Gennaro Auletta

The Quantum Mechanics Conundrum

Interpretation and Foundations

Springer

The Quantum Mechanics Conundrum

Gennaro Auletta

The Quantum Mechanics Conundrum

Interpretation and Foundations

Springer

Gennaro Auletta
University of Cassino and Southern Lazio
Cassino, Frosinone, Italy

Pontifical Gregorian University
Rome, Italy

ISBN 978-3-030-16648-9 ISBN 978-3-030-16649-6 (eBook)
https://doi.org/10.1007/978-3-030-16649-6

© Springer Nature Switzerland AG 2019
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

The starting point of this book was my tutorship for a doctoral dissertation in the year 2014 at the Pontifical Gregorian University in Rome. The now Ph.D. Father Lluc Torcal wrote a thesis titled *In Search of an Ontology that Underlies Quantum Mechanics: Which Vision of the Physical World Could We Get from Quantum Mechanics?* It was essentially a philosophical assessment of the main problems and solutions in interpreting quantum theory with a look at the proposals that I have developed in my own research across more than 20 years starting from the publication of *Foundations and Interpretation of Quantum Mechanics: A Critical–Historical Analysis of the Problems and a Synthesis of the Results*, published in 2000 by World Scientific. At that time, I spent a summer in the monastery of Poblet working together with Father Torcal. Since this gave us the opportunity for examining many crucial points in interpreting quantum mechanics, after the discussion of the mentioned dissertation, both of us envisaged the project of making a book on this subject. However, when I started to work on it, I became soon aware that the issue at the stake went much further than a simple philosophical assessment. Moreover, Father Torcal, due to his duties in the Cistercian order, could not follow this project. Nevertheless, the current structure of the first three chapters partly mirrors that of the dissertation.

And that was why metaphysics and science were
such courageous enterprises, such startling inventions,
bigger than the wheel, bigger than agriculture,
human artifacts set right against the grain of human nature.
Disinterested truth.
Ian McEwan, *Enduring Love*

Rome, Italy — Gennaro Auletta

Introduction

As well-known, it was the epistemologist K. Popper to speak of a schism in physics determined by quantum theory.[1] Retrospectively it seems opportune to ask: does the schism consist of the introduction of the subjectivity in the mid of physics, as Popper (following Einstein) seems inclined to think, or in accepting that events occur randomly without any cause or even ground as stressed by Bohr? Or is it rooted in the uncontrolled interactions of the experimenter with quantum systems as suggested by the young Heisenberg? Or even does it consist in the acceptance of non-local correlations as pointed out by Schrödinger? The answer may vary according to personal taste. These aspects essentially deal with the following problems (keeping the same order): reality, causality, measurement, status of formalism. They represent a challenge to our rational understanding of Nature. They are all fundamental issues, but I am inclined to think that the crucial problem that is at the root of all these puzzles is the notion of reality. So, we may rephrase Popper's definition of the schism in physics as the possible break down of the notion of reality as such.

It is thus not by chance that quantum mechanics still today (after a century from its beginning) is poorly understood and not universally accepted by the community of scientists and philosophers. As a matter of fact, many physicists still today consider it as a kind of metaphysical speculation devoid or poor of empirical import, and when they use it in neighbouring fields (e.g. cosmology or statistical mechanics), they limit such a connection to what is strictly necessary and often, in these applications, forget or minimise what is really typical and fundamental of quantum theory. The present book is addressed to the problem of making quantum mechanics understandable and viable in its rational foundations to the community of scholars.

Now, the issue of reality has in turn its roots in a categorial problem. In fact, if our relation with reality is disturbed, it must be related somehow with the way in which we categorise reality with our notions of physical object, law of nature, probability, cause and so on. Moreover, if, in order to circumvent the problem, we

[1] Popper (1982).

deny that there is any reality as such, then we are saying that our categorial structure be simply the way in which we humans (and likely in different ways) build our "story". These are two sides of the same coin: to deny reality and to reduce categories to mere subjective constructions or games. Thus, although such a critical philosophical problem arises in the mid of science for the first time with quantum mechanics, it, directly or indirectly, affects the whole of our knowledge.

The part of the work that is devoted to interpretation is dominated by four giants: Bohr, Einstein, Heisenberg and Schrödinger. On many issues they did not agree. They had strong different scientific backgrounds and even philosophical ideas: Bohr was very much influenced by Kant and Kirkegaard and interested in epistemic problems of the theory, Einstein by Spinoza and, for what his realism concerns, by a kind of Aristotelian background, the late Heisenberg was also influenced by Aristotle but with a constant operational approach, and Schrödinger was fascinated by Schopenhauer and the Oriental philosophy. What they have in common is that they showed a significant interest in general ideas, a character that seems to be lost in most of the actual scientific debate, apart from very rare exceptions. In the following, we shall see that each of these great scientists was somehow right and that their views can be integrated. In fact, meanwhile, we have cumulated a so huge amount of experimental results (and of theoretical developments) to be able to solve most of the questions that were object of discussion at that time. However, for doing this, a philosophical critical assessment is necessary.

It could be objected that to satisfy everybody seems too ambitious, especially with a theory that has raised so many different standpoints and even dissatisfaction. Nevertheless, the essence of the scientific enterprise is to find common rational explanations of our world.[2] It is a permanent duty of scholars involved in these problems, and I hope that the following examination could show at least elements that could help for building such a common view according to quantum mechanics.

The other issue is rather a foundational inquiry. We shall show that quantum mechanics is basically an information theory. However, such a reversible theory of information needs to be integrated with a theory of signals when interactions are involved. Here, the relationship with both special and general relativity is crucial. Finally, we shall deal with Category theory and display the categorial structure of quantum mechanics. The method of this work is not axiomatic and in general I prefer to proceed pragmatically instead of from definitions and axioms. This method, as we shall see, was anticipated by Einstein when he said that physics is a science "going in the direction of increasing simplicity of the logical bases".[3] I interpret this statement as meaning that, instead of trying to lead consequences from first truths, the right method is to start from theories and explanations that are less general and less grounded and to proceed backwards to better foundations and theories that are more general. As we shall see, this is the method followed today by Category theory and this explains its central role in the present work.

[2] As pointed out in Deutsch (1997, Chap. 1). Also C. Rovelli has insisted on the relevance of foundational issues (Rovelli 2004, p. 20).

[3] Einstein (1936, p. 96). See also Auletta (2011, Sects. 5.2, 2.1.2 and 5.4.2).

Such a subject necessarily involves many aspects and disciplines and certainly a massive use of philosophical methods and notions. In so doing, it is mandatory to fulfil the five criteria of philosophical adequacy set by Abner Shimony: coherence, fineness of reasoning, comprehensiveness, openness to evidence, richness of contents, of which the latter three require scrupulous taking into account of scientific results.[4] As Sir A. Eddington said, "the compartments into which human thought is divided are not so water–tight that fundamental progresses in one is a matter of indifference to the rest".[5] I completely support this point of view and have taken inspiration from it for the present book. In this context, he recalled that physics was traditionally called natural philosophy and that a specialisation that has made philosophy and science dumb to each other is not sane.

References

Auletta, Gennaro (in collaboration with I. Colagè, P. D'ambrosio, and L. Torcal). 2011. *Integrated Cognitive Strategies in a Changing World*. Rome: G and B Press.

Deutsch, David. 1997. *The Fabric of Reality: The Science of Parallel Universes and Its Implications*. London: Penguin Books.

Eddington, Arthur Stanley. 1939. *Philosophy of Physical Science*. Cambridge: University Press.

Einstein, Albert. 1936. Physik und Realität. *Journal of Franklin Institute* 221: 349–382; Eng. tr. in Einstein (1956), 59–97.

Einstein, Albert. 1956. *Out of My Later Years, Estate of A. Einstein*. New York: Wings Books.

Popper, Karl R. 1982. *Quantum Theory and the Schism in Physics*. London: Unwin Hyman Ltd; Routledge, 1992, 1995.

Rovelli, Carlo. 2004. *Quantum Gravity*. Cambridge: University Press.

Shimony, Abner. 1970. Scientific Inference. In R. G. Colodny (Ed.). *The Nature and Function of Scientific Theories*, 79–172, Pittsburgh; rep. in Shimony (1993), I, 183–273.

Shimony, Abner. 1993. *Search for a Naturalistic Point of View*. Cambridge: University Press.

[4]Shimony (1970, p. 187).

[5]Eddington (1939, p. 8).

Contents

1 Summary of the Basic Elements of the Theory 1
1.1 The Beginning of Quantum Mechanics 1
1.1.1 From the Classical to the Quantum World 1
1.1.2 The Quantum Postulate and Matter's Discontinuity 5
1.1.3 The Road to the Schrödinger Equation 9
1.2 The Basic Principles of Quantum Mechanics 24
1.2.1 The Superposition Principle 24
1.2.2 The Quantisation Principle 30
1.2.3 Physical Observables and Different Representations 44
1.2.4 Commutativity and Uncertainty Relations 61
1.2.5 Unitary Transformations and Symmetries 88
1.3 Some Further Principles 95
1.3.1 Pauli's Exclusion Principle and Spin 95
1.3.2 The Correspondence Principle 105
1.3.3 The Complementarity Principle 108
1.4 Density Operator and Compound Systems 110
1.4.1 Pure and Mixed States 110
1.4.2 Entangled and Product States 116
1.4.3 Total and Marginal States 120
1.5 Summary 124
References 124

2 The Main Problems 129
2.1 Quantum-Mechanical Formalism 129
2.1.1 Two Pictures of Quantum Dynamics 129
2.1.2 Probability and Interference 139
2.1.3 What Does the Quantum-Mechanical Formalism Tell Us? 150
2.2 The Measurement Problem 150

2.2.1 It is Not a Problem for Classical Physics . . . 150
2.2.2 It is a Double Problem for Quantum Mechanics . . . 152
2.2.3 von Neumann's Problem . . . 156
2.2.4 What Are the Consequences? . . . 161
2.2.5 Schrödinger Cat . . . 162
2.3 The Problem of Non-locality . . . 164
2.3.1 The Paradox Proposed by Einstein–Podolsky–Rosen . . . 164
2.3.2 Classical Physics and Locality . . . 171
2.3.3 Reality and Correlations . . . 187
2.3.4 Possible Replies to EPR . . . 191
2.4 The Problem of Causality . . . 191
2.4.1 Classically . . . 191
2.4.2 Quantum-Mechanically . . . 195
2.4.3 Delayed Choice . . . 201
2.5 Summary of the Chapter . . . 206
References . . . 207

3 The Main Interpretations . . . 215
3.1 Solutions to the Measurement Problem . . . 216
3.1.1 Subjectivist Interpretation . . . 216
3.1.2 Objectivist Interpretation . . . 226
3.1.3 Role of the Environment . . . 235
3.1.4 Solution of the Cat's Paradox . . . 244
3.1.5 Summarising . . . 247
3.2 Formalism and Reality: Information . . . 247
3.2.1 Objectivism . . . 247
3.2.2 A Formalism That Denotes a Reality? . . . 248
3.2.3 Dealing with Information . . . 261
3.2.4 Two Points of View on Entropy and Information . . . 271
3.3 Hidden Variables and Non-locality . . . 275
3.3.1 Logical Restatement of the EPR Argument . . . 275
3.3.2 Bohr's Reply to EPR . . . 277
3.3.3 Schrödinger's Reply to EPR . . . 282
3.3.4 Hidden Variables . . . 285
3.3.5 Bell Inequalities for Testing EPR . . . 300
3.3.6 Entanglement Swapping, Teleportation, Quantum Cryptography . . . 308
3.3.7 What Is Non-separability . . . 320
3.4 Back to the Problem of Causality . . . 333
3.4.1 Form, Structure and Information . . . 333
3.4.2 Correlations and Causal Constraints . . . 339
3.4.3 Potentiality . . . 346

3.5 Summary . . . 352
References . . . 353

4 Quantum Computation . . . 365
4.1 A Reversible Multipath Computation . . . 365
4.2 Quantum Gates . . . 372
4.2.1 The Basic Gates . . . 372
4.2.2 Generation of States . . . 376
4.2.3 Relations Among Gates . . . 379
4.3 Quantum Algorithms for Computing . . . 384
4.3.1 Deutsch's Algorithm . . . 384
4.3.2 Shor's Theorem . . . 387
4.3.3 Grover's Algorithm . . . 389
4.4 Heisenberg Picture and Teleportations . . . 392
4.4.1 Heisenberg Picture for Computation . . . 392
4.4.2 Generalised Teleportation . . . 395
4.4.3 Logic Gate Teleportation . . . 397
4.5 Energy and Time in Quantum Computation . . . 401
4.6 Summary . . . 402
References . . . 403

