No. 4/5 of 5 June 1926).

HEITLER, WALTER, AND FRITZ LONDON

1927 Wechselwirkung neutraler Atome und homöopolare Bindung nach der Quantentheorie, *Z. Phys.* **44**, 455–472 (received 30 June 1927, published in issue No. 6/7 of 16 August 1927).

HELLINGER, ERNST

1907 *Die Orthogonalinvarianten quadratischer Formen von unendlichvielen Variablen*, doctoral dissertation, University of Göttingen.

1909 Neue Begründung der Theorie quadratischer Formen von unendlichvielen Veränderlichen, *J. reine & angew. Math.* **136**, 210–271 (*Habilitationsschrift*, University of Marburg, March 1909; published in issue No. 3/4).

HELLINGER, ERNST, AND OTTO TOEPLITZ

1906 Grundlagen für eine Theorie der unendlichen Matrizen, *Nachr. Ges. Wiss. Göttingen*, pp. 351–355 (presented by D. Hilbert at the meeting of 28 July 1906).

1910 Grundlagen für eine Theorie der unendlichen Matrizen, *Mathematische Annalen* **69**, 289–330.

1927 Integralgleichungen und Gleichungen mit unendlichvielen Unbekannten, *Encykl. d. math. Wiss. II/3, Part 2*, pp. 1335–1597 (published in issue No. 9 of 15 December 1927).

HELMHOLTZ, HERMANN VON

1898–1907 *Vorlesungen über Theoretische Physik* (A. König, O. Krigar-Menzel, M. Laue, F. Richarz, C. Runge, eds.), Leipzig: J. A. Barth.

I. Band, 1. Abteilung: Einleitung zu den Vorlesungen (A. König and C. Runge, eds.), 1903;

I. Band, 2. Abteilung: Dynamik diskreter Massenpunkte (O. Krigar-Menzel, ed.), 1898;

II. Band: Dynamik kontinuierlich verbreiteter Massen (O. Krigar-Menzel, ed.), 1902;

III. Band: Mathematische Prinzipien der Akustik (A. König and C. Runge, eds.), 1898;

IV. Band: Elektrodynamik und Theorie des Magnetismus (O. Krigar-Menzel and M. Laue, eds.), 1907;

V. Band: Elektromagnetische Theorie des Lichtes (A. König and C. Runge, eds.), 1897;

VI. Band: Theorie der Wärme (F. Richarz, ed.), 1903.

HERMITE, CHARLES

1854 Remarques sur une mémoire de M. Cayley relatif aux déterminants gauches, *Cambridge and Dublin Mathematical Journal* **9**, 63–67 (published in the issue of February 1854).

1855 Remarque sur un théorème de M. Cayley, *Comptes rendus (Paris)* **41**, 181–183 (letter to the meeting of 6 August 1855).

HERTZ, GUSTAV

1925 Das Spektrum des Neons im kurzwelligen Ultraviolett, *Naturwiss.* **13**, 489 (letter dated 1 May 1925, published in the issue of 29 July 1925).

HILBERT, DAVID

1904a Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen. (Erste Mitteilung.), *Nachr. Ges. Wiss. Göttingen*, pp. 49–91 (presented at the meeting of 5 March 1904).

1904b Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen. (Zweite Mitteilung.), *Nachr. Ges. Wiss. Göttingen*, pp. 213–259 (presented at the meeting of 25 June 1904).

1905 Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen. (Dritte Mitteilung.), *Nachr. Ges. Wiss. Göttingen*, pp. 307–338 (presented at the meeting of 22 July 1905).

1906a Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen. (Vierte Mitteilung.), *Nachr. Ges. Wiss. Göttingen*, pp. 157–227 (presented at the meeting of 3 March 1906).

1906b Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen. (Fünfte Mitteilung.), *Nachr. Ges. Wiss. Göttingen*, pp. 439–480 (presented at the meeting of 28 July 1906).

1910 Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen. (Sechste Mitteilung.), *Nachr. Ges. Wiss. Göttingen*, pp. 355–419 (presented at the meeting of 28 May 1910).

1912c *Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen*, Leipzig and Berlin: B. G. Teubner.

1915 Die Grundlagen der Physik. (Erste Mitteilung.), *Nachr. Ges. Wiss. Göttingen*, pp. 395–407 (presented at the meeting of 20 November 1915, published in issue No. 3 of 1915).

1917 Die Grundlagen der Physik. (Zweite Mitteilung.), *Nachr. Ges. Wiss. Göttingen*, pp. 53–76 (presented at the meeting of 23 December 1916, published in issue No. 1 of 1917).

HILBERT, DAVID, JOHANNES VON NEUMANN AND LOTHAR NORDHEIM

1928 Über die Grundlagen der Quantenmechanik, *Mathematische Annalen* **98**, 1–30 (received 6 April 1927).

HILL, GEORGE WILLIAM

1877 *On the Part of the Motion of the Lunar Perigee which is a Function of the Mean Motions of the Sun and Moon*, Cambridge (Mass.): Wilson; reprinted with some additions in *Acta Mathematica* **8**, 1–36 (1886).

1905 *The Collected Mathematical Papers, Volume 1*, Washington, D.C.: Carnegie Institution.

HÖNL, HELMUT

1925 Die Intensitäten der Zeemankomponenten, *Z. Phys.* **34**, 340–354 (received 26 November 1924, published in No. 1 of 11 February 1925).

HOPMANN, JOSEF

1923 Die Deutung der Ergebnisse der amerikanischen Einsteinexpedition, *Phys. Zs.* **24**, 476–484 (extended version of a paper presented at the Bonn meeting of the German Physical Society, 16–22 September 1923, published in issue No. 21/22 of 1/15 November 1923).

HUND, FRIEDRICH

1925c Zur Deutung verwickelter Spektren, insbesondere der Elemente Scandium bis Nickel, *Z. Phys.* **33**, 345–371 (received 22 June 1925, published in issue No. 5/6 of 8 August 1925).

1927 *Linienspektren und periodisches System der Elemente* (Vol. IV of the series *Struktur der Materie in Einzeldarstellungen*, M. Born and J. Franck, eds.), Berlin: J. Springer Verlag.

JORDAN, CAMILLE

1870 *Traité des substitutions et des équations algébriques*, Paris: Gauthier-Villars.

JORDAN, PASCUAL

1924 Zur Theorie der Quantenstrahlung, *Z. Phys.* **30**, 297–319 (received 30 November 1924, published in issue No. 4/5 of 29 December 1924).

1925a Bemerkungen zu einer Arbeit von K. W. Meissner: Absorption im angeregten Neon, *Z. Phys.* **31**, 877–881 (received 23 February 1925, published in issue No. 11 of 28 March 1925).

1925b Zur Quantentheorie der aperiodischen Vorgänge. II. Bemerkungen über die Integration der Störungsgleichungen, *Z. Phys.* **33**, 506–508 (received 11 June 1925, published in issue No. 7 of 15 August 1925).

1925c Bemerkung zur Theorie der Atomstruktur, *Z. Phys.* **33**, 563–570 (received 8 July 1925, published in issue No. 8 of 18 August 1925).

1925d Über das thermische Gleichgewicht zwischen Quantenatomen und Hohlraumstrahlung, *Z. Phys.* **33**, 649–655 (received 8 July 1925, published in issue No. 9 of 22 August 1925).

1926 Bemerkungen über einen Zusammenhang zwischen Duanes Quantentheorie der Interferenz und den de Broglieschen Wellen, *Z. Phys.* **37**, 376–382 (received 22 April 1926, published in issue No. 4/5 of 5 June 1926).