5 Ontological Ascription and Operations . . . 405
5.1 Quantum Events and Features . . . 405
5.1.1 Quantum Events . . . 405
5.1.2 Quantum Features . . . 413
5.1.3 Quantum Features and Discord . . . 429
5.1.4 Again on Complementarity . . . 435
5.1.5 A Participatory Universe: Relativity and Quantum Mechanics . . . 440
5.2 Three Subsequent Steps in the Measurement Process . . . 445
5.2.1 General Considerations . . . 445
5.2.2 Preparation . . . 446
5.2.3 Premeasurement and Reversibility . . . 450
5.2.4 Detection . . . 460
5.2.5 Generalised Operations . . . 481
5.2.6 Some Lessons . . . 486
5.3 Four Convergent Interpretations . . . 489
5.4 Summary . . . 491
References . . . 492

6 Information and Signal Exchanging in Universal Network . . . 497
6.1 A Quantum Theory of Information . . . 498
6.1.1 Information is Conserved in Our Universe . . . 498
6.1.2 No-programming Theorem . . . 504

6.1.3 No-Bit Commitment . . . 507
6.1.4 Information Accessibility . . . 510
6.1.5 Final Considerations . . . 512
6.2 Signal Exchanging . . . 516
6.2.1 Channels for Exchanging Information . . . 517
6.2.2 No Information Acquisition Without Disturbance . . . 520
6.2.3 Informational Completeness and Tomography . . . 526
6.2.4 Green's Functions, Path Integrals, Quantum Trajectories . . . 537
6.2.5 Information Sources, Causality, Reversibility . . . 547
6.3 Networks . . . 558
6.3.1 D-Shaped Circuits . . . 558
6.3.2 Error Correction . . . 560
6.3.3 Teleportation and Entanglement Swapping as Example of POVM Networks . . . 567
6.3.4 Information Selection, Information Sharing, Information Acquiring . . . 574
6.4 Quantum Mechanics Out of Six Informational Principles . . . 581
6.4.1 Principle of Causality . . . 582
6.4.2 Atomicity of the Composition . . . 586
6.4.3 Perfect Discriminability . . . 587
6.4.4 Local Discriminability . . . 587
6.4.5 Ideal Compression . . . 588
6.4.6 Purification . . . 590
6.5 Summary . . . 593
References . . . 594

7 Quantum Mechanics, Fields and Geometry . . . 599
7.1 Fields, Matters, Geometry . . . 600
7.1.1 Classical Electromagnetic Field . . . 600
7.1.2 Second Quantisation . . . 606
7.1.3 Quantum-Field Theory . . . 612
7.1.4 Matters . . . 627
7.1.5 A Basic Exploration of General Relativity . . . 634
7.2 Space-time, Gravity, Signals . . . 650
7.2.1 Quantum and Space-time . . . 650
7.2.2 Loop Quantum Gravity and Entropic Gravity . . . 659
7.2.3 A Dynamical View . . . 670
7.3 Emerging World . . . 674
7.3.1 The Hydrogen Atom . . . 675
7.3.2 Approximation Methods . . . 690
7.3.3 The Emergence of the Molecular World . . . 695

7.4 Summary of the Chapter . . . 711
References . . . 712

8 Category Theory and Quantum Mechanics . . . 717
8.1 Category Theory . . . 717
8.1.1 Introduction . . . 717
8.1.2 Preliminary Notions: Sets and Maps . . . 719
8.1.3 Categories and Functors . . . 727
8.1.4 Universal Constructions . . . 739
8.1.5 Natural Transformations . . . 745
8.1.6 Category Theory and Quantum Information . . . 751
8.2 Logic Assessment: Boolean Algebra and Quantum Mechanics . . . 757
8.2.1 Basic Notions . . . 757
8.2.2 Combinatorial Approach . . . 762
8.2.3 Building Logical Spaces . . . 771
8.2.4 Is Quantum Logic Non-classical? . . . 782
8.3 Epistemological–Gnoseological Assessment . . . 786
8.3.1 Physical (and Operational) Connection . . . 787
8.3.2 Inferential Connection . . . 799
8.3.3 Interpreted and Uninterpreted Ontology . . . 819
8.4 Summary . . . 828
References . . . 829

Bibliography . . . 837

Author Index . . . 839

Subject Index . . . 849

Abbreviations

BM	Beam merger
BS	Beam splitter
CHSH	Clauser, Horne, Shimony, Holt
CM	Classical mechanics
D	Detector
EPR	Einstein, Podolsky, Rosen
GHSZ	Greenberger, Horne, Shimony, Zeilinger
GR	General relativity
HV	Hidden variable
iff	if and only if
LASER	Light amplification by stimulated emission of radiation
LCAO	Linear combination of atomic orbitals
LHS	Left-hand side
LIF	Local inertial frames
LQC	Loop quantum gravity
PS	Phase shifter
QFT	Quantum field theory
QM	Quantum mechanics
RHS	Right-hand side
SR	Special relativity
VBM	Valence bond method

List of Figures

Fig. 1.1 Phase, amplitude, and wavelength. 3
Fig. 1.2 Snapshot of an electromagnetic wave 4
Fig. 1.3 The duality wave and particle 10
Fig. 1.4 Constructive and destructive interference 11
Fig. 1.5 Oscillations of the electron 12
Fig. 1.6 Real and imaginary parts 14
Fig. 1.7 Plot of the exponential function 17
Fig. 1.8 Partial derivative of a function of two variables 20
Fig. 1.9 Quantum state as superposition of vectors 25
Fig. 1.10 Projection 35
Fig. 1.11 Circular polarisation 40
Fig. 1.12 Change of basis in Hilbert space. 40
Fig. 1.13 Gradient of a function. 51
Fig. 1.14 De Broglie's waves as a 1–form 53
Fig. 1.15 Sums of 1–forms 54
Fig. 1.16 Gradient of a scalar field as a 1–form and 2–form 55
Fig. 1.17 Action of several polarisation filters 64
Fig. 1.18 Angular momentum of a classical particle 66
Fig. 1.19 The right-hand rule 67
Fig. 1.20 Levi–Civita Tensor 69
Fig. 1.21 Graphical representation of a matrix 70
Fig. 1.22 Graphical representation of a tensor 71
Fig. 1.23 Quantisation of angular momentum 75
Fig. 1.24 Relation between rectangular and spherical coordinates 76
Fig. 1.25 Uncertainty relation. 85
Fig. 1.26 Stern–Gerlach Experiment 96
Fig. 1.27 Planck's spectrum 107
Fig. 1.28 Representation on a sphere of pure and mixed states 113
Fig. 2.1 Active and passive transformations 134
Fig. 2.2 Mach–Zehnder interferometer 142
Fig. 2.3 Mach–Zehnder interferometer with object. 145

Fig. 2.4 Repeated interaction–free measurements . . . 146
Fig. 2.5 Mach–Zehnder interferometer with a blocked path . . . 148
Fig. 2.6 Unitary transformation in Hilbert space . . . 153
Fig. 2.7 Time evolution in phase space and the uncertainty relation . . . 154
Fig. 2.8 Schrödinger cat thought experiment . . . 162
Fig. 2.9 EPRB Experiment . . . 168
Fig. 2.10 Relation between two different reference frames $\mathcal{R}$ and $\mathcal{R}'$ under Galilei transformations . . . 172
Fig. 2.11 The Michelson–Morley Interferometer . . . 174
Fig. 2.12 Fibre bundles . . . 179
Fig. 2.13 Light cones . . . 180
Fig. 2.14 Boscovich's alternation of forces . . . 192
Fig. 2.15 Delayed choice experiment . . . 202
Fig. 2.16 The realised delayed–choice experiment . . . 203
Fig. 2.17 Results of delayed–choice experiment . . . 204
Fig. 3.1 Decohering wave packets . . . 238
Fig. 3.2 Decohering histories . . . 239
Fig. 3.3 Pictorial representation of a coherent state and separation between the two components . . . 245
Fig. 3.4 Plot of logarithms . . . 251
Fig. 3.5 Entropic uncertainty relations . . . 256
Fig. 3.6 Poincaré sphere representation of quantum states . . . 263
Fig. 3.7 Unit interval [0, 1] as limiting case of digitized coding . . . 269
Fig. 3.8 Mandel's experiment on Empty Waves . . . 287
Fig. 3.9 Magnetic and Electric AB-effect . . . 291
Fig. 3.10 10-node Kochen–Specker graph . . . 296
Fig. 3.11 117-node Kochen–Specker graph . . . 297
Fig. 3.12 Optimal orientation for $\mathbf{a}, \mathbf{a}', \mathbf{b}, \mathbf{b}'$ for testing CHSH inequality . . . 305
Fig. 3.13 Aspect and coworkers' experiment . . . 306
Fig. 3.14 Alley–Shih and Ou–Mandel's experiment . . . 307
Fig. 3.15 Schematic set-up of entanglement swapping . . . 309
Fig. 3.16 Scheme of entanglement swapping . . . 311
Fig. 3.17 Scheme of quantum teleportation . . . 312
Fig. 3.18 Schematic setup of quantum teleportation . . . 315
Fig. 3.19 Experimental realisation of quantum teleportation . . . 316
Fig. 3.20 PR Boxes . . . 325
Fig. 3.21 Information causality . . . 329
Fig. 3.22 The GHSZ Proposed Experiment . . . 335
Fig. 3.23 Conditional entanglement . . . 337
Fig. 3.24 Borromean rings and Hopf rings . . . 338
Fig. 4.1 Scheme of a Turing machine . . . 366
Fig. 4.2 Logical circuits . . . 367
Fig. 4.3 Representation of computational complexity . . . 368

Fig. 4.4 Poincaré–sphere representation of Hadamard gate 373
Fig. 4.5 Phase transformations on the x–y plane of the Poincaré sphere 380
Fig. 4.6 Representation of Shor Theorem 388
Fig. 4.7 Geometric representation of the Grover's algorithm 392
Fig. 5.1 Wheeler's representation of reality 410
Fig. 5.2 Bénard cells 414
Fig. 5.3 Classical correlation, superposition, and entanglement 415
Fig. 5.4 Interference in Mach-Zehnder interferometer 419
Fig. 5.5 Experiment proposed for detecting features each experimental run 424
Fig. 5.6 Parallel transport 428
Fig. 5.7 Geometric phase 429
Fig. 5.8 Mutual information 430
Fig. 5.9 Entanglement between a system and a fragment of the universe 433
Fig. 5.10 Relation Holevo quantity/discord 435
Fig. 5.11 Mach–Zehnder interferometer with variable beam splitters 436
Fig. 5.12 Wheeler's participatory universe 444
Fig. 5.13 Schematic representation of Mandel's experiment 452
Fig. 5.14 Interference and Visibility in Mandel's experiment 453
Fig. 5.15 Scully, Englert and Walther's proposed experiment 455
Fig. 5.16 Bloch–sphere representation of states 471
Fig. 5.17 Amplitude damping contraction of the Bloch sphere 473
Fig. 5.18 Example of POVMs 474
Fig. 6.1 Negative conditional entropies 514
Fig. 6.2 Schematic representation of quantum non–separability 516
Fig. 6.3 Potential and energy levels of the harmonic oscillator 522
Fig. 6.4 Displacement operator for coherent states 531
Fig. 6.5 Representation of the Wigner function of a coherent, number and squeezed state 533
Fig. 6.6 The different paths 539
Fig. 6.7 The sum over paths 540
Fig. 6.8 Two possible paths from i to f both passing through the same central point c 542
Fig. 6.9 What "has happened" in the past is influenced by interactions made in the present 547
Fig. 6.10 Fundamental information triad 579
Fig. 7.1 Magnetic flux through an open surface $\mathbf{S}$ 602
Fig. 7.2 The three directions of the electromagnetic field 609
Fig. 7.3 Standard model 627
Fig. 7.4 Colours and anticolours 628
Fig. 7.5 Baryon and meson octets 629
Fig. 7.6 Colour exchange 629