1928 Die Lichtquantenhypothese. Entwicklung und gegenwärtiger Stand, *Ergeb. d. exakt. Naturwiss.* **7**, 158–208.

KEMBLE, EDWIN CRAWFORD

1926b *Problems of Atomic Dynamics. (Lectures delivered at the M.I.T.)*. Max Born, *Phys. Rev. (2)* **28**, 423–424 (book review, published in issue No. 2 of August 1926).

KLEIN, OSKAR

1924a The simultaneous action on a hydrogen atom of crossed homogeneous electric and magnetic fields, *Phys. Rev. (2)* **23**, 308 (abstract of a paper presented at the Cincinnati meeting of the American Physical Society, 27–29 December 1923, published in issue No. 2 of February 1924).

1924b Über die gleichzeitige Wirkung von gekreuzten homogenen elektrischen und magnetischen Feldern auf das Wasserstoffatom. I., *Z. Phys.* **22**, 109–118 (dated 23 December 1923, received 23 January 1924, published in No. 1/2 of 6 March 1924).

KNOTT, CARGILL GILSTON

1911 *Life and Scientific Work of Peter Guthrie Tait*, Cambridge: Cambridge University Press.

KOCH, HELGE VON

1892 Sur les déterminants infinis et les équations differentielles linéaires, *Acta Mathematica* **16**, 217–295.

KÖTTERITZSCH, TH.

1870 Über die Auflösung eines Systems von unendlich vielen linearen Gleichungen, *Zeitschrift für Mathematik und Physik* **15**, 1–15, 229–268.

KRAMERS, HENDRIK ANTHONY

1924a The law of dispersion and Bohr's theory of spectra, *Nature* **113**, 673–674 (letter dated 25 March 1924, published in the issue of 10 May 1924); reprinted in *Collected Scientific Papers*, pp. 290–291, and in *Sources of Quantum Mechanics* (van der Waerden, 1967), pp. 177–180.

1924b The quantum theory of dispersion, *Nature* **114**, 310–311 (letter dated 22 July 1924, published in the issue of 30 August 1924); reprinted in *Collected Scientific Papers*, p. 292, and in *Sources of Quantum Mechanics* (van der Waerden, 1967), pp. 199–201.

1925b On the behaviour of atoms in an electromagnetic wave field, *6e Skand. Mat. Kongress*, pp. 145–153; reprinted in *Collected Scientific Papers*, pp. 321–331.

1925c Eenige Opmerkingen over de Quantenmechanica van Heisenberg, *Physica*, **5**, 369–376 (dated November 1925, published in issue No. 11/12 of November–December 1925).

1956 *Collected Scientific Papers*, Amsterdam: North Holland Publishing Company.

KRAMERS, HENDRIK ANTHONY, AND WERNER HEISENBERG

1925 Uber die Streuung von Strahlung durch Atome, *Z. Phys.* **31**, 681–708 (received 5 January 1925, published in issue No. 9 of 17 March 1925); reprinted in *Collected Scientific Papers* (Kramers, 1956), pp. 293–320. English translation: On the dispersion of radiation by atoms, in *Sources of Quantum Mechanics* (van der Waerden, 1967), pp. 223–251.

KRAMERS, HENDRIK ANTHONY, AND WOLFGANG PAULI

1923 Zur Theorie der Bandenspektren, *Z. Phys.* **13**, 351–367 (received 3 January 1923, published in issue No. 6 of 26 February 1923); reprinted in *Collected Scientific*

Papers (Kramers, 1956), pp. 231–247, and in *Collected Scientific Papers 2* (Pauli, 1964), pp. 134–150.

KRATZER, ADOLF

1922 Störungen und Kombinationsprinzip im System der violetten Cyanbanden, *Sitz.-ber. Bayer. Akad. Wiss. (München)*, pp. 107–118 (presented by A. Sommerfeld at the meeting of 4 March 1922).

KRONIG, RALPH DE LAER

1926d The dielectric constant of diatomic dipole gases on the new quantum mechanics, *Proc. Nat. Acad. Sci. (U.S.A.)* **12**, 488–493 (communicated 3 July 1926, published in issue No. 8 of 15 August 1926).

KUHN, WERNER

1925a Über die Gesamtstärke der von einem Zustande ausgehenden Absorptionslinien, *Z. Phys.* **33**, 408–412 (received 14 May 1924, published in issue No. 5/6 of 8 August 1925). English translation: On the total intensity of absorption lines emanating from a given state, in *Sources of Quantum Mechanics* (van der Waerden, 1967), pp. 253–257.

LADENBURG, RUDOLF

1926b Die quantentheoretische Dispersionsformel und ihre experimentelle Prüfung, *Naturwiss.* **14**, 1208–1213 (published in the issue of 10 December 1926).

LAGRANGE, JOSEPH LOUIS, COMTE DE

1772 Sur une nouvelle espèce de calcul relatif à la differentiation et à l'intégration des quantités variables, *Nouvelles Mémoires de l'Académie de Berlin*, p. 441ff; reprinted in *Oeuvres* **3**, pp. 441–476.

LANCZOS, CORNELIUS

1925a Zum Problem der unendlich schwachen Felder in der Einsteinschen Gravitationstheorie, *Z. Phys.* **31** 112–132 (received 1 November 1924, published in issue No. 1–4 of 11 February 1925).

1925b Zum Problem der Ätherstrahlung in einer räumlich geschlossenen Welt, *Z. Phys.* **32**, 135–149 (received 26 February 1925, published in issue No. 2 of 28 April 1925).

1925c Stationäre Elektronenbahnen und die Methode der Eigenfunktionen, *Z. Phys.* **33**, 128–143 (received 20 May 1925, published in issue No. 1/2 of 18 July 1925).

1926a Über eine feldmäßige Darstellung der neuen Quantenmechanik, *Z. Phys.* **35**, 812–830 (received 22 December 1925, published in issue No. 11/12 of 26 February 1926).

1926b Variationsprinzip und Quantenbedingung in der neuen Quantenmechanik, *Z. Phys.* **36**, 401–409 (received 1 March 1926, published in issue No. 6 of 1 April 1926).

1926c Über die komplexe Beschaffenheit der quantenmechanischen Matrizen, *Z. Phys.* **37**, 405–413 (received 28 April 1926, published in issue No. 6 of 10 June 1926).

1929 Die Erhaltungssätze in der feldmäßigen Darstellung der Diracschen Theorie, *Z. Phys.* **57**, 484–493 (received 13 August 1929, published in issue No. 7/8 of 10 September 1929).

1930 Zur Theorie des Starkeffekts in hohen Feldern, *Z. Phys.* **62**, 518–544 (received 7 April 1930, published in issue No. 7/8 of 18 June 1930).

LANDAU, LEV DAVIDOVICH

1926 Zur Theorie der Spektren der zweiatomigen Moleküle, *Z. Phys.* **40**, 621–627 (received 13 November 1926, published in issue No. 8 of 22 December 1926).

LANDÉ, ALFRED

1923a Termstruktur und Zeemaneffekt der Multipletts, *Z. Phys.* **15**, 189–205 (received 5 March 1923, published in issue No. 4/5 of 26 May 1923).

1924a Das Wesen des relativistischen Röntgendubletts, *Z. Phys.* **24**, 88–97 (received 15 March 1924, published in issue No. 2 of 22 May 1924).

LANDÉ, ALFRED, AND WERNER HEISENBERG

1924 Termstuktur der Multipletts höherer Stufe, *Z. Phys.* **25**, 279–286 (received 18 May 1924, published in issue No. 4–6 of 4 August 1924).

LAUE, MAX VON

1911 *Das Relativitätsprinzip* (Volume 38 of the series *Die Wissenschaft*), Braunschweig: Fr. Vieweg & Sohn.

LEBESGUE, HENRI

1924a Conditions de régularité, conditions d'irrégularité, conditions d'impossibilité dans le problème de Dirichlet, *Comptes rendus (Paris)* **178**, 349–354 (presented at the meeting of 21 January 1924).