Fig. 7.7 Quark–Antiquark fields 630
Fig. 7.8 Tidal and contraction effects of gravity 638
Fig. 7.9 Geodesic deviation 639
Fig. 7.10 Intrinsicity of geodesics 640
Fig. 7.11 Parallel transport in curved space 642
Fig. 7.12 Relation between basis vectors and dual 1–forms 643
Fig. 7.13 Family of geodesics 644
Fig. 7.14 Length and rotation in geodesic deviation 645
Fig. 7.15 Interferometer for detecting gravitational waves 647
Fig. 7.16 A graphical representation of the clock–shutter device 648
Fig. 7.17 Ma et al. experiment 652
Fig. 7.18 Ma et al. experiment in details 653
Fig. 7.19 Loops and electric field 659
Fig. 7.20 Spin networks 660
Fig. 7.21 Spin foam 661
Fig. 7.22 Black hole's entropy 668
Fig. 7.23 Atomic levels 676
Fig. 7.24 Atomic suborbitals 677
Fig. 7.25 Electron in a hydrogen atom 679
Fig. 7.26 Plot of the radial eigenfunctions of the hydrogenoid atom 686
Fig. 7.27 Plot of the radial probability densities 687
Fig. 7.28 Periodic table of elements 689
Fig. 7.29 Schematic diagram of the LCAO function $f(E)$ 692
Fig. 7.30 Spheroidal coordinates for the H_2^+ ion 692
Fig. 7.31 LCAO energy solutions for the H_2^+ molecular ion 694
Fig. 7.32 Symmetric and antisymmetric states of the ground level of the H_2^+ molecule 694
Fig. 7.33 Singular limit for light 699
Fig. 7.34 Bonding and antibonding 701
Fig. 7.35 Bonding and antibonding in phase and out of phase 702
Fig. 7.36 Water molecule 702
Fig. 7.37 Electronegativity 703
Fig. 7.38 Entropy with absolute zero temperature 707
Fig. 8.1 Connection categories–logic–computation 718
Fig. 8.2 Example of olog 722
Fig. 8.3 Example of pullback 724
Fig. 8.4 Example of pushout 725
Fig. 8.5 Example of Equaliser 726
Fig. 8.6 The structure of category as an olog 728
Fig. 8.7 A deterministic finite state machine 732
Fig. 8.8 Different models of finite state machine 747
Fig. 8.9 Example of cobordism 752
Fig. 8.10 Cobordism representing a Feynman diagram 753
Fig. 8.11 The Boolean algebra sa a graph 760

Fig. 8.12 Logical linear subspaces 760
Fig. 8.13 The $\mathcal{B}_1$ network 766
Fig. 8.14 The sets X_1 and X_2 767
Fig. 8.15 The $\mathcal{B}_2$ network 768
Fig. 8.16 Three basic sets of $\mathcal{B}_3$ 770
Fig. 8.17 2D LVS 775
Fig. 8.18 Universal chart of metabolism 789
Fig. 8.19 Selenocysteine 790
Fig. 8.20 The world of the *Escherichia coli* 792
Fig. 8.21 Markov blanket 793
Fig. 8.22 The organism as a Markov blanket 794
Fig. 8.23 Signal–code–function 795
Fig. 8.24 *E. coli*'s motion 796
Fig. 8.25 Code–signal–function 803
Fig. 8.26 Inference circle 815
Fig. 8.27 Interpreted ontology 820
Fig. 8.28 Olog of the generalised quantum theory 821

List of Tables

Table 1.1 Derivatives of some basic functions 16
Table 1.2 Actions of the Pauli matrices 98
Table 1.3 Elementary particles 103
Table 3.1 Decoherence rate and timescale 237
Table 3.2 Truth table for conjunction 276
Table 3.3 Truth table for inclusive disjunction 276
Table 3.4 Truth table for exclusive disjunctions 276
Table 3.5 Actions of the Pauli matrices 314
Table 3.6 Possible events in a teleportation process 315
Table 3.7 Example of the BB84 protocol in the absence of eavesdropper 319
Table 3.8 Example of the BB84 protocol in the presence of eavesdropper 319
Table 4.1 Truth table for Toffoli gate 370
Table 4.2 The Clifford group 393
Table 4.3 The transformations displayed in circuit (4.80) 394
Table 4.4 The transformations displayed in circuit (4.81) 395
Table 7.1 The six quarks 630
Table 8.1 Table of actions of a finite-state machine 731
Table 8.2 Original model X 746
Table 8.3 New model Y 747
Table 8.4 Naturality table for a : ▲ ⟶ ▲ 748
Table 8.5 Naturality table for b : ▲ ⟶ ▲ 749
Table 8.6 Naturality table for some transformations of diagram (8.77). The three columns on the left after the first (ID) column and the three columns on the right bring to the same result through different paths 749
Table 8.7 Sum and product of IDs 762
Table 8.8 The IDs of the self-dual variables X_1 and X_2 for $\mathcal{B}_2$ 767
Table 8.9 The different classes generated in $\mathcal{B}_2$ 768

Table 8.10 The generation of the basic sets of $\mathcal{B}_3$ and their complements 769
Table 8.11 The two main codes of basic sets of $\mathcal{B}_3$ 769
Table 8.12 The 16 main codes for $\mathcal{B}_4$ 772

Chapter 1
Summary of the Basic Elements of the Theory



The aim of this chapter is to introduce some basic notions of quantum mechanics (=QM). The physicist or the scholar who already knows this background may skip it. After having introduced the basic historical reasons that led to quantum theory with the final result of the Schrödinger equation, I introduce the reader to the two basic principles of this theoretical building: the superposition and the quantisation principles. This will be the appropriate context for introducing the basic physical observables, like position and momentum as well as time and energy. It will be shown that all basic transformations in QM are unitary (reversible). After an examination of the non-commutativity among quantum observables (strictly related to quantisation) with the definition of angular momentum and the consequent uncertainty relations, some further principles are presented: Pauli's exclusion principle with the notion of spin and the two statistics (Fermi–Dirac and Bose–Einstein) ruling the two families of quantum particles: fermions and bosons, respectively; a first and rough version of the complementarity principle; finally a brief presentation of the heuristic correspondence principle follows. In the final section of the chapter, we shall deal with the formalism of the density matrix, very helpful for the notion of quantum-state space allowing the distinction between pure states and mixtures. Moreover, this formalism will be particularly relevant for distinguishing between entangled and product states and for dealing with compound systems (total and marginal states).

1.1 The Beginning of Quantum Mechanics

1.1.1 From the Classical to the Quantum World

Planck's Solution to the Black Body Problem

Let us summarise in few words some of the main discoveries that led to the quantum theory. The radiation of a black body played an essential role in this history. A black

© Springer Nature Switzerland AG 2019

G. Auletta, *The Quantum Mechanics Conundrum*,
https://doi.org/10.1007/978-3-030-16649-6_1

body is an idealised physical body (like a sphere) that absorbs all incident electromagnetic radiation falling on it. The German physicist Gustav Kirchhoff (1824–1887) established that "the quality of heat radiation produced in an enclosure surrounded by any emitting or absorbing bodies whatsoever is entirely independent of the nature of such bodies".[1] Moreover, he introduced, between 1859 and 1861, the concept of a black body in the modern sense of the term and achieved results that contain a series of elements, such as the constant between emission and absorption, basic also for quantum physics.[2] Nevertheless, classical physics could not provide a solution to the problem of the balance of energy inside the black body because, according to the calculations based on its laws, energy density (i.e. energy per unit volume) would grow exponentially in contrast with experimental results.

The work representing the first step of quantum physics is the two articles published on 1900 by the German physicist Max Planck (1858–1947) in the Discussions Paper of the Society of German Physics.[3] In the aforesaid articles, he offered a theoretical explanation of the behaviour of the black body by means of the famous (Plank) equation

$$E = nh\nu, \tag{1.1}$$

where E is the energy, ν the frequency of the radiation (Fig. 1.1), $n = 1, 2, \ldots$ (i.e. n is a positive integer: $n \in \mathbb{N}$, where $\in$ denotes set membership), and

$$h = 6.626069 \times 10^{-34}\,\mathrm{J \cdot s}, \tag{1.2}$$

as expressed in Joule $\mathrm{J} = \mathrm{kg\,m^2 s^{-2}}$ times seconds, is a new nature constant that has the name of its discoverer.[4] Note that Planck's constant, as stressed by Planck himself, who called it the quantum of action,[5] has the same physical dimensions of action: energy in time, $[\mathrm{M}][L^2][T^{-1}]$, where energy has the dimensions $[\mathrm{M}][\mathrm{L}^2][\mathrm{T}^{-2}]$.[6] Equation (1.1), instead of postulating (as classical physics did) that the exchange of energy between the heated object and the radiation this object emits is made in a continuous way, tells us that there are no frequencies except those possessing a discrete amount of energy determined by the number n: "the continuous transformation of energy into radiation could be avoided assuming that energy is forced, from the beginning, to remain grouped in certain quanta".[7] Planck assumed in fact that the emission of radiation could be modelled treating the elements of the surface

[1] Planck (1922).

[2] See Kuhn (1978, Chap. 1). Another important source of Planck contribution is the work on irreversibility and entropy of Ludwig Boltzmann (Kuhn 1978, Chap. 2).

[3] Planck (1900a, b). The reader interested into the general historical issues of QM may have a look at Jammer (1966) and Mehra and Rechenberg (1982–2001), to which I shall also sometimes refer in the following.

[4] Throughout this book I shall use the SI system of units.

[5] Planck (1906).

[6] The basic physical units are length L, time T, mass M, electric current I, temperature Θ, quantity of substance N, and luminous intensity J.

[7] Letter of Planck to R. Williams, 7 October 1931.

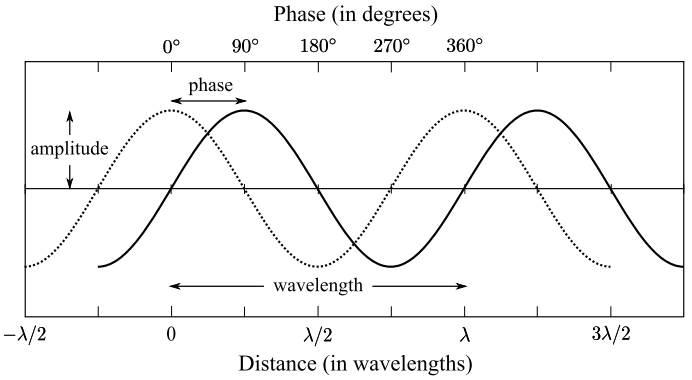


Fig. 1.1 Snapshot of two waves of the same wavelength and amplitude but with different phases at a particular moment in time. The *amplitude* of the wave is the distance between a peak and the baseline. Its *wavelength* $\lambda = c/\nu$, where c is the speed of light, is the distance between adjacent peaks and is inversely proportional to the *frequency* ν, which is the number of peaks that go through a given point in a given time interval (the larger is the frequency, the shorter the wavelength). It has, therefore, the dimension inverse to time: $[T^{-1}]$. The *phase* of a wave (solid line) relative to a reference wave (dotted line), whose peak corresponds to 0° phase, is the distance between their respective peaks (modulo the wavelength). The two waves shown here are 90° out of phase. Adapted from Auletta and Wang (2014, p. 19)

as a collection of resonators each oscillating at a specific frequency ν. As a consequence, Planck introduced a new formula for the (discontinuous) spectrum of the black body radiation (on which I shall come back). Thus, energetic discontinuity was introduced, for the first time, although in Planck's mind this was only a "purely formal hypothesis". Note finally that Eq. (1.1) connects for the first time a dynamical quantity (energy) with a kinematic one (frequency).[8]

The Photoelectric Effect

In 1905 the German–Jew physicist[9] Albert Einstein (1879–1955) demonstrated that the photoelectric effect, consisting in the emission of electrons from a metal plate when lighted up, could be explained if one was willing to admit that the light that produces the effect consisted in discontinuous quanta of energy according to Planck's equation. The consequence is that the kinetic energy of these electrons does not depend on the intensity of the incident radiation (as it was classically expected), but solely on the composition of the spectrum of the light that illuminates the plate: the smaller the wavelength, the greater the kinetic energy, showing that there is a maximum wavelength (lowest frequency) over (under) which no electron can be

[8] Planck (1931). Dynamical variables are momentum, angular momentum, energy, while kinematic ones are space, angle, time.

[9] But, as we shall see in the following, he also developed a robust philosophical reflection.