1924b Observations au sujet de la Note de M. N. Wiener, *Comptes rendus (Paris)* **178**, 1053–1054 (presented at the meeting of 24 March 1924).

LEIBNIZ, GOTTFRIED WILHELM

1693 Letter to G. F. A. de l'Hospital of 29 April 1693, published in Leibniz: *Mathematische Schriften, Erste Abteilung, II* (C. J. Gerhard, ed.), Berlin, 1850, pp. 236–241.

1710 Symbolismus memorabilis calculi algebraici et infinitesimalis in comparatione potentiarum et differentiarum, *Berolina Miscellania* **1**, 160ff.

LENZ, WILHELM

1924 Über den Bewegungsablauf und die Quantenzustände der gestörten Keplerbewegung, *Z. Phys.* **24**, 197–207 (received 2 April 1924, published in issue No. 3/4 of 4 June 1924).

LICHTENSTEIN, LEON

1924 Neuere Entwicklung der Theorie partieller Differentailgleichungen zweiter Ordnung vom elliptischen Typus, *Encykl. d. math. Wiss. II/3, Part 2*, pp. 1279–1334 (dated March 1924, published in issue No. 8 of 10 September 1924).

LONDON, FRITZ

1926a Energiesatz und Rydbergprinzip in der Quantenmechanik, *Z. Phys.* **36**, 775–777 (received 17 March 1926, published in issue No. 9/10 of 30 April 1926).

LYMAN, THEODORE

1924 The spectrum of helium in the extreme ultra-violet, *Astrophysical Journal* **60**, 1–14 (dated 5 May 1924, published in issue No. 1 of July 1924).

MACH, ERNST

1883 *Die Mechanik in ihrer Entwicklung*, Leipzig: F. A. Brockhaus (seventh edition 1912). English translation: *The Science of Mechanics*, Chicago, 1893.

1886 *Beiträge zur Analyse der Empfindungen*, Jena: Gustav Fischer.

1896 *Die Prinzipien der Wärmelehre*, Liepzig: J. A. Barth.

1921 *Prinzipien der physikalischen Optik, historisch und erkenntnis-psychologisch dargestellt*, Leipzig: J. A. Barth.

MADELUNG, ERWIN

1925 *Die mathematischen Hilfsmittel des Physikers*, second edition, Berlin: J. Springer Verlag.

MANNEBACK, CHARLES

1926 Die Dielektrizitätskonstante der zweiatomigen Dipolgase nach der Wellenmechanik, *Phys. Zs.* **27**, 563–569 (received 20 July 1926, published in issue No. 17 of 1 September 1926).

MAXWELL, JAMES CLERK

1873 *A Treatise on Electricity and Magnetism*, 2 volumes, Oxford: Oxford University Press.

MEHRA, JAGDISH (EDITOR)

1973b *The Physicist's Conception of Nature*, Dordrecht: D. Reidel Publishing Company.

MEISSNER, KARL W.

1925a Absorption in angeregtem Neon, *Ann. d. Phys. (4)* **76**, 124–144 (received 18 September 1924, published in issue No. 2/3 of January 1925, dedicated to F. Paschen).

1925b Absorption in angeregten Gasen, *Phys. Zs.* **26**, 687–689 (presented on 15 September 1925 at the meeting of the German Physical Society, Danzig, published in issue No. 19 of 1 October 1925).

MENSING, LUCIE (LUCY)

1925a Zur Störungsmechanik der Molekularmodelle, *Z. Phys.* **34**, 602–610 (received 1 September 1925, published in issue No. 8 of 2 November 1925).

1925b Beitrag zur Theorie der Verbreiterung von Spektrallinien, *Z. Phys.* **34**, 611–621 (received 1 September 1925, published in issue No. 8 of 2 November 1925).

1926a Die Rotations-Schwingungsbanden nach der Quantenmechanik, *Z. Phys.* **36**, 814–823 (received 29 March 1926, published in issue No. 11/12 of 11 May 1926).

1926b Die Intensitäten der Zeemankomponenten beim partiellen Paschen-Back-Effekt, *Z. Phys.* **39**, 24–28 (received 27 July 1926, published in issue No. 1 of 21 September 1926).

MENSING, LUCIE (LUCY), and WOLFGANG PAULI

1926 Über die Dielektrizitätskonstante von Dipolgasen nach der Quantenmechanik, *Phys. Zs.* **27**, 509–512 (received 25 June 1926, published in issue No. 15 of 1 August 1926); reprinted in *Collected Scientific Papers 2* (Pauli, 1964), pp. 280–283.

MIE, GUSTAV

1912a Grundlagen einer Theorie der Materie. Erste Mitteilung., *Ann. d. Phys. (4)* **37**, 511–534 (received 9 January 1912, published in issue No. 3 of 8 March 1912).

1912c Grundlagen einer Theorie der Materie. Dritte Mitteilung., *Ann. d. Phys. (4)* **40**, 1–66 (received 2 November 1912, published in issue No. 1 of 31 December 1912).

MILLIKAN, ROBERT ANDREWS, AND IRA SPRAGUE BOWEN

1925a The significance of the discovery of X-ray laws in the field of optics, *Proc. Nat. Acad. Sci. (U.S.A.)* **11**, 119–122 (communicated 19 December 1924, published in issue No. 2 of 15 February 1925).

MINKOWSKI, HERMANN

1908 Die Grundgleichungen für die elektromagnetischen Vorgänge in bewegten Körpern, *Nachr. Ges. Wiss. Göttingen*, pp. 53–111 (presented at the meeting of 21 December 1907); reprinted in *Mathematische Annalen* **68**, 472–525 (1910).

MÜLLER, CONRAD HEINRICH, AND GEORG PRANGE

1923 *Allgemeine Mechanik*, Hanover: Helwigsche Verlagsbuchhandlung.

MURPHY, ROBERT

1837 First memoir on the theory of analytical operations, *Phil. Trans. Roy. Soc. (London)* **127**, 179–210 (received 8 December 1836, read 22 December 1836, published 1837).

NERNST, WALTHER, AND ARTHUR SCHOENFLIES

1895 *Einführung in die mathematische Behandlung der Naturwissenschaften: Kurzgefaßtes Lehrbuch der Differential- und Integralrechnung mit besonderer Berücksichtigung der Chemie*, Munich and Berlin: R. Oldenburg (ninth edition 1919).

NEUMANN, JOHANNES VON

1927a Mathematische Begründung der Quantenmechanik, *Nachr. Ges. Wiss. Göttingen*, pp. 1–57 (presented at the meeting of 20 May 1927); reprinted in *Collected Works 1*, pp. 151–207.

1927b Wahrscheinlichkeitstheoretischer Aufbau der Quantenmechanik, *Nachr. Ges. Wiss. Göttingen*, pp. 245–272 (presented at the meeting of 11 November 1927); reprinted in *Collected Works 1*, pp. 208–235.

1961 *Collected Works, Volume 1*, London-New York-Paris-Frankfurt/Main: Pergamon Press.

NÖRLUND, NIELS ERIK

1924 *Vorlesungen über Differenzenrechnung*, Berlin: J. Springer Verlag.

PALEY, R. E. C. A., AND NORBERT WIENER

1934 *Fourier Transforms in the Complex Domain* (American Mathematical Society Publications, Vol. 19), Providence, R.I.: American Mathematical Society.

PASCHEN, FRIEDRICH, AND RAIMUND GÖTZE

1922 *Seriengesetze der Linienspektren*, Berlin: J. Springer Verlag.

PAULI, WOLFGANG

1921a Zur Theorie der Dielektrizitätskonstante zweiatomiger Dipolgase, *Z. Phys.* **6**, 319–327 (received 30 July 1921, published in issue No. 5/6 of 3 October 1921); reprinted in *Collected Scientific Papers 2*, pp. 39–47.

1921b Relativitätstheorie, *Encykl. d. math. Wiss. V/2*, pp. 539–775 (dated December 1920, published in issue No. 4 of 15 September 1921); also published as a separate book *Relativitätstheorie*, with a preface by A. Sommerfeld, Leipzig: B. G. Teubner, 1921; reprinted in *Collected Scientific Papers 1*, pp. 1–237. English translation: *Theory of Relativity* (with supplementary notes), New York: Pergamon Press Inc., 1958.

1923b Über das thermische Gleichgewicht zwischen Strahlung und freien Elektronen, *Z. Phys.* **18**, 272–286 (received 9 August 1923, published in issue No. 5 of 18 October 1923); reprinted in *Collected Scientific Papers 2*, pp. 161–175.

1924b Bemerkungen zu den Arbeiten "Die Dimensionen der Einsteinschen Lichtquanten" und "Zur Dynamik des Stoßes zwischen einem Lichtquant und einem Elektron" von L. S. Ornstein und H. C. Burger, *Z. Phys.* **22**, 261–265 (received 23 January 1924, published in issue No. 4 of 17 March 1924); reprinted in *Collected Scientific Papers 2*, pp. 193–197.

1924c Zur Frage der theoretischen Deutung der Satelliten einiger Spektrallinien und ihrer Beeinflussung durch magnetische Felder, *Naturwiss.* **12**, 741–743 (letter dated 17 August 1924, published in the issue of 12 September 1924); reprinted in *Collected Scientific Papers 2*, pp. 198–200.

1925a Über den Einfluß der Geschwindigkeitsabhängigkeit der Elektronenmasse auf den Zeemaneffekt, *Z. Phys.* **31**, 373–385 (received 2 December 1924, published in issue No. 5/6 of 19 February 1925); reprinted in *Collected Scientific Papers 2*, pp. 201–213.

1925b Über den Zusammenhang des Abschlusses der Elektronengruppen im Atom mit der Komplexstruktur der Spektren, *Z. Phys.* **31**, 765–783 (received 16 January 1925, published in issue No. 10 of 21 March 1925); reprinted in *Collected Scientific Papers 2*, pp. 214–232. English translation: On the connection between the completion of electron groups in an atom with the complex structure of spectra, in D. ter Haar: *The Old Quantum Theory*, Oxford–London–Edinburgh–New York–Toronto–Sydney–Paris–Braunschweig: Pergamon Press, 1967, pp. 184–203.

1925d Über die Intensitäten der im elektrischen Feld erscheinenden Kombinationslinien, *Kgl. Danske Vid. Selsk. Math.-fys. Medd.* **7**, No. 3 (ready for printing 9 November 1925); reprinted in *Collected Scientific Papers 2*, pp. 233–250.

1926a Über das Wasserstoffspektrum vom Standpunkt der neuen Quantenmechanik, *Z. Phys.* **36**, 336–363 (received 17 January 1926, published in issue No. 5 of 27 March 1926); reprinted in *Collected Scientific Papers 2*, pp. 252–279. English

translation: On the hydrogen spectrum from the standpoint of the new quantum mechanics, in *Sources of Quantum Mechanics* (van der Waerden, 1967), pp. 387–415.

1926b Quantentheorie, in *Handbuch der Physik* (H. Geiger and K. Scheel, eds.), *Vol. 23, Part 1*, Berlin: J. Springer Verlag, pp. 1–278; reprinted in *Collected Scientific Papers 1*, pp. 269–548.

1927 Über Gasentartung und Paramagnetismus, *Z. Phys.* **41**, 81–102 (received 16 December 1926, published in issue No. 1 of 26 January 1927); reprinted in *Collected Scientific Papers 2*, pp. 284–305.

1933c Die allgemeinen Prinzipien der Wellenmechanik, in *Handbuch der Physik* (H. Geiger and K. Scheel, eds.), *Vol. 24, Part 1*, second edition, Berlin: J. Springer Verlag, pp. 83–272.

1946 Remarks on the history of the exclusion principle, *Science* **103**, 213–215 (amplified address, given on 10 December 1945 at the Institute for Advanced Study, Princeton, during a dinner honouring the award of the 1945 Nobel prize in physics to Pauli); reprinted in *Collected Scientific Papers 2*, pp. 1073–1075.

1955 Exclusion principle, Lorentz group and reflection of space-time and charge, in *Niels Bohr and the Development of Physics* (W. Pauli, ed.), London: Pergamon Press, Inc., pp. 30–51; reprinted in *Collected Scientific Papers 2*, pp. 1239–1260.

1964 *Collected Scientific Papers* (R. Kronig and V. F. Weisskopf, eds.), 2 volumes, New York–London–Sydney: Interscience Publishers.

PEACOCK, GEORGE

1834 Report on the recent progress and present state of certain branches of analysis, *Reports Brit. Assoc. Adv. Sci.* **3**, 185–352.

PINCHERLE, SALVATORE

1906 Funktionaloperatoren und -gleichungen, *Encykl. d. math. Wiss. II/1, Part 2*, pp. 761–817 (dated December 1905, published in issue No. 6 of 27 March 1906).

PLANCK, MAX

1906 *Vorlesungen über die Theorie der Wärmestrahlung*, Leipzig: J. A. Barth.

POINCARÉ, HENRI

1885 Remarque sur l'emploi de la méthode précédente, *Bulletin de la Societé Mathématique de France* **13**, 19–27 (presented at the meeting of 19 December 1884).

1886 Sur les déterminants d'ordre infini, *Bulletin de la Societé Mathématique de France* **14**, 77–90 (presented at the meeting of 17 February 1886).

1892, 1893, 1899 *Les Méthodes Nouvelles de la Mechanique Céleste. Tome I: Solutions Périodiques.—Non-Existence des Intégrales Uniformes. Solutions Asymptotiques; Tome II: Méthodes de M. M. Newcomb, Lindstedt & Bohlin; Tome III: Invariants Intégraux.—Solutions Périodiques du Deuxième Genre. Solutions Doublement Asymptotiques*, Paris: Gauthier-Villars; paperback edition New York: Dover Publications Inc., 1957.

1894 Sur l'équations de la physique mathématique, *Rend. Circ. Mat. Palermo* **8**, 57–156 (dated March 1894, published in part 1ª of 3 April 1894).

1895 Sur la méthode de Neumann et le problème de Dirichlet, *Comptes rendus (Paris)* **120**, 347–352 (presented at the meeting of 18 February 1895).

1910 *Sechs Vorträge über ausgewählte Gegenstände der reinen Mathematik und der mathematischen Physik*, (Wolfskehl lectures at Göttingen, 22–28 April 1909), Leipzig & Berlin: B. G. Teubner.

RAMSAUER, CARL

1921b Über den Wirkungsquerschnitt der Edelgase gegenüber langsamen Elektronen, *Phys. Zs.* **22**, 613–615 (presented at the Jena meeting of the German Physical Society, 18–24 September 1921, published in issue No. 21/22 of 1/15 November 1921).

1922 Über den Wirkungsquerschnitt der Gasmoleküle gegenüber langsamen Elektronen. I. Fortsetzung., *Ann. d. Phys. (4)* **66**, 546–558 (received 22 December 1921, published in issue No. 24 of 22 March 1922).

1923 Über den Wirkungsquerschnitt der Gasmoleküle gegenüber langsamen Elektronen. II. Fortsetzung und Schluß., *Ann. d. Phys. (4)* **72**, 345–352 (received 26 April 1923, published in issue No. 21 of 25 September 1923).