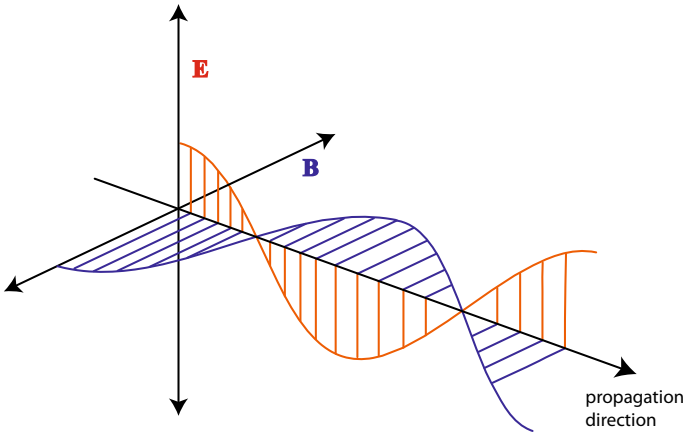


Fig. 1.2 The classical picture of light. Note that both the electric (in red) and the magnetic (in blue) components of the electromagnetic fieldoscillate into directions that are orthogonal to the propagation direction (as well as orthogonal to each other). Adapted from Auletta et al. (2009, p. 20)

extracted. The ultraviolet light (which is particularly energetic due to its short wavelength and high frequency) has the capability of extracting electrons off the surface of a metal plate. Moreover, the effect is almost instantaneous, a behaviour that is not understandable if the light was a wave. Thus, Einstein retook Planck's hypothesis and adapted it to a new physical problem. Supposing that monochromatic light (having a fixed wavelength), that is to say, light of a single frequency, was formed of millions of corpuscles carrying the same quantum of energy (whose number at a fixed time represents the intensity of light), the ejection of the electron off the metallic plate could be interpreted as a collision between those corpuscles of light and the electron, so that photons (this was the name given later on to the quanta of light) would communicate their quantum of energy to the electron, which in turn would use this energy to evade the attraction of the nucleus and to depart from the metal plate with a certain velocity. "If all quanta of energy from the incident light transfer their energy to electrons independently from the rest of quanta, then the distribution of velocity of electrons, that is to say, the quality of the resulting cathodic radiation, will be independent of the intensity of the incident light; on the other hand, the number of electrons emitted by the body will be proportional to the intensity of the incident light."[10]

Thus, the photoelectric effect was another one of the problems that could not be explained by means of the classical theories of British physicists Isaac Newton (1643–1727) and James C. Maxwell (1831–1879). In particular, with Einstein's explanation of the photoelectric effect, we have the quantisation of electromagnetic radiation, which until that moment was believed to be continuous (Fig. 1.2) also by

[10]Einstein (1905).

Planck, who indeed used Maxwell's theory of classical electromagnetic fields for the resonators' absorption and emission.[11] In fact, the Planck's postulate dealt with *emitted* radiation only (and so it was still possible to interpret quantisation as due to the light's interaction with the atomic-molecular structure of the internal surface of the black body[12]), while Einstein dealt with *incident* radiation, which suggests that quantisation is a property of light as such and not a particular behaviour of light in very specific circumstances.[13] It is important to stress that (according to the American physicist, historian and philosopher of science Thomas Kuhn (1922–1996)), Planck did not consider energy as restricted to discrete set of values, although accepting the reality of quantisation.[14] Thus, according to Einstein's result, radiation was made up of quanta of energy that abided by the same equation introduced by Planck. Since photons are kind of particles, we can associate to their energy a linear momentum, although strictly speaking they are massless corpuscles:

$$E = cp, \tag{1.3}$$

where c is the speed of light (ca. 299, 792, 458 m/s) and

$$p := mv \tag{1.4}$$

is the classical magnitude of the momentum (I do not consider its direction), where v is the velocity of an ordinary body (whose physical dimension is $[\mathrm{L}][\mathrm{T}^{-1}]$) and m its mass. The symbol ":=" here and in the following means a definition.

1.1.2 The Quantum Postulate and Matter's Discontinuity

Quantum Postulate

The quantum postulate was formulated for the first time in 1913 by the Danish physicist Niels Bohr (1885–1962), who generalised the solution offered by Planck to the problem of the atom's stability (although doubts could be cast on whether he fully accepted the idea of a light corpuscle).[15] It is the basis of the later quantisation principle. We can distinguish two aspects here:

- A quantum of radiation has a frequency equal to its energy divided by Plank's constant h. This is the part due to Planck's contribution.

[11] See Kuhn (1978, Chaps. 3–4).

[12] As clearly pointed out in Born (1949, p. 80). As we shall see, Planck performed an inference that is called abduction.

[13] See also Bohr (1949, p. 202).

[14] See Kuhn (1978, Chap. 5).

[15] Bohr (1913).

- An atomic system can exist in particular or discrete states, each of which corresponds to a definite energy of the system, with electrons emitting or absorbing energy quanta when they leap from one level to another.[16] This is Bohr's contribution reflecting also the progress made on the explanation of the photoelectric effect.

We have already considered the first point of the postulate in the previous subsection. Bohr developed the second part when he brought Planck's hypothesis into the model of the atom, trying to give a solution to the instability of Rutherford's atomic model.

Rutherford's Atomic Model

The New Zealander-British physicist Ernest Rutherford (1871–1937) proposed a very ingenious planetary model. In fact, already at the beginning of the twentieth century, it was clear that atoms had internal structure, ceasing to be that spheres of homogeneous, elementary and indivisible matter of the atomists, a model that classical physics had inherited. At that time, it was also experimentally evident that negatively charged particles cloud around positively charged nuclei. Rutherford proposed that the atom should be considered as a kind of planetary system in which, in the simplest case of the hydrogen atom, a negatively charged electron rotates around a nucleus formed by a positively charged proton. But this model presented a problem because an electrical charge in accelerated motion (induced by a rotation about the opposite charge) emits radiation and, therefore, loses energy, then the negatively charged electrons that circled round the positively charged nucleus were bound to lose their energy and to fall in a spiral towards the nucleus. Now, this would contradict the high stability of matter (apart from radioactivity).

Bohr's Atomic Model

For this reason, Bohr introduced discontinuity into the very heart of the atom, according to which the electrons that "gravitate" around the nucleus can only do so in certain specific "orbits"; in addition, it is impossible for them to take an "orbit" below the one called the ground level (that corresponds to the minimal quantum of energy), so that they cannot crash into the nucleus. With these assumptions, Bohr not only conferred coherence to Rutherford's model, but he could simultaneously explain, although partially, the discrete spectrum of electromagnetic emission of atoms when bombarded by an external source of heat and the regularity of the chemical elements properties. However, his solution did not yet allow for an adequate explanation of the atom's structure, because it still conceives the paths of electrons in a classical form. Only the ripe quantum theory could solve this problem in a satisfactory way, as we shall see, by introducing non-commutativity of the conjugate observables and the Schrödinger equation.

[16] Auletta (2000, p. 17).

The Compton Effect

Quantisation was also confirmed by the discovery of the so-called Compton effect.[17] This effect consists of the diffusion of x-rays on the part of electrons of a certain material, so that the scattered wavelength of the emitted x-rays (electromagnetic radiation with short wavelength) was longer (denoting a loss of energy) than that of the incident radiation. This phenomenon can only be explained if the collision between the electromagnetic radiation of high frequency and the electron is interpreted as if it was an elastic collision between two particles, such that there is energy transfer between electron and photon: first a photon is absorbed by the electron and then the latter emits a photon with lower energy. Light, traditionally thought of as an undulatory phenomenon, with interference as the typical behaviour, acquired now, as a consequence of Einsteins's, Bohr's, and Compton's results, the typical behaviour of particles, i.e. that of collision. Light seemed to be a compound of particles without mass, particles that Compton will baptise "photons".[18]

The Principle of Continuity

At this point, it is already possible to venture that one of the most important consequences of the quantum postulate (and of the results summarised in the previous subsection) was the violation of the classical–mechanical Principle of continuity: with the words of the great German philosopher, logician, mathematician (one of the inventors of both the infinitesimal and combinatorial calculus), jurist, and theologian Gottfried W. Leibniz (1646–1716), *Numquam fit per saltum*.[19] Now, if there are discrete amounts of energy, if there are leaps, a continuous evolution of a system in time cannot be upheld. This rupture of continuity has important consequences as regards the understanding of the transition from the quantum world into the macroscopic world, a problem that will be treated more extensively later once we have set out the principles that govern the quantum world.[20]

Another consequence is that the finite value of Planck's constant makes impossible to reduce to zero the margin of error for any given measurement, as it was still assumed by classical mechanics (=CM). If the Principle of continuity collapses, it will not be possible to indefinitely increase the precision of measurements approximating more and more the exact value of the measured parameter: therefore, Planck's constant imposes a maximum limit to the reduction of error.[21] Also classically, to measure a system perfectly, is unjustified, as pointed out by the German physicist

[17] Compton (1923).

[18] Note that Bohr was dismissive of the concept of a light's particle up to the Compton's experiment (Home and Whitaker 2007, p. 25).

[19] Leibniz (1715–16, p. 25). See also Boscovich (1754).

[20] The reader interested in this subject may have a look at Auletta (2004); see also Auletta (2000, pp. 565–67).

[21] Auletta (2000, p. 23).

and mathematician Max Born (1882–1970).[22] In fact, even classically, if we try to make a metre smaller and smaller, we shall reach a point in which it is perturbed by Brownian motion or the molecular structure of matter makes the required continuity impossible.

Thus, the assumption that an infinitely precise measurement (a null error) be possible is beyond physical foundation.[23] It is not by chance that the American physicist John A. Wheeler (1911–2008) affirmed that the physical continuum which CM takes for granted is rather an idealisation.[24] A real physical continuum cannot be experienced; at the opposite, one can make experience of singularities, symmetry ruptures, and points of crises.[25]

Divisibility of Matter

Another revolutionary consequence in relation to CM is the fact that, through the application of the quantum notion to physics, the principle of divisibility of physical matter ad infinitum, shared by both natural philosophy and CM, inevitably crumbles. In fact, taking into account the Principle of convertibility mass–energy, according to which mass and energy are interchangeable due to Einstein's formula (see also Eq. (1.3))

$$E = mc^2, \tag{1.5}$$

where I recall that E is the energy, m the mass of the system, and c the speed of light (note that c^2 plays the role of a conversion constant between mass and energy),[26] it is evident that infinite divisibility is no longer possible if we assume that nature is made of quanta. In other words, if energy has a lower limit in the quantum according to Eq. (1.1), applying the principle of convertibility of mass and energy, we see that quantisation can be extended to the manifestations of matter as mass. As a matter of fact, all components of matter that appear by now to be irreducible have a specific mass.

A Bit of History

These problems are not new and have a long tradition in philosophy. We may recall in this context the Aristotelian doctrine on the impossibility of the infinite division of

[22] Born (1949, Chaps. 6–7). "How small a fraction of a grain of millet must I demand is put on the first square of the chess board if after doubling up at every square I end up having to payout only a pound of millet? It would be a figure of such smallness as to have no meaning as a figure for a margin of error" (Anscombe and Elizabeth 1971, pp. 95–96).

[23] See also Margenau (1950, Sect. 6.5).

[24] Wheeler (1988). See also Poincaré (1902, p. 51) or Poincaré (1905, pp. 61–62).

[25] As recalled in Auletta (2000, pp. 567–68).

[26] For a summary of the historical background of the formula see Rindler (2001, p. 114).

matter, by which, the ancient Greek philosopher Aristotle (385–322 B. C.) affirmed that matter cannot be equated with extension, which, at the opposite, is by definition, infinitely divisible.[27] At the opposite, the famous French mathematician and metaphysician René Descartes (1596–1650), equating matter with extension (*res extensa*) maintained the divisibility of matter ad infinitum.[28] Leibniz accepted this Cartesian vision of the infinite divisibility of extension but he felt the need to postulate the indivisible monadic unit in order to preserve continuity (a subject on which we shall come back).[29] The quantum postulate gives us grounds to refute the identification of matter with extension: matter cannot be equated with extension as long as the latter is taken to be infinitely divisible. We may then take infinite divisibility for granted as mathematical notion and as far as it may contribute to deal with specific problems in certain situations. But this divisibility, as it happens with the classical physical continuum, is purely ideal. The Universe bears a minimum of subdivision, which depends on the quantum of action.

1.1.3 The Road to the Schrödinger Equation

Young's Experiment

We have seen how Einstein solved the photoelectric effect problem by assuming that light was discontinuous. This collided with the—at that time established—conception, which held that the nature of light was wave-like and not corpuscular-like. The Compton effect came to strengthen the corpuscular view. On the other hand, the experimentally ascertained interference and diffraction effects of light could not just be explained corpuscularly alone (Fig. 1.3): Panel (b) shows the famous experiment first carried out in 1803 by the British polymath Thomas Young (1773–1829). A source of monochromatic light emits light which is made to pass through two vertical slits in a screen to a second screen behind it, which actually is a photographic plate. If light is cast long enough to impress the plate, but not too long so as not to overexpose it, we will observe on this plate alternating vertical bands, more or less bright or more or less dark, which will allow us to describe a curve representing the intensity of the incoming light: this is the interference phenomenon (Fig. 1.4). If, at the opposite, light was made of quanta, it should display the behaviour shown in

[27] The physical continuum, according to Aristotle, cannot be made up of indivisibles. An indivisible object should not have parts. Continuum is always divisible in other divisible parts. See Aristotle (1950, VI, 1, 231a18–b19). The extension is a continuous one and, therefore, it is divisible. See Aristotle (1950, III, 6, 206b1-20).

[28] Descartes (1641). Descartes was the father of analytical geometry succeeding in unifying mathematical analysis and geometry thanks to the introduction of the Cartesian coordinate system. However, we shall deal here mainly with his metaphysical theses. It may be also noted that equating matter and extension deprives the former of its main causal resources of which we shall have to say more in the following.

[29] Leibniz (1712–14).