RASCHEVSKY, N. VON

1923 Die Relativitätstheorie als eine der möglichen Darstellungen der physikalischen Erscheinungen, *Phys. Zs.* **24**, 438–441 (presented on 19 September 1923 at the Bonn meeting of the German Physical Society, published in issue No. 20 of 15 October 1923).

RAYLEIGH, JOHN WILLIAM STRUTT, LORD

1880 On the resultant of a large number of vibrations of the same pitch and of arbitrary phase, *Phil. Mag. (5)* **10**, 73–78 (dated June 1880, published in issue No. 60 of August 1880).

REID, CONSTANCE

1970 *Hilbert*, Berlin–Heidelberg–New York: Springer-Verlag.

1976 *Courant in Göttingen and New York: The Story of an Improbable Mathematician*, New York–Berlin–Heidelberg–Paris: Springer-Verlag.

ROSANES, JAKOB

1904 Charakteristische Züge in der Mathematik des 19. Jahrhunderts, Jahresbericht der *Deutschen Mathematiker Vereinigung* **13**, 17–30 (rector's address delivered on 15 October 1903 at the University of Breslau, published in issue No. 1).

RUNGE, CARL

1919 *Vektoranalysis*, Leipzig: S. Hirzel

SCHLICK, MORITZ

1917 *Raum und Zeit in der gegenwärtigen Physik: Zur Einführung in das Verständnis der allgemeinen Relativitätstheorie*, Berlin: J. Springer Verlag.

SCHMIDT, ERHARD

1905 *Entwicklung willkürlicher Funktionen nach Systemen vorgeschriebener*, doctoral dissertation, University of Göttingen.

1908 Über die Auflösung linearer Gleichungen mit unendlich vielen Unbekannten, *Rend. Circ. Mat. Palermo* **25**, 53–77.

SCHRÖDINGER, ERWIN

1926c Quantisierung als Eigenwertproblem. (Erste Mitteilung.), *Ann. d. Phys. (4)* **79**, 361–376 (received 27 January 1926, published in issue No. 4 of 13 March 1926). English translation: Quantization as a problem of proper values. (Part I), in *Collected Papers on Wave Mechanics*, pp. 1–12.

1926d Quantisierung als Eigenwertproblem. (Zweite Mitteilung.), *Ann. d. Phys. (4)* **79**, 489–527 (received 23 February 1926, published in issue No. 6 of 6 April 1926). English translation: Quantization as a problem of proper values. (Part II), in *Collected Papers on Wave Mechanics*, pp. 13–40.

1926e Über das Verhältnis der Heisenberg-Born-Jordanschen Quantenmechanik zu der meinen, *Ann. d. Phys. (4)* **79**, 734–756 (received 18 March 1926, published in issue No. 8 of 4 May 1926). English translation: On the relation between the quantum mechanics of Heisenberg, Born and Jordan and that of Schrödinger, in *Collected Papers on Wave Mechanics*, pp. 45–61.

1926f Quantisierung als Eigenwertproblem. (Dritte Mitteilung: Störungstheorie, mit Anwendung auf den Starkeffekt der Balmerlinien.), *Ann. d. Phys. (4)* **80**, 437–490 (received 10 May 1926, published in issue No. 13 of 13 July 1926). English translation: Quantization as a problem of proper values. (Part III) Perturbation theory, with application to the Stark effect of the Balmer lines, in *Collected Papers on Wave Mechanics*, pp. 62–101.

1928 *Collected Papers on Wave Mechanics* (translated by J. F. Shearer and W. M. Deans), Glasgow: Blackie & Son.

SCHUR, ISSAI

1909 Über die charakteristischen Wurzeln einer linearen Substitution mit einer Anwendung auf die Theorie der Integralgleichungen, *Mathematische Annalen* **66**, 488–510 (published in No. 4 of 27 February 1909).

SCHUSTER, ARTHUR

1894 On interference phenomena, *Phil. Mag. (5)* **37**, 509–545 (published in issue No. 229 of June 1894).

1897 On lunar and solar periodicities of earthquakes, *Proc. Roy. Soc. (London)* **61**, 455–465 (received 18 May 1897, read 18 June 1897 and published in issue No. 377 of October 1897).

SCHWARZSCHILD, KARL

1916 Zur Quantenhypothese, *Sitz.ber. Preuss. Akad. Wiss. (Berlin)*, pp. 548–568 (presented at the meeting of 30 March 1916).

SERVOIS, FRANÇOIS JOSEPH

1814 Essai sur un nouveau mode d'exposition des principes du calcul différentiel, *Annales de Mathématique (Gergonne)* **5**, 93–114.

SMEKAL, ADOLF

1923c Zur Quantentheorie der Dispersion, *Naturwiss.* **11**, 873–875 (letter dated 15 September 1923, published in the issue of 26 October 1923).

SOMMERFELD, ARNOLD

1910b Zur Relativitätstheorie. I. Vierdimensionale Vektoralgebra, *Ann. d. Phys. (4)* **32**, 749–776 (received 17 March 1910, published in issue No. 9 of 21 June 1910); reprinted in *Gesammelte Schriften II*, pp. 189–216.

1910c Zur Relativitätstheorie. II. Vierdimensionale Vektoranalysis, *Ann. d. Phys. (4)* **33**, 649–689 (received 13 July 1910, published in issue No. 14 of 25 October 1910); reprinted in *Gesammelte Schriften II*, pp. 217–257.

1915b Zur Theorie der Balmerschen Serie, *Sitz.ber. Bayer. Akad. Wiss. (München)*, pp. 425–458 (presented at the meeting of 6 December 1915).

1915c Die Feinstruktur der Wasserstoff- und der Wasserstoff-ähnlichen Linien, *Sitz.ber. Bayer. Akad. Wiss. (München)*, pp. 459–500 (presented at the meeting of 6 January 1916, published in the volume of 1915).

1916d Zur Theorie des Zeeman-Effekts der Wasserstofflinien mit einem Anhang über den Stark-Effekt, *Phys. Zs.* **17**, 491–507 (received 7 September 1916, published in issue No. 20 of 15 October 1916); reprinted in *Gesammelte Schriften III*, pp. 309–325.

1919 *Atombau und Spektrallinien*, Braunschweig: Fr. Vieweg & Sohn.

1922a Quantentheoretische Umdeutung der Voigtschen Theorie des anomalen Zeeman-Effektes vom D-Linientypus, *Z. Phys.* **8**, 257–272 (received 12 December 1921, published in issue No. 5 of 15 February 1922); reprinted in *Gesammelte Schriften III*, pp. 609–624.

1968 *Gesammelte Schriften*, 4 volumes (*I–IV*), Braunschweig: Fr. Vieweg & Sohn.

SOMMERFELD, ARNOLD, AND ALBRECHT UNSÖLD

1926 Über das Spektrum des Wasserstoffs, *Z. Phys.* **36**, 259–275 (received 5 February 1926, published in issue No. 4 of 22 March 1926); reprinted in *Gesammelte Schriften III* (Sommerfeld, 1968), pp. 770–786.

SOMMERFELD, ARNOLD, AND GREGOR WENTZEL

1921 Über reguläre und irreguläre Dubletts, *Z. Phys.* **7**, 86–92 (received 11 August 1921, published in issue No. 2 of 21 October 1921); reprinted in *Gesammelte Schriften III* (Sommerfeld, 1968), pp. 594–600.

STONER, EDMUND CLIFTON

1924 The distribution of electrons among atomic levels, *Phil. Mag. (6)* **48**, 719–736 (dated July 1924, published in No. 286 of October 1924).

STRUIK, DIRK JAN, AND NORBERT WIENER

1927 Quantum theory and gravitational relativity, *Nature* **119**, 853–854 (letter dated 2 April 1927, published in issue of 11 June 1927).