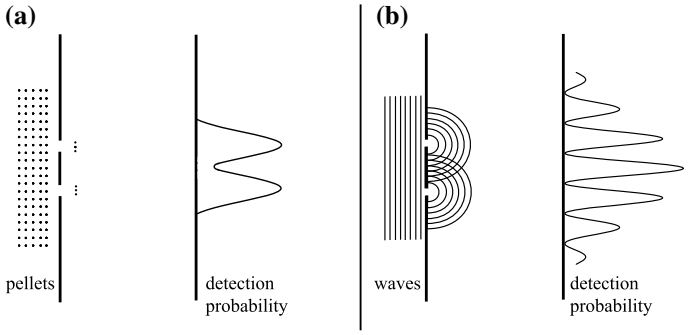


Fig. 1.3 Suppose that we let particles **a** and waves **b** go to wall with two slits and then see what is the distribution on the screen located on the right. In the case of particles, we have two bump-like distributions. At the opposite, when a wave goes through the two slits, it gives rise to typical interference phenomena that result in the distribution shown in Panel (**b**). Adapted from Auletta and Wang (2014, p. 75)

Panel (a) of Fig. 1.3[30]: interference is no longer present here. Thus, light became, for the early twentieth-century physics, a puzzle, for it displayed, at once, a particle-like behaviour and a wave-like behaviour that seemed to contradict each other.

de Broglie's Matter Waves

But oddities would not stop there. In 1925 a French physicist, Louis de Broglie (1892–1987), in his doctoral dissertation proposed an original mathematical hypothesis which dealt with matter and, particularly, the electrons as if they were waves, i.e. endowing the 'particles' with a wavelength that is mathematically related, through Plank's constant, to the momentum of the particle (see also Eq. (1.3)):

$$p = \frac{h}{\lambda} = \frac{h\nu}{c}, \tag{1.6}$$

where I recall that p is the magnitude of the momentum, and therefore related to the (kinetic) energy through the equation

$$E = \frac{p^2}{2m}, \tag{1.7}$$

which is kinetic counterpart of Einstein's formula (1.5). In accordance with Eq. (1.6), and thanks to Eqs. (1.1) and (1.5), Compton[31] was able to derive that in the Compton

[30] Among the first interpretations of light, a considerable stress was made on its geometric properties of e.g. reflection (Newton 1704).

[31] Compton (1923).

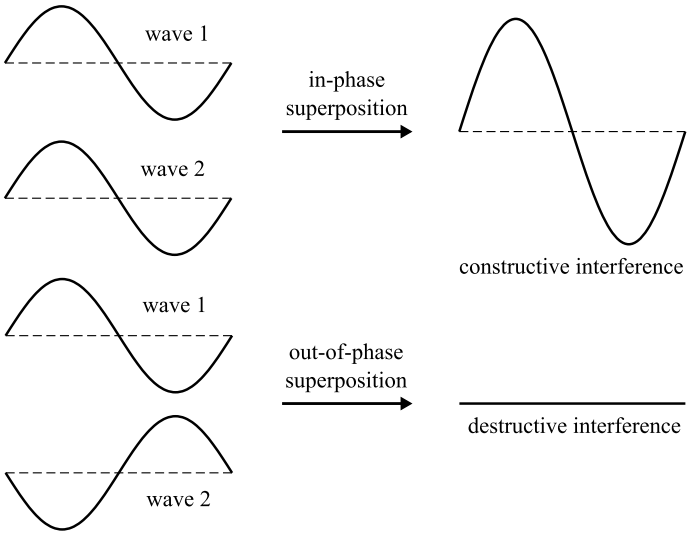


Fig. 1.4 Constructive and destructive interference depending on the phases of the waves. In the former case, they are in phase, in the latter case they are completely out of phase (intermediate possibilities result in intermediate interferences). Adapted from Auletta (2011, p. 12)

effect a specular change occurs in the wavelength of the electron

$$\lambda_C = \frac{h}{mc}, \tag{1.8}$$

relative to that of the radiation, and this is called the Compton wavelength of the electron.

Of course, if matter quanta are kind of waves, they should display typical interference phenomena. However, De Broglies's argument was more complex.[32] Given that many forms of energy are reducible to motion, and that, according to the principle of convertibility (1.5), mass and energy are interchangeable (even in repose), there must be a certain motion associated with mass. Since a corpuscle is a traceable structure liable to being in repose, the motion associated with mass can be defined in a very small region. This motion is somewhat analogous to a rotation or a vibration, i.e. to a periodical motion with a certain frequency. De Broglie interpreted this oscillation as a kind of internal clock of the particle. That is to say, the wave ascribed here to the particle is not a monochrome one, with a limitless extension in space, but rather a wave packet (a wave concentrated in a relatively small region), the maximum amplitude of which shifts at the same velocity as that of the monochrome wave. In other words, de Broglie liked to keep distinct the internal clock from the external wave determining diffraction effects, which is produced by the motion of the particle.

[32] de Broglie (1924).

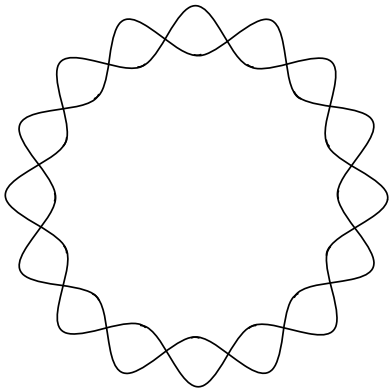


Fig. 1.5 The oscillation of an electron needs to be a certain precise frequency to close the curve at a certain orbital, otherwise, the unclosed orbit would self-destroy itself. Adapted from Auletta (2011, p. 163)

Postponing the analysis of de Broglie's standpoint, the crucial point is that De Broglie conferred to the classical material particles an undulatory behaviour, and he expected to prove that material bodies were subject to the phenomenon of diffraction. This hypothesis was confirmed thanks to a famous experiment performed in 1927 by the American physicists Clinton J. Davisson (1881–1958) and Lester H. Germer (1896–1971), further confirmed in the same year by the British physicist George P. Thomson (1892–1975), definitively proving the phenomenon of electron diffraction.[33] There is here an important consequence for the model of the bounded electron in an atom: we can now understand, through de Broglie's formulation, that the only stable states of this system are those for which the electron has a wave with a wavelength such that the circumference of the orbital shell is its multiple (with the result of a stable orbit, as displayed in Fig. 1.5). In such a way, the electron's orbit could no longer be understood in classical terms.

Wave and Particle

Thus, matter can display properties traditionally attributed to light, and this allows us to treat the quantisation of light as a particular case of a general behaviour (bearing in mind that the mass of the photon is null), an example of how closely the properties of a particle are to those of the electromagnetic radiation. Therefore, the main result of de Broglie is that *both* undulatory and corpuscular behaviour can be attributed to the *same* quantum system as well as to *all* quantum systems (it does not matter whether light or matter): by means of de Broglie's theory we can interpret mathematically this double phenomenology. However, this is quite puzzling from a physical point of view. Classical physics had treated the phenomena of radiation and matter in two different disciplines (electromagnetism, from which also the special relativity theory stems, and CM, respectively) and ascribed to the former wave-like properties whilst

[33] Davisson and Germer (1927). Note that de Broglie did not agree with their statistical interpretation of the phenomenon.

it attributed to the latter corpuscular properties. In fact, Einstein thought that these two disciplines were not fully compatible.[34] Now, what is amazing with QM was its capability to unify these different phenomena but by reproducing such a duality inside a single theory for *all* physical systems. In fact, coming back to the experiment with photons displayed in Fig. 1.3, it is evident that by closing alternatively either slit, what we obtain are two curves the sum of which is by no means the initial curve, that is, the interference figure. In other words, a photon does not behave the same, when the two slits are open or when either is closed. And in the same way behave matter particles.

To avoid this duality, one could introduce the notion of extended particles, particles that can behave either as a wave or as a corpuscle according to the context. If this is understood in classical terms, it is untenable, as I shall show. On the other hand, if it is understood as a reformulation of some basic aspects of quantum theory, it is not totally clear what is the gain relative to it. It is not by chance that the British astronomer and physicist Arthur S. Eddington (1882–1944) spoke of wavicles for denoting that conundrum.[35] Another possibility is represented by the Complementarity principle, which will occupy us later. Finally, as shown in the following, a reformulation of this principle will be perhaps the best solution to the problem capturing also the basic tenet of Eddington's wavicles.

Dirac's Notation

Let us take a short break and introduce a formal instrument that will play a crucial role in the following, although its discovery was much later than the developments we are discussing now. In fact, Paul Dirac (1902–1984), a Cambridge physicist who gave fundamental contributions to both QM and quantum field theory, formulated a simple algebra for the treatment of the quantum-mechanical problems.[36] He introduced the general concept of state vector (i.e. a vector in the Hilbert space $\mathcal{H}$, conventionally represented as a column vector) $|\cdot\rangle$, called *ket* (where $\cdot$ is a placeholder). A *Hilbert space*, after the name of the German mathematician David Hilbert (1862–1943), is a (in QM generally complex) linear vector space endowed with scalar product that is complete and separable having a countable dense subset (i.e. there exists a sequence of elements of the space such that every nonempty open subset of the space contains at least one element of the sequence[37]), as it happens for the Euclidean space. Hilbert spaces can have finite or infinite dimensions. In all generality, a *vector space* consists of a set of vectors, a rule for summing them, and a rule that allows product between complex numbers.

---

[34] Einstein (1993, pp. 160–61). It is interesting to remark that the first hint towards the wave particle duality, at least for light, can be found in Einstein (1909).

[35] Eddington (1929, p. 201).

[36] Dirac (1939). Thus, I do not follow here a strict historical reconstruction.

[37] https://en.wikipedia.org/wiki/Separable_space.

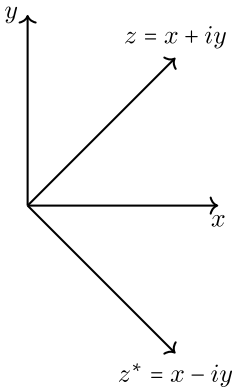


Fig. 1.6 Real ($x = \Re(z)$) and imaginary ($y = \Im(z)$) parts of the complex number z. Note that z^* is given by a reflexion of the complex plane in the real axis

Dirac also introduced the dual $\langle \cdot |$, called *bra* (or covector), so that each ket $|\psi\rangle$ corresponds to a state of the system, and the corresponding bra $\langle\psi| = |\psi\rangle^\dagger$ is its complex conjugated with transposition from a columnar vector to a row one. Thus, the term *transposed conjugate* or *adjoint* units two properties:

- The complex conjugate z^* of a complex number z is obtained by inverting the sign of the imaginary parts: $(x + iy)^* = x - iy$, with $x, y \in \mathbb{R}$ (where $\mathbb{R}$ is the set of real numbers) and $i = \sqrt{-1}$ being the imaginary unit.
- The transposed of a vector transforms a row vector into a column vector or vice versa.

The same is true for operators, as we shall see. Note that the real part x and the imaginary part y of z (Fig. 1.6) are given by

$$\Re(z) = \frac{z + z^*}{2} \quad \text{and} \quad \Im(z) = \frac{z - z^*}{2i}, \tag{1.9}$$

respectively. Mathematically, bras are vectors in the dual Hilbert space $\mathcal{H}^*$.

Note that this formalism was an attempt to supply a neutral quantum-mechanical formulation[38] with respect to an undulatory mechanics (set forth by Schrödinger[39]) and a matrix mechanics (set forth by the German physicists Werner Heisenberg (1901–1976), Max Born (1882–1970), and Pascual Jordan (1902–1980)[40]). These formulations were eventually proven to be equivalent thanks to the work of Jordan, Klein, Wigner and von Neumann. Let us now deal with these two formulations.

[38] Dirac (1930, Chap. 1). For a long time, this has been a reference textbook on QM.

[39] On this see also Schrödinger (1926e).

[40] Born and Jordan (1925), Born et al. (1926).

Schrödinger Equation

De Broglie's ideas were promptly taken up, developed and modified by Erwin Schrödinger (1887–1961), an Austrian polymath, thereby giving birth to undulatory QM and to the equation bearing his name[41]: since particles have a wave-like behaviour, the obvious inference is that they can be described by means of a wave-like equation.[42] Schrödinger equation exemplifies and governs the behaviour of both material particles and light, and, applied to De Broglie's waves, it allows the description of the behaviour of the electron as well as the reconstruction of the atoms' spectrum. From this new standpoint, the atom's planetary model was definitively substituted by its undulatory version, in which electrons do not trace specific trajectories on their orbital plane.