SYLVESTER, JAMES JOSEPH

1850 Additions to the articles in the September number of this Journal "On a new class of theorems" and on Pascal's theorem, *Phil. Mag. (3)* **37**, 363–370 (published in issue No. 251 of November 1850).

1851 On the general theory of associated algebraical forms, *Cambridge and Dublin Mathematical Journal* **6**, 289–293.

1884 The genesis of an idea, or story of a discovery relating to equations in multiple quantity, *Nature* **31**, 35–36 (letter dated 26 October 1884, published in the issue of 13 November 1884).

TAMM, IGOR

1926 Zur Quantenmechanik des Rotators, *Z. Phys.* **37**, 685–698 (received 23 April 1926, published in issue No. 9 of 21 June 1926).

THIRRING, HANS

1928 Die Grundgedanken der neuen Quantentheorie. Erster Teil: Die Entwicklung bis 1926, *Ergeb. d. exakt. Naturwiss.* **7**, 384–431.

THOMAS, LLEWELLYN HILLETH

1926 The motion of the spinning electron, *Nature* **117**, 514 (letter dated 20 February 1926, published in the issue of 10 April 1926).

TOEPLITZ, OTTO

1907a Die Jacobische Transformation der quadratischen Formen von unendlichvielen Veränderlichen, *Nachr. Ges. Wiss. Göttingen*, pp. 101–109 (presented by D. Hilbert at the meeting of 23 February 1907).

1907b Zur Transformation der Scharen bilinearer Formen von unendlichvielen Veränderlichen, *Nachr. Ges. Wiss. Göttingen*, pp. 110–115 (presented by D. Hilbert at the meeting of 23 February 1907).

1911 Zur Theorie der quadratischen und bilinearen Formen von unendlichvielen Veränderlichen. I. Teil: Theorie der *L*-Formen, *Mathematische Annalen* **70**, 351–376.

UHLENBECK, GEORGE EUGENE, AND SAMUEL GOUDSMIT

1925 Ersetzung der Hypothese vom unmechanischen Zwang durch eine Forderung bezüglich des inneren Verhaltens jedes einzelnen Elektrons, *Naturwiss.* **13**, 953–954 (letter dated 17 October 1925, published in the issue of 20 November 1925).

1926 Spinning electrons and the structure of spectra, *Nature* **117**, 264–265 (letter dated December 1925, published in the issue of 20 February 1926).

VAN DER WAERDEN, BARTEL LEENDERT

1967 *Sources of Quantum Mechanics*, Amsterdam: North-Holland Publishing Company; paperback reprint in the *Classics of Science Series, Vol. V*, New York: Dover Publications, Inc.

VAN VLECK, JOHN HASBROUCK

1924a The absorption of radiation by multiply periodic orbits, and its relation to the

correspondence principle and the Rayleigh–Jeans law. Part I. Some extensions of the correspondence principle, *Phys. Rev. (2)* **24**, 330–346 (dated 19 June 1924, published in issue No. 4 of October 1924); reprinted in *Sources of Quantum Mechanics* (van der Waerden, 1967), pp. 203–222.

1924b The absorption of radiation by multiply periodic orbits, and its relation to the correspondence principle and the Rayleigh–Jeans law. Part II. Calculations of absorption by multiply periodic orbits, *Phys. Rev. (2)* **24**, 347–365 (dated 19 June 1924, published in issue No. 4 of October 1924).

1926a Magnetic susceptibilities and dielectric constants in the new quantum mechanics, *Nature* **118**, 226–227 (published in the issue of 14 August 1926).

WEIERSTRASS, KARL

1868 Zur Theorie der bilinearen und quadratischen Formen, *Monatsberichte der Preussischen Akademie der Wissenschaften (Berlin)*, pp. 310–338; reprinted in *Mathematische Werke 2*, pp. 19–44.

WEISS, PIERRE

1907 L'hypothèse du champ moléculaire et la propriété ferromagnétique, *J. phys. (Paris) (4)* **6**, 661–690 (presented on 4 April 1907 before the French Physical Society, published in issue No. 9 of September 1907).

WENTZEL, GREGOR

1921 Zur Systematik der Röntgenspektren, *Z. Phys.* **6**, 84–99 (received 5 July 1921, published in issue No. 2 of 27 August 1921).

1924d Zur Quantentheorie des Röntgenbremsspektrums, *Z. Phys.* **27**, 257–284 (received 24 July 1924, published in issue No. 4 of 12 September 1924).

1925b Zum Termproblem der Dublettspektren, insbesondere der Röntgenspektren, *Ann. d. Phys. (4)* **76**, 803–828 (received 23 February 1925, published in issue No. 8 of May 1925).

1925c Die Theorien des Compton-Effektes. I., *Phys. Zs.* **26**, 436–454 (received 22 May 1925, published in issue No. 12 of 15 June 1925).

1926a Die mehrfach periodischen Systeme in der Quantenmechanik, *Z. Phys.* **37**, 80–94 (received 27 March 1926, published in issue No. 1/2 of 22 May 1926).

1926b Eine Schwierigkeit für die Theorie des Kreiselelektrons, *Z. Phys.* **37**, 911–914 (received 22 May 1926, published in issue No. 12 of 10 July 1926).

1926c Eine Verallgemeinerung der Quantenbedingungen für die Zwecke der Wellenmechanik, *Z. Phys.* **38**, 518–529 (received 18 June 1926, published in issue No. 6/7 of 10 August 1926).

WEYL, HERMANN

1908 *Singuläre Integralgleichungen mit besonderer Berücksichtigung des Fourierschen Integraltheorems*, doctoral dissertation, University of Göttingen. Abridged and extended as: Singuläre Integralgleichungen, *Mathematische Annalen* **66**, 273–324 (published in issue No. 3 of 15 December 1908).

1909 Über beschränkte quadratische Formen, deren Differenz vollstetig ist, *Rend. Circ. Mat. Palermo* **27**, 373–392 (dated 11 November 1908, published in the issue of 19 January 1909).

1918c *Raum-Zeit-Materie: Vorlesungen über allgemeine Relativitätstheorie*, Berlin: J. Springer Verlag.

1928 *Gruppentheorie und Quantenmechanik*, Leipzig: S. Hirzel. English translation (by H. P. Robertson): *The Theory of Groups and Quantum Mechanics*, Methuen & Co., London, 1931; paperback reprint New York: Dover Publications, Inc.

WHITROW, GERALD JAMES

1972 *Einstein: The Man and His Achievement*, New York: Dover Publications, Inc. (reprint of the publication edited by the British Broadcasting Corporation, London, 1967).

WIENER, NORBERT

1921a The average of an analytical functional and the Brownian movement, *Proc. Nat. Acad. Sci. (U.S.A.)* **7**, 294–298 (communicated by A. G. Webster on 22 March 1921, published in issue No. 10 of 15 October 1921).

1921b On the theory of sets and points in terms of continuous transformations, in *Comptes Rendus du Congrès International des Mathématiciens* (Strasbourg, September 1920), Toulouse, pp. 312–315.

1924a Certain notions in potential theory, *J. Math. & Phys. M.I.T.* **3**, 24–51 (published in January 1924).

1924b Une condition nécessaire et suffisante de possibilité pour le problème de Dirichlet, *Comptes rendus (Paris)* **178**, 1050–1053 (presented by H. Lebesgue at the meeting of 24 March 1924).

1926a The operational calculus, *Mathematische Annalen* **95**, 557–584 (received 20 April 1925).

1926b The harmonic analysis of irregular motion, *J. Math. & Phys. M.I.T.* **5**, 99–121.

1926c The harmonic analysis of irregular motion. Second paper, *J. Math. & Phys. M.I.T.* **5**, 158–189.

1953 *Ex-prodigy: My Childhood and Youth*, New York: Simon & Schuster; reprinted Cambridge (Mass.): The M.I.T. Press, 1965; also paperback edition. (The quotations are taken from the latter.)