Let us consider an arbitrary quantum state $|\psi(t)\rangle$, where $\psi(t)$ is a function of time. Then, we can formulate the Schrödinger equation as:

$$i\hbar\frac{d}{dt}|\psi(t)\rangle = \hat{H}|\psi(t)\rangle\ , \tag{1.10}$$

which tells us that this state evolves in time (due to the first-order differential on the left-hand side (LHS) that expresses the variation of the state across time) in proportion with the reduced Planck constant (particularly helpful for periodic phenomena)

$$\hbar := \frac{h}{2\pi} \tag{1.11}$$

times the imaginary unit i in a way that is ruled by the energy operator $\hat{H}$, also called Hamiltonian operator, after the name of the Irish physicist and astronomer William Rowan Hamilton (1805–1865), present on the right-hand side (RHS).[43] The concept of operator will be dealt with below. By now, let us fix the idea that they rule state transformations. In order to distinguish operators from numbers we shall put a hat on them as I already did for energy. Some examples of derivatives are shown in Table 1.1.

In the one-dimensional (=1D) case, i.e. a quantum moving along a line, the Hamiltonian, representing the total energy of the system, takes the form

$$\hat{H} := \frac{\hat{p}_x^2}{2m} + V(\hat{x}, t), \tag{1.12}$$

[41] Schrödinger (1926a, b, c, d). For historical reconstruction see Jammer (1966, Sect. 5.3).

[42] For a presentation of the Schrödinger equation the reader may have a look at Auletta et al. (2009, Chap. 3), Auletta and Wang (2014, Chap. 7). These two textbooks will be very often quoted since they can provide both formal understanding and further reference about several topics of the present book: the former gives an account of QM for physicists, the latter is addressed to a wider audience and therefore I suggest the inexpert reader have a look at it for a better understanding of what follows. Since they are very much quoted, the final index of authors does not report these entries.

[43] To the reader who needs a refresh in derivation and integration I strongly recommend (Auletta and Wang 2014, Sects. 6.2 and 6.4–6.5).

Table 1.1 Derivatives of some basic functions

Type	Function	Derivative
Power	x^{α}	$\alpha x^{\alpha-1}$
Exponential	e^{x}	e^{x}
Logarithmic	$\ln x$	x^{-1}
Trigonometric	$\sin x$	$\cos x$
	$\cos x$	$-\sin x$

where the first component represents the kinetic energy $\hat{T}$ ($\hat{p}_x$ is the 1D momentum while $\hat{x}$ is the 1D position), while the second component represents a potential (depending on some external force or field) that in the case in which the particle is free (by definition not subjected to any force or field) is obviously 0 reducing to a counterpart of Eq. (1.7). In short, what Schrödinger equation tells us is that quantum systems evolve ruled by a "wave" equation but according to a quantised expression (due to the presence of the reduced Planck constant).

I would like to point out an important property of the quantum state $|\psi\rangle$ (and of the Schrödinger's equation), i.e. its linearity.In its generality, *linearity* is a combination of two mathematical properties called additivity and homogeneity: for any function f and arbitrary variables x, y, *additivity* means that

$$f(x+y) = f(x) + f(y), \tag{1.13a}$$

while *homogeneity*, for any number $\alpha \in \mathbb{C}$ (where $\mathbb{C}$ denotes the set of complex numbers), means that

$$f(\alpha \cdot x) = \alpha \cdot f(x). \tag{1.13b}$$

Whatever attempts to put non-linear terms into the quantum wave equation have had no success experimentally, as we shall see later.

Reversibility and Unitarity

As in CM, all relevant transformations in QM (apart from measurement, as we shall see) are linear and reversible. In fact, the Schrödinger equation is linear and reversible. I have already summarised what is linearity. We express the second concept by means of the operator $\hat{U}_t$ that is called *unitary*, where the subscript t denotes that this unitary operator rules the time evolution of quantum systems. Indeed, a product of $\hat{U}_t$ with its inverse $\hat{U}_t^{-1}$ will give the identity operator $\hat{I}$, i.e. the null transformation (the transformation that does not change the initial state). Since in the specific case of this operator its inverse $\hat{U}_t^{-1}$ coincides with its transposed conjugate or *adjoint* $\hat{U}_t^{\dagger}$, we have

$$\hat{U}_t \hat{U}_t^{\dagger} = \hat{U}_t^{\dagger} \hat{U}_t = \hat{I}. \tag{1.14}$$

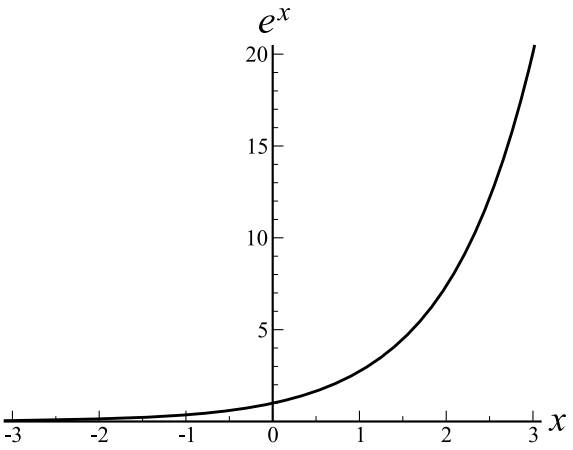


Fig. 1.7 Plot of the exponential function e^x

Thus, we can write the solution of the Schrödinger equation (1.10) by using unitary operators:

$$|\psi(t)\rangle = \hat{U}_t \, |\psi(0)\rangle \, , \tag{1.15}$$

where on the RHS we have the initial state at time 0 and on the LHS we have the evolved state at time t. For the Schrödinger equation (1.10), the unitary operator takes the form:

$$\hat{U}_t = e^{-\frac{i}{\hbar}\hat{H}t}. \tag{1.16}$$

Exponential Function

Allow me to explain some mathematical notions introduced so far. The quantity $e^1 = 2.718281828\ldots$ is the exponential function and is the inverse of the natural logarithm function, i.e. $\forall x$ (Fig. 1.7),

$$\ln e^x = x \text{ and } e^{\ln x} = x, \tag{1.17}$$

where $\forall$ means "for every …". Note that the complex conjugate of e^{-in}, $n \in \mathbb{R}$, is e^{+in}. The exponential can be expanded as

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots = \sum_{n=0}^{\infty} \frac{x^n}{n!}, \tag{1.18}$$

where $n! = 1 \cdot 2 \cdot 3 \cdots (n-1) \cdot n$ is called the *factorial*. This leaves x^n in the numerator and $n!$ in the denominator for each term in the infinite sum. In the limit x approaches zero, we can keep only the lowest order term in the above expansion and

obtain

$$e^x \approx 1 + x \quad \text{for} \quad x \to 0, \tag{1.19}$$

where the symbol $\approx$ denotes that the LHS is approximately equal to the RHS. Similarly, the trigonometric functions $\sin x$ and $\cos x$ in the limit x approaches zero (measured in radians) can be approximated by

$$\sin x \approx x, \quad \cos x \approx 1 \quad \text{for} \quad x \to 0. \tag{1.20}$$

Equation (1.18) is an example of *Taylor series*, after the name of the British mathematician Brook Taylor (1685–1731). In general, a function $f(x)$ having an infinite number of derivatives can be Taylor-expanded about the origin ($x = 0$) as follows[44]:

$$f(x) = f(0) + \frac{f^{(1)}(0)}{1!}x + \frac{f^{(2)}(0)}{2!}x^2 + \frac{f^{(3)}(0)}{3!}x^3 + \cdots = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!}x^n, \tag{1.21}$$

where the indexes denote the progressive order of derivation. When a function can be expanded in this way it is said to be *analytic* at $x = 0$. This can be generalised to a series about any point a by replacing x by $x - a$ so to obtain

$$f(x) = f(a) + \frac{f^{(1)}(a)}{1!}(x - a) + \frac{f^{(2)}(a)}{2!}(x - a)^2 + \cdots = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x - a)^n. \tag{1.22}$$

Action and Lagrangian

In conclusion of the present section, I shall give a short account of the main elements of CM (a kind of counterpart of QM). Its current formulation, which is centred on energy functions (following a view that was originally proposed by G. Leibniz)[45] rather than on the notion of external force, as for the previous Galilean–Newtonian mechanics, is due to the path-breaking work of Italian mathematician and astronomer Joseph–Louis Lagrange (1736–1813),[46] followed by the contributions of the already mentioned W. Hamilton, the German mathematician Carl G. J. Jacobi (1804–1851), the French mathematician and physicist Siméon Denis Poisson (1781–1840), and others, arriving to the French mathematician, physicist, engineer, and philosopher of science J. Henri Poincaré (1854–1912) who represents somehow the threshold between this classical theory and the twentiethth century revolutions in physics. The

[44] Byron and Fuller (1969–70, I, Sect. 5.6). I strongly recommend this classical book for several problems of the mathematical calculus and notation used here. For the reader needing a more basic teaching in mathematics, I recommend the three-volume (Apostol 1969).

[45] Leibniz (1686, 1695).

[46] Lagrange (1788–89).

reason for interest is that the Hamiltonian is not the only classical function that relates kinetic and potential energy.[47] Another one is the Lagrangian,which plays also an important role in quantum field theory since it is Lorentz (or relativistically) invariant.[48] Moreover, this quantity is strictly connected with an important physical notion: action (I recall that its physical dimensions are energy in time, i.e. $[M][L^2][T^{-1}]$). In fact, all classical systems obey the *principle of least action*:

$$S = \int_{t_1}^{t_2} L(q, \dot{q}, t)dt, \tag{1.23}$$

where S is the action and the trajectory of the system must be such that it takes the minimal value in the time interval $t_2 - t_1$. In other words, we expect a classical system to take the trajectory that 'costs' less action. The expression above is the (Riemann) integral over this time interval and the quantity dt, called the *differential* of t, represents infinitesimal increments or differences of the integration variable t (an integral is a continuous 'summation' and is, therefore, the opposite operation relative to derivation). The quantity $\dot{q}$ is the first time derivative of the (generalised) spatial coordinate q and it, therefore, represents the speed. Generalised coordinates, like $q, \dot{q}$, represent the minimum number of independent coordinates that define the configuration of a (e.g. multiparticle) system. For simplicity, I am dealing here with pure magnitudes. L is the *Lagrangian function*,after the name of J.-L. Lagrange, and is defined as the difference between the kinetic energy, expressed in terms of the momentum p (Eq. (1.7)):

$$T(p) := \frac{p^2}{2m} \tag{1.24}$$

and the potential energy $V(q)$ expressed as a function of q and whose forms depends on the fields and forces to which to system is subject, i.e.

$$L := T(p) - V(q). \tag{1.25}$$

Suppose now that there is some small variation in the trajectory such that, at each time t, is a variation δq of the coordinate q as well as $\delta\dot{q}$ of the velocity $\dot{q}$. The condition of the minimisation of the action requires that this small fluctuation is 'absorbed' by the system, that is,

$$\delta S = \delta \int_{t_1}^{t_2} L(q, \dot{q}, t)dt = 0, \tag{1.26}$$

which, thanks to (with t kept constant in the first equation)

[47] For this examination I consider still the classical text (Landau and Lifshitz 1976) as fundamental.

[48] The reason is that the Hamiltonian selects energy as a privileged quantity, but since energy is conjugate with time, it would also select a special time coordinate, which is in contrast with Einstein's relativity (Kastner 2013, p. 51).

$$\delta f(q,t) = \frac{\partial f}{\partial q}\delta q \text{ and } \delta \int f(q,t)dt = \int \delta f(q,t)dt, \tag{1.27}$$

can be rewritten as

$$\int_{t_1}^{t_2} \left(\frac{\partial L}{\partial q}\delta q + \frac{\partial L}{\partial \dot{q}}\delta \dot{q} \right) dt = 0, \tag{1.28}$$

where

$$\delta \dot{q} = \delta \frac{dq}{dt} = \frac{d}{dt}\delta q. \tag{1.29}$$

I recall that, like ordinary derivatives, the partial derivative also has a simple geometric representation as the slope of the tangent at a given point. Assuming that $f(x, y)$ is a function of two variables, the graph of $f(x, y)$ defines a surface $z = f(x, y)$ in a three-dimensional space as shown in Fig. 1.8. To every point on this surface, there are an infinite number of tangent lines. Partial derivation corresponds to selecting a particular one of these lines and finding its slope. For the partial derivative of $f(x, y)$ with respect x, we are considering y as constant and x as variable, and for partial derivative with respective to y, we are considering x as constant and y as variable. In the former case, we trace a single curve $z = f(x, y_0)$ and the partial derivative $\partial f(x, y)/\partial x$ taken at a certain point gives the slope of the tangent to the curve at that point along the x direction. Similarly, in the latter case, we trace a curve $z = f(x_0, y)$, and the partial derivative $\partial f(x, y)/\partial y$ taken at a certain point gives the slope of the tangent to the curve at that point along the y direction. The rules for partial derivatives are the same as those for ordinary derivatives, except for holding some variables as constant.

Let us use the integration by parts, which is the transformation of an integral of a product of functions into the integral of their derivatives and anti-derivatives. For functions $u = u(x)$ with $du = u'(x)dx$ and $v = v(x)$ with $dv = v'(x)dx$ (where $f'(x)$ denotes the first derivative in x), we have

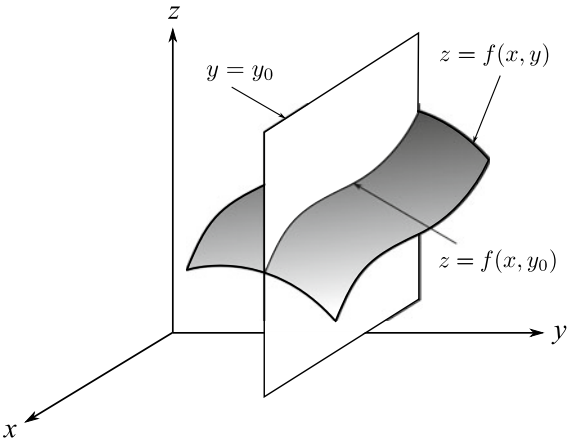


Fig. 1.8 Partial derivative with respect to x of the function $z = f(x, y)$, whose graph is represented here by the surface in grey scale. The partial derivative with respect to x with $y = y_0$ held constant is the slope of the tangent to the curve $z = f(x, y_0)$, which is the intersection of the plane $y = y_0$ and the surface $z = f(x, y)$. Adapted from Auletta and Wang (2014, p. 132)

$$\int udv = uv - \int vdu. \tag{1.30}$$

Setting $u = \partial L/\partial\dot{q}$, $dv = d(\delta q)/dt$ and using the expression (1.29), we get from the second term in Eq. (1.28):

$$\int_{t_1}^{t_2} \frac{\partial L}{\partial \dot{q}} \frac{d}{dt} \delta q dt = \frac{\partial L}{\partial \dot{q}} \delta q \Big|_{t_1}^{t_2} - \int_{t_1}^{t_2} \left(\frac{d}{dt} \frac{\partial L}{\partial \dot{q}} dt \right) \delta q. \tag{1.31}$$

Involving also the first term of Eq. (1.28), we get

$$\frac{\partial L}{\partial \dot{q}} \delta q \Big|_{t_1}^{t_2} + \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial q} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} \right) \delta q dt = 0, \tag{1.32}$$

The first term in Eq. (1.32) cancels out because we are varying the path and not the extremal points, so that $\delta q(t_1) = \delta q(t_2) = 0$. Since the integral of the second term is zero, also the integrand must be zero, and we derive

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0. \tag{1.33}$$

This is the *Euler–Lagrange equation*, after the names of the Swiss mathematician and physicist Leonhard Euler (1707–1783) and J.–L. Lagrange, and can be considered as the equation of motion. Taking into account diverse subsystems j, from this equation, and from the fact that the kinetic term of Eq. (1.25) does not depend on q, we can deduce:

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_j} = \frac{\partial L}{\partial q_j} = -\frac{\partial V}{\partial q_j} = \dot{p}_j, \tag{1.34}$$

and therefore

$$p_j = \frac{\partial L}{\partial \dot{q}_j} = \dot{q}_j m_j, \tag{1.35}$$

in agreement with the momentum's definition (1.4). From the last two equations, we immediately derive

$$m_j \ddot{q}_j = -\frac{\partial V}{\partial q_j}, \tag{1.36}$$

where $\ddot{q}$ is the second time derivative of q. This equation, when generalised to a force F, is Newton's second law, after the name of Isaac Newton,

$$ma = F. \tag{1.37}$$

Hamiltonian, Lagrangian and Action

Let us now consider the differential of the Lagrange function as a function of position and speed[49]

$$dL = \sum_j \left(\frac{\partial L}{\partial q_j} dq_j + \frac{\partial L}{\partial \dot{q}_j} d\dot{q}_j \right), \tag{1.38}$$

and take the total time derivative

$$\frac{dL}{dt} = \sum_j \left(\frac{\partial L}{\partial q_j} \dot{q}_j + \frac{\partial L}{\partial \dot{q}_j} \ddot{q}_j \right). \tag{1.39}$$

By taking advantage of Eq. (1.34) we get

$$\begin{aligned} \frac{dL}{dt} &= \sum_j \left(\dot{q}_j \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_j} + \frac{\partial L}{\partial \dot{q}_j} \ddot{q}_j \right) \\ &= \sum_j \frac{d}{dt} \left(\dot{q}_j \frac{\partial L}{\partial \dot{q}_j} \right), \end{aligned} \tag{1.40}$$

what allows us to write the following equation:

$$\frac{d}{dt} \left(\sum_j \dot{q}_j \frac{\partial L}{\partial \dot{q}_j} - L \right) = 0. \tag{1.41}$$

Since the term

$$\sum_j \dot{q}_j \frac{\partial L}{\partial \dot{q}_j} - L \tag{1.42}$$

is invariant during time translation, we take it as the definition of the energy of the system, the Hamiltonian, which, thanks to Eqs. (1.24) and (1.34)–(1.35), can be written as

$$\begin{aligned} H(p,q,t) &= \sum_j p_j \dot{q}_j - L = \sum_j p_j \dot{q}_j - \frac{1}{2} \sum_j p_j \dot{q}_j + V(q) \\ &= T(p) + V(q). \end{aligned} \tag{1.43}$$

In other words, it is the sum of the kinetic and potential energy, as expected from Eq. (1.12). Consider now that the kinetic energy (1.24) can be rewritten as $p\dot{q}/2$ while the potential term, thanks to Eqs. (1.36)–(1.37), must be $V(q) = -\dot{p}q$; this allows us to write the differential equality

[49] Landau and Lifshitz (1976, Sects. 40 and 43).

$$dH = -\sum_j \left(\dot{p}_j dq_j - \dot{q}_j dp_j\right) \tag{1.44}$$

from which we can derive *Hamilton's equations* (that are the equation of motion by making use of the Hamiltonian):

$$\frac{\partial H}{\partial p_j} = \dot{q}_j, \text{ and } \frac{\partial H}{\partial q_j} = \dot{p}_j. \tag{1.45}$$

Let us now write the total time derivative of the Hamiltonian (which is the sum of three components):

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} + \sum_j \frac{\partial H}{\partial q_j}\dot{q}_j + \sum_j \frac{\partial H}{\partial p_j}\dot{p}_j. \tag{1.46}$$

By making use of Eqs. (1.45) it is clear that the last two terms on the right disappear, so that we get

$$\frac{dH}{dt} = \frac{\partial H}{\partial t}, \tag{1.47}$$

what, in the particular case in which there is no dependence on time, reduces to

$$\frac{dH}{dt} = 0, \tag{1.48}$$

that can be considered an expression of the law of energy conservation. Since from Eq. (1.23) we have

$$\frac{dS}{dt} = L, \tag{1.49}$$

we also have—compare Eq. (1.34)—

$$\frac{\partial S}{\partial q_j} = p_j. \tag{1.50}$$

From the latter result the expression

$$L = \frac{\partial S}{\partial t} + \sum_j p_j \dot{q}_j \tag{1.51}$$

follows, which, thanks to Eq. (1.43), allows us to deduce

$$\frac{\partial S}{\partial t} = -H. \tag{1.52}$$

1.2 The Basic Principles of Quantum Mechanics

1.2.1 The Superposition Principle

Quantum Superposition

We can grasp the quantum superposition principle through a generalisation of the so-called double-slit or Young's experiment shown in Panel (b) of Fig. 1.3, Sect. 1.1.2, according to which there is not only the probability of passing through either slit (corresponding to the classical view of particles), but also that of interference (see also Fig. 1.4, Sect. 1.1.3). The superposition principle, likely formulated for the first time by Paul Dirac, establishes that if a quantum system can e.g. be in a state (or ket) $|\psi_1\rangle$, or in another state (or ket) $|\psi_2\rangle$, then it can be also in any linear combination of both. Mathematically, the principle can be formulated as follows[50]:

$$|\psi\rangle = c_1 |\psi_1\rangle + c_2 |\psi_2\rangle \,, \tag{1.53}$$

which tells us that an arbitrary state $|\psi\rangle$ can be understood as a sum of at least two components, here $|\psi_1\rangle$ and $|\psi_2\rangle$, where $c_1, c_2 \in \mathbb{C}$ are some numerical coefficients that express the amount of the contributions of the components $|\psi_1\rangle$ and $|\psi_2\rangle$, respectively, to the global state $|\psi\rangle$ (Fig. 1.9). Note, therefore, that the components $|\psi_1\rangle$ and $|\psi_2\rangle$ are themselves state vectors. Their superposition gives rise to another state vector ($|\psi\rangle$), so that they can be understood as kinds of reference axes (like x and y in the ordinary Cartesian plane). Thus, they turn out to be a *basis* in which we can expand or express this state vector (or even other ones).

By making use of explicit vectorial notation, we can write the two components as the following kets

$$|\psi_1\rangle := \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |\psi_2\rangle := \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{1.54a}$$

while their bras are given by

$$\langle\psi_1| := \begin{pmatrix} 1 & 0 \end{pmatrix}, \quad \langle\psi_2| := \begin{pmatrix} 0 & 1 \end{pmatrix}, \tag{1.54b}$$

so that the state $|\psi\rangle$ given by Eq. (1.53) can be rewritten in vectorial form as

$$\begin{aligned} |\psi\rangle &= c_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \end{aligned} \tag{1.55}$$

[50] I follow here Auletta et al. (2009, Chap. 1). See also Auletta and Wang (2014, Chap. 2).

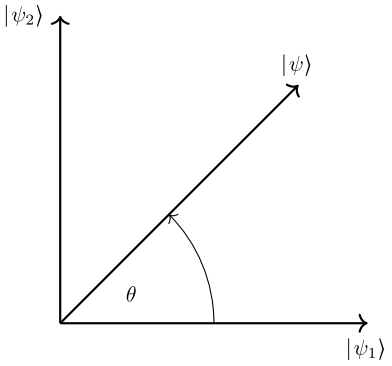


Fig. 1.9 The state $|\psi\rangle$ is represented as a vector and can be geometrically understood as the sum of its two components $|\psi_1\rangle$ and $|\psi_2\rangle$. If the former is represented on the horizontal axis, its coefficient c_1 gives the angle θ that with $|\psi\rangle$ (as well as the coefficient c_2 describes the complementary angle). Note that whenever we have $\theta = 45°$, we also have $c_1 = c_2 = 1/\sqrt{2}$ (which can be conceived as cosine and sine of θ)

and its bra as

$$\langle\psi| = \left(c_1^\star \; c_2^\star\right). \tag{1.56}$$

A Problem with *Omnimoda Determinatio*

All this, according to Dirac, entails "that the initial state must be as the result of a superposition of two or more other states, in a way that cannot be conceived on classical ideas".[51] In fact, in classical physics, we can have a superposition of, e.g. two different waves, while here we have a *single* system (e.g. a photon) in a single determined (not statistical) state $|\psi\rangle$ that is in superposition. Therefore, what we have is a superposition of a system with itself, something that is classically unknown. This is why Dirac introduced the concept of *self-interference*.[52] This shows that QM cannot satisfy the classical principle of *omnimoda determinatio*.[53] In fact, according to classical physics and philosophy any natural system is in a state such that all of the properties that could be assigned to it are either instantiated or are not, what means that the state of the system can be conceived as determined by all values of the physical quantities (energy, momentum, position, …) that describe it. Now, if

[51] See Dirac (1930, p. 9).

[52] Self-interference was experimentally verified for the first time in the 1960s (Pfleegor and Mandel 1967a, b). Successively, further confirmations have come from the experiments performed by different teams (Grangier et al. 1986; Franson and Potocki 1988).

[53] As e.g. formulated in Baumgarten (1739), Kant (1763). Peirce understood that modern philosophy is characterised by this assumption (Peirce 1903, p. 180) and is likely to be the first philosopher to have considered that individual beings are not perfectly determined (Peirce 1877). He also acknowledged that the *omnimoda determinatio* is a necessary condition of classical determinism (Peirce 1892, p. 116). On this see also Auletta (2004).

the state of the system shows self-interference we cannot say that, e.g. it occupies a specific position. Below, we shall deal with this problem extensively.