1956 *I Am a Mathematician: The Later Life of a Prodigy*, Garden City, N.Y.: Doubleday; paperback reprint Cambridge (Mass.): The M.I.T. Press. (The quotations are taken from the latter.)

WIENER, NORBERT, AND ARMAND SIEGEL

1953 A new form of the quantum postulate of quantum mechanics, *Phys. Rev. (2)* **91**, 1551–1560 (received 15 May 1953, published in issue No. 6 of 15 September 1953).

WREDE, ERWIN

1927 Über die magnetische Ablenkung von Wasserstoffatomstrahlen, *Z. Phys.* **41**, 569–575 (received 22 December 1926, published in issue No. 8/9 of 26 February 1927).

Author Index

Note: Pages where biographical data are given are indicated in italics.

Alexander, James Waddell 230
Allen, Herbert Stanley 62
Allen, L. vii
Appell, Paul-Émile 25
Artin, Emil 48
Autonne, Léon 122

Back, Ernst 279, 281, 282
Banach, Stefan 224
Barnett, Irving 223
Bell, Eric Temple 16
Bernoulli, Daniel 170
Bernoulli, Johann 13, 173
Bernstein, Felix 222
Birkhoff, George David 223
Blüh, Otto 193
Bôcher, Maxime 22–23, 59, 64, 65, 66, 94
Bodenstein, Max *47*
Bohr, Aage vii
Bohr, Harald 232, 246
Bohr, Niels 3–4, 11, 48, 50, 51, 55, 59, 63, 80, 91, 92, 95–96, 97, 102, 106, 118–120, 142, 143–144, 145, 149, 150, 164, 167, 168, 169, 171, 172, 173, 174, 178, 181, 182, 183, 184–185, 187, 191, 198, 200, 201–202, 203, 207, 208, 220, 232, 242–243, 244, 246, 250–251, 254, 256, 257, 258, 263, 268–269, 271–272, 276, 282, 283, 284, 285, 286, 287, 288
Bölsche, Wilhelm *45*
Boltzmann, Ludwig 46
Boole, George 17, 225
Born, Hedwig 5
Born, Max vii, 3, 4, 5–12, 34, 36–44, 48, 50–51, 53, 54–59, 61–148, 149, 150, 152, 153, 156, 157–158, 160–161, 163, 165–166, 167, 169, 170–172, 173–174, 175, 177–178, 187, 188, 189, 190, 195, 201, 204, 209, 211–214, 217, 220–221, 230–251, 254, 257, 258, 261, 265, 268, 277, 283, 285, 287, 288, 289–293, 294, 295–299, 300
Bose, Satyendra Nath 149–150, 207, 299
Bothe, Walther 51, 150
Bouligand, Georges 224
Bowen, Ira Sprague 288
Brioschi, Francesco 17
Broglie, Louis de 6, 150, 244–245

Carnap, Rudolf 46
Carson, John Renshaw *227*, 228
Carst, Agathe 284
Cauchy, Augustin Louis *14*, 16, 22, 25, 261
Cayley, Arthur 3, 11, 15, *16–17*, 18–20, 21, 22, 32, 34–35, 65, 66–67, 68, 119
Cayley, Henry 16
Cayley, Maria Antonia, *née* Doughty 16
Cayley, Susan, *née* Moline 17
Christina, Queen of Sweden 13
Chrystal, George 34
Clebsch, Alfred 17
Cohn, Emil 205
Compton, Arthur Holly 273, 280
Courant, Richard 10, 23, 38, 47, 50, 53, 59, 64, 65, 67, 68, 94, 131, 206–207, 230, 246, 286
Cramer, Gabriel *24*

Darwin, Charles 45
Debye, Peter 152, 181, 191, 193, 221
Descartes, René *13*, 33
Dewey, John 223
Dieudonné, Jean 22
Dirac, Paul Adrien Maurice vii, 78, 109, 110, 111, 118, 134, 143, 149, 167, 219, 229, 245, 253, 258, 299, 300

Eddington, Arthur Stanley 49

Ehrenfest, Paul 5, 50, 52, 149, 150–152, 154
Einstein, Albert 3, 5, 6, 7, 8, 50, 51–52, 54, 55, 57, 90, 149, 150, 155, 156, 167, 202, 205, 206, 207, 208, 219, 222, 244, 301
Eisele, Carolyn 25
Elsasser, Walter 6, 150, 244
Eötvös, Loránd 205, 206
Epstein, Paul Sophus 181
Euler, Leonhard *13*, 16, 30

Fejér, Lipot (Leopold) 205
Fermi, Enrico 299
Föppl, August 37
Föppl, Ludwig *37*–38
Forman, Paul vii
Fourier, Joseph Baptiste *23*–24, 26, 27
Franck, James 7, 49, 50, 53, 54, 55, 150, 202, 285
Frank, Philipp 5, 46
Franklin, Philip 224
Fréchet, Maurice 223
Fredholm, Erik Ivar *29*–30, 34, 119
Frenkel, Jakow 270
Frobenius, Ferdinand Georg *21*, 119

Galois, Évariste 22
Gauss, Carl Friedrich 14, 15
Geiger, Hans 150
Gibbs, Josiah Willard 223
Gödel, Kurt 46
Goldstein, Herbert 173
Gordan, Paul 17, 33
Götze, Raimund 289
Goudsmit, Samuel 4, 165, 179, 195, 199–203, 247, 266–269, 270, 271–273, 280, 284–285, 287
Graeffe, Karl Heirich *170*
Grossmann, Marcel 3

Haeckel, Ernst 45
Hahn, Hans 46
Hamilton, William Rowan 15, 17, 20, 21, 34, 43, 64, 68, 139
Hanle, Wilhelm 8, 11
Hardy, Godfrey Harold 222, 247
Hartmann, Johannes Franz *49*
Heaviside, Oliver 225, *226*–229, 231, 235, 245
Heaviside, Rachel Elizabeth, *née* West 226
Heaviside, Thomas 226
Heilbron, John L. vii
Heisenberg, Werner vii, 3, 4, 5, 7–12, 43–44, 48, 52, 54, 55, 57, 59, 61, 62, 63, 64, 66, 68, 72, 74, 78–80, 81, 83–84, 85, 87–88, 90, 91–121, 123, 124, 128–129, 131–134, 136–149, 150, 156, 157–166, 167, 168–172, 174, 175, 177–178, 180, 181, 182, 187, 188, 189, 195, 196–204, 209, 217, 219, 221, 231, 232, 235, 244, 245, 247–255, 257–258, 261, 265, 266, 267–270, 272–283, 285, 286–301
Heitler, Walter 300–301
Hellinger, Ernst *32*–34, 38, 120–121, 124, 125, 130–135
Helmholtz, Hermann von 48–49
Hermann, Jakob 173
Hermite, Charles 3, 17, *73*, 120, 123
Hertz, Gustav 54
Hilbert, David 3, 5, 10, 23, 30–34, 35, 43, 44, 49–50, 59, 64, 65, 66, 67, 68, 73, 76, 77, 92, 94, 103, 110, 119–125, 129, 130–132, 206–207, 209, 221, 222, 230, 232
Hill, George William *25*–27, 34, 37–38, 39, 77
Himstedt, Franz 205
Holmgren, E. 30
Hönl, Helmut 165, 179
Hopf, Eberhard 246
Hopmann, Josef 206
Horace, Quintus Flaccus 15
Hund, Friedrich 8, 11, 48, 51, 54, 57, 287
Husserl, Edmund 222