Born's Probabilistic Formulation

Let us note that in Schrödinger's initial understanding of the homonymous equation, a quantum state should represent a true wave: Schrödinger wanted to portray the electrons as pure waves moving in the ordinary space and to attribute the observed corpuscular properties to a little undulatory disturbance. But this interpretation was subject to big difficulties: the (quick) dispersion of the assumed wave packet (an objection that was also addressed to de Broglie),[54] the problem of localisation (to be dealt with below), and the fact that quantum "waves" are "located" in a Hilbert space and not in an ordinary position space or (Euclidean 3D) configuration space. In CM, a single particle has a configuration space of $\mathbb{R}^3$ (according to three spatial coordinates). Of course, for n particles, this configuration space is $\mathbb{R}^{3n}$. However, the state of a *single* quantum system often spans a Hilbert space with more than three dimensions; additionally, such a space needs to be complex. For this reason, the German physicist and mathematician Max Born, who, as seen, gave fundamental contributions to the building of the new theory together with W. Heisenberg, introduced the *probabilistic interpretation*, thanks to which quantum-mechanical processes were ruled by probabilities.[55] Born assumed that the wave in Schrödinger's equation was nothing but a kind of probability function: in particular, it would be associated with the probability of detect a particle in some state at a certain time as represented by one of the components of the superposition (e.g. either $|\psi_1\rangle$ or $|\psi_2\rangle$). This is evident when combining Eq. (1.53) with Eq. (1.10). Then, the coefficients c_1 and c_2 turn out to be *probability amplitudes*, such that their square moduli represent the probability to get the components $|\psi_1\rangle$ and $|\psi_2\rangle$ when measuring the system, respectively. In the case displayed in Fig. 1.9, we have equal probability ($|c_1|^2 = |c_2|^2 = \left|1/\sqrt{2}\right|^2 = 1/2$) to get one or the other of the two components of the superposition. Since the sum of the probabilities of all possible alternative events must be equal to 1 (because at least one of those events will occur so that the sum of the probabilities of all possible events express certainty), we have that also the sum of the square moduli of c_1 and c_2 must be equal to 1. This is called the *normalisation condition*.

In other words, the state vector, according to Born, did not stand for the properties of the physical world, neither for the possible vibrations of the universe's ultimate substrate, as Schrödinger (and, from a different perspective, de Broglie) had presumed, but was interpreted as representing probability amplitudes rather than amplitudes of physically real waves or real fields. As I have summarised elsewhere, the state vector would only determine the probability that a particle—with its energy and momentum—may choose a certain trajectory; but neither the energy nor

[54] Schlosshauer (2007, pp. 116–17).

[55] Born (1926). For historical reconstruction see Jammer (1966, Sect. 6.1).

momentum belong to the (probability) wave.[56] The motion of a particle was thus related to a law of probability, which was expressed by the Schrödinger's equation, and since there is no other way of defining the path (or the location) of a particle independently of the probabilistic distribution (or probability density) of the positions, QM becomes essentially a *probabilistic theory*. I stress in fact that quantum-mechanical probabilities do not depend on a partial knowledge, but rather they are something intrinsic to the quantum entities themselves. We shall see later how an interpretation that considers probabilities as depending on subjective ignorance of the actual properties of the system is untenable. Nevertheless, note that the Schrödinger equation is still *deterministic*, although of a very different kind relative to the equations of CM due to its probabilistic character. We shall come back on this problem showing how to integrate these two aspects.

Thus, from a formal point of view, we express in general terms the probability $\wp_1$ to get the component say $|\psi_1\rangle$ as the square modulus of the product between the ket $|\psi\rangle$ and the bra $\langle\psi_1|$:

$$\wp_1 = |\langle\psi_1 \mid \psi\rangle|^2 = \langle\psi_1 \mid \psi\rangle \langle\psi_1 \mid \psi\rangle^* , \tag{1.57}$$

and similarly for $\wp_2$. The expression $\langle\psi_1 \mid \psi\rangle$ is called the scalar product between (bra and ket), where the two central lines have been contracted to a single one for the sake of notation, and $\langle\psi_1 \mid \psi\rangle^* = \langle\psi \mid \psi_1\rangle$ is its complex conjugate. In other words, the two coefficients of the superposition (1.53) are given by

$$c_1 = \langle\psi_1 \mid \psi\rangle \text{ and } c_2 = \langle\psi_2 \mid \psi\rangle , \tag{1.58}$$

so that the state (1.55) can be rewritten as

$$\begin{aligned} |\psi\rangle &= \langle\psi_1 \mid \psi\rangle\, |\psi_1\rangle + \langle\psi_2 \mid \psi\rangle\, |\psi_2\rangle \\ &= \begin{pmatrix} \langle\psi_1 \mid \psi\rangle \\ \langle\psi_2 \mid \psi\rangle \end{pmatrix}, \end{aligned} \tag{1.59}$$

where i have used the vector formulation (1.54a). Note that all unitary transformations preserve the scalar product and therefore probability amplitudes, making reversibility possible.

Scalar Product

The *scalar product* (or inner product) of two state vectors is a kind of the dot product of two Euclidean vectors.[57] For two vectors $\mathbf{a}$ and $\mathbf{b}$, their dot product $\mathbf{a} \cdot \mathbf{b}$ is a real number defined by

$$\mathbf{a} \cdot \mathbf{b} := \|\mathbf{a}\|\, \|\mathbf{b}\|\, \cos\theta, \tag{1.60}$$

[56] Auletta (2000, pp. 104–105).

[57] I follow here Auletta and Wang (2014, Chap. 3).

where $\|\mathbf{a}\|$ and $\|\mathbf{b}\|$ are the magnitudes of the vectors $\mathbf{a}$ and $\mathbf{b}$, respectively, and θ is the angle between $\mathbf{a}$ and $\mathbf{b}$. The dot product $\mathbf{a} \cdot \mathbf{b}$ expresses how much the vectors $\mathbf{a}$ and $\mathbf{b}$ are aligned (or overlap). When two vectors are aligned (i.e. $\mathbf{a} = k\mathbf{b}$ with $k \neq 0$), their dot product is obviously different from zero. When the vectors are said to be orthogonal since they are pointing in perpendicular directions, their dot product is zero. The magnitude (length) or also (Euclidean) norm of a vector $\|\mathbf{a}\|$ can be expressed in terms of the dot product as

$$\|\mathbf{a}\| = \sqrt{\mathbf{a} \cdot \mathbf{a}}. \tag{1.61}$$

It is suitable to introduce here Cartesian coordinates. By denoting with $\mathbf{e}_x$, $\mathbf{e}_y$, and $\mathbf{e}_z$ the unit vectors (versors) in the x, y, and z directions, respectively, we can decompose any vector in their components and write the dot product of the vectors $\mathbf{a} = a_x\mathbf{e}_x + a_y\mathbf{e}_y + a_z\mathbf{e}_z$ and $\mathbf{b} = b_x\mathbf{e}_x + b_y\mathbf{e}_y + b_z\mathbf{e}_z$ as

$$\mathbf{a} \cdot \mathbf{b} = a_x b_x + a_y b_y + a_z b_z. \tag{1.62}$$

Coming back to QM, the scalar product of two arbitrary (2D) state vectors $|\,a\rangle$ and $|\,b\rangle$ is a complex number and is defined by

$$\langle a \mid b\rangle := \begin{pmatrix} a_1^* & a_2^* \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = a_1^* b_1 + a_2^* b_2. \tag{1.63}$$

The general rule for calculating the scalar product is to multiply the first and second columns in the row vector by the corresponding rows in the column vector and then add the results together. It is important to discriminate between Cartesian dimension of a system, which can be between 1 and 3 and give e.g. the direction of a physical quantity like momentum, and its Hilbert-space dimension that goes from 2 to infinity and represents the number of dimensions of one of the bases expanding the state of the system under consideration. Here, we are speaking of the latter case. In the following, I shall specify when an ambiguity may arise.

In addition, the scalar product satisfies the following properties:

- $\langle b \mid a\rangle = \langle a \mid b\rangle^*$, (1.64a)
- $\langle a \mid a\rangle \geq 0$ (equality holds iff $|\,a\rangle = 0$), (1.64b)
- $(\langle b| + \langle c|)\,|a\rangle = \langle b \mid a\rangle + \langle c \mid a\rangle$, (1.64c)
- $\langle a\,|\,(|b\rangle + |c\rangle) = \langle a \mid b\rangle + \langle a \mid c\rangle$, (1.64d)
- $\langle b\,|\,(\alpha\,|a\rangle) = \alpha\,\langle b \mid a\rangle$, (1.64e)
- $(\alpha\,\langle b|)\,|a\rangle = \alpha\,\langle b \mid a\rangle$, (1.64f)

where $\alpha \in \mathbb{C}$.

Orthonormal Bases

The previous formalism can be generalised to any n-dimensional state as follows.[58] In QM we only use orthonormal bases,[59] whose vectors are orthogonal to each other (in order to span the Hilbert space of the system) and have length =1. The reason for the latter requirement is that, if we assume a system to be in a given state, it has a probability =1 to be in that state, and therefore also its magnitude must be =1. Formally, a basis is *orthonormal* if

$$\forall j, \langle b_j \mid b_j \rangle = 1 \text{ and } \forall\, j \neq k, \langle b_j \mid b_k \rangle = 0. \tag{1.65a}$$

Synthetically,

$$\langle b_j \mid b_k \rangle = \delta_{jk}, \tag{1.65b}$$

where δ_{jk} is the Kronecker delta, after the name of the German mathematician Leopold Kronecker (1823–1891), defined by

$$\delta_{jk} = \begin{cases} 1 & \text{for } j = k, \\ 0 & \text{for } j \neq k. \end{cases} \tag{1.66}$$

Any quantum state $|\psi\rangle$ can be expanded thanks to an orthonormal basis $\{|b_j\rangle\}$ in a Hilbert space with a certain dimension n:

$$\begin{aligned} |\psi\rangle &= \sum_{j=1}^{n} c_j \, |b_j\rangle \\ &= \sum_{j=1}^{n} \langle b_j \mid \psi \rangle \, |b_j\rangle \ , \end{aligned} \tag{1.67}$$

where the $c_j = \langle b_j \mid \psi \rangle$ are (in general) complex numbers. Introducing another vector

$$|\psi'\rangle = \sum_{k=1}^{n} |b'_k\rangle \, , \tag{1.68}$$

the scalar product between them is

$$\langle \psi' \mid \psi \rangle = \sum_{j,k} c_j (c'_k)^* \langle b'_k \mid b_j \rangle . \tag{1.69}$$

[58] In the following, when there is no lack of generality, I shall introduce the formalism for the 2D Hilbert space. Sometimes, I shall explicitly show the generalisation to the nD case.

[59] Byron and Fuller (1969–70, I, Sect. 4.2).

The vectors constituting the basis can be written as a generalisation of Eqs. (1.54):

$$|b_1\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \dots \\ 0 \end{pmatrix}, \quad |b_2\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \dots \\ 0 \end{pmatrix}, \quad |b_3\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ \dots \\ 0 \end{pmatrix}, \quad \dots, \quad |b_n\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \dots \\ 1 \end{pmatrix}, \tag{1.70}$$

so that the ket $|\psi\rangle$ can be written as a generalisation of Eq. (1.55)

$$|\psi\rangle = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ \dots \\ c_n \end{pmatrix}, \tag{1.71}$$

and similarly for $\left|\psi'\right\rangle$. Note that we can meet state vectors that are not normalised. Suppose that such a vector is $|\psi\rangle$. To get its normalised form it suffices to take into account its norm given by (1.61), that we rewrite here as

$$\| \, |\psi\rangle \, \| = \sqrt{\langle\psi \mid \psi\rangle}. \tag{1.72}$$

Then, the normalised vector is given by

$$\left|\psi'\right\rangle = \frac{|\psi\rangle}{\| \, |\psi\rangle \, \|} \, . \tag{1.73}$$

1.2.2 The Quantisation Principle

Quantum Observables and Operations

Although Planck's solution had a rather specific scope (to "technically" solve the black body problem), its consequences were really revolutionary. As we have seen in Sect. 1.1.1, Planck did not reject the classical electromagnetic theory and so did not assume that the source light was discontinuous (as it was hypothesised by Einstein for explaining the photoelectric effect). So, Planck assumed that the discontinuity properties are only acquired in consequence of the emission of radiation out of the atomic and molecular structure of the internal surface of the black body that at that time was already known to be discontinuous. As said, Planck's assumption was confined to a particular behaviour of light in certain specific circumstances. However, there is also a general lesson, here: the idea that, in order to display discontinuous properties,

quantum systems need to interact with other systems.[60] This is the physical ground of the quantisation principle, which could be considered an extension or even the formal ground of the quantum postulate (Sect. 1.1.2).[61] This represents Heisenberg's great contribution (as the result of the mentioned collaboration of Born and Jordan) and the birth of matrix mechanics (the approach alternative to Schrödinger's wave mechanics) that is in fact also considered to be the official birth of QM as such.[62]

In fact, energy having a discrete spectrum, as suggested by experimental evidence in both cases of light and matter, as seen in the previous section, has important consequences in the general definition of physical quantities in QM. Note that these quantum-mechanical physical quantities, like energy, position, momentum, angular momentum, and so on, that allows us to describe (the state of) a system are called *observables*. Moreover, by