Jackson, Dugald C. 225
Jacobi, Carl Gustav Jacob 43, 139
Jaffé, George 286
Johnson, Lyndon Baines 247
Jordan, Camille 17, *22*
Jordan, Ernst Pascual 44
Jordan, Eva, *née* Fischer 44
Jordan, Pascual vii, 3, 4, 5, 6, 7, 8, 9, 11, 12, *44*–59, 61–119, 121, 123–124, 128–129, 131–134, 136–150, 152–156, 157–166, 167, 168, 169, 170–172, 174, 175, 177–178, 187, 188, 189, 201, 204, 209, 211–214, 217, 220, 221, 231, 232, 235,

258, 261, 265, 266, 267, 272–283, 285, 286, 288, 290–291, 292, 294, 301

Kant, Immanuel 222
Kapitza, Peter 5
Kármán, Theodor (Theodore) von 38–43, 111, 120, 127
Kelvin, William Thomson, Lord 34
Kemble, Edwin Crawford 62
Kepler, Johannes 13
Klein, Felix 3, 119, 222
Klein, Oskar 181
Knopp, Konrad 46
Knott, Cargill Gilston 20, 34
Koch, Helge von *27–29*
Kohlrausch, Wilhelm Friedrich *47*
Koopman, Bernard Osgood 223
Kopfermann, Hans 284
Kötteritzsch, Th. 24–25, 27
Kramers, Hendrik Anthony 5, 6, 12, 51, 55, 93, 94, 103, 139–143, 148, 150, 164, 181, 191, 207, 208, 217, 253, 271–272, 285, 289
Kratzer, Adolf 163
Kronig, Ralph de Laer 11, 76, 165, 167–168, 174, 179, 181, 193, 196–197, 200, 201, 202
Kühn, Alfred *49*
Kuhn, Thomas S. vii
Kuhn, Werner 142

Ladenburg, Rudolf 284
Lagrange, Joseph Louis, Comte de 34, 225
Lamarck, Jean Baptiste 45
Lanczos, Cornelius (Kornel) vii, 4, 195, *204–220*, 251, 300
Lanczos, Elmar 220
Lanczos, Ilse, *née* Hildebrand 220
Lanczos, Maria, *née* Rupp 220
Landau, Edmund 50, 222
Landau, Lev Davidovich 193
Landé, Alfred 12, 54, 199, 201, 279, 281, 282, 284
Lange, Friedrich Albert *45*
Langevin, Paul 301
Laplace, Pierre Simon de 173
Laue, Max von 36
Lebesgue, Henri 224
Leibniz, Gottfried Wilhelm *14*, 16, 17, 18, 20, 23, 225
Lenz, Wilhelm 172–173, 181, 183, 184, 188, 191
Levi–Civita, Tullio 3
Lichtenstein, Leo 224
Loewy, Alfred 205
Loewy, Carolus 204
London, Fritz 100, 300–301
Lorentz, Hendrik Antoon 202
Lyman, Theodore 289

Mach, Ernst 46, 57
Madelung, Erwin 23, 206–207
Manneback, Charles 193
Maxwell, James Clerk 34, 87, 88, 89
Mehra, Jagdish 55, 167, 218
Meissner, Karl Wilhelm 53–54
Menger, Karl 46
Mensing, Lucie (Lucy) 12, 187–193, 281–282
Mersenne, Marin 13
Mie, Gustav 36–37, 205
Millikan, Robert Andrews 288
Minkowski, Hermann 3, 10, 35–37, 38, 119
Mittag-Leffler, Gösta 27, 29
Moulton, Forest Ray 25
Müller, Conrad Heinrich *47*
Murphy, Robert 225

Napoleon (Bonaparte) 23, 44
Nernst, Walther 46
Neumann, John (Johannes) von 122–123, 131, 223, 245
Neurath, Otto 46
Newcomb, Simon 25
Newton, Isaac 3, 6, 14, 23, 64, 68
Noether, Emmy 50, 93
Nordheim, Lothar 122
Nörlund, Niels Erik 6, 50, 56
North, J. D. 17
Norton, C. L. 221

Ornstein, Leonard Salomon 150
Ortvay, Rudolph 205
Osgood, William Fogg 223

Paley, R. E. C. A. 246
Paschen, Friedrich 289

Pauli, Wolfgang vii, 3, 4, 7, 8, 11–12, 44, 47, 48, 50, 51–52, 54, 58, 59, 62, 76, 94, 96, 104, 105, 109, 115, 117, 119, 131, 137, 139–140, 143, 147, 149, 160, 164, 165, 166–187, 189, 191–193, 195, 196–198, 200–204, 217, 219, 220, 244, 245, 247–258, 261, 267–274, 280, 281, 282–289, 291, 299–300, 301
Pauly, A. 45
Peacock, George *17*
Peter the Great, Czar of Russia 14
Pincherle, Salvatore 226
Planck, Max 46, 49, 52, 90, 152
Plato 32
Pohl, Robert Wichard 47–48, 49
Poincaré, Jules Henri 3, 6, 22, 25, 27–28, 29, 34, 48
Prandtl, Ludwig 6, 8
Prange, Georg *47*
Precht, Carl Christian Julius *47*

Ramsauer, Carl 150
Raschevsky, N. von 206
Rayleigh, John William Strutt, Lord 230
Reid, Constance 30, 38, 230
Ricci, Curbastro Gregorio 3
Richard III, King of England 269
Riemann, Bernhard 17, 25
Rosanes, Jakob 9, *38*, 44, 119
Runge, Carl 173
Russell, Bertrand 222, 223
Rydberg, Janne Robert 288

Salmon, George 17
Schlick, Moritz *46*
Schmidt, Erhard *33*, 170
Schoenfliess, Arthur 46
Schönborn, Johann Philipp, Elector of Mainz 14
Schrödinger, Erwin vii, 4, 61, 138, 167, 218, 219, 220, 243–244, 245, 283, 284, 300
Schur, Issai 122
Schuster, Arthur 230
Schütz, Wilhelm 188
Schwarzschild, Karl 48, 181
Servois, François Joseph 225
Siegel, Armand 246
Slater, John Clarke 51, 150, 207, 208
Smekal, Adolf 150
Sommerfeld, Arnold 36, 48, 164, 169, 172, 173, 178, 181, 182, 185, 191, 200, 205, 207, 253, 254, 257, 259–266, 267, 276–278, 280, 283, 287, 292
Sponer, Hertha 8, 51, 57
Stern, Otto 167, 198
Stoner, Edmund Clifton 54
Struik, Dirk Jan 246
Sylvester, James Joseph *15*, 16, 17, 20, 21
Szasz, Otto 222
Szilard, Leo 219

Tait, Peter Guthrie 20, 34
Tamm, Igor 193
Thilly, Frank 222
Thirring, Hans 62
Thomas, Llewellyn Hilleth *270*–273, 282, 285, 286
Toeplitz, Otto *32*–34, 38, 39–40

Uhlenbeck, George Eugene 4, 195, 199–203, 247, 266–269, 273, 280, 296
Unsöld, Albrecht 283

Van der Waerden, Bartel Leendert 8, 63, 90, 91, 100, 101, 102, 108, 149, 156, 157, 170
Van Vleck, John Hasbrouck 55, 193
Voigt, Woldemar 203, 276

Waismann, Friedrich 46
Weierstrass, Karl Theodor Wilhelm *21*, 25
Weiss, Pierre 301
Wentzel, Gregor 4, 253–265, 281, 282, 286, 300
Weyl, Hermann 10, 63, 100, 109, 110, 120, 123, 129–130, 161, 217
Wheatstone, Charles 226
Whitrow, Gerald James 205, 219
Wiener, Bertha, *née* Kahn 221
Wiener, Leo 221, 246
Wiener, Norbert vii, 4, 195, 204, *221–225*, 227–*246*, *247*–251, 254, 257–258, 268, 283, 298, 300
Wrede, Erwin 198

Zernike, Frits 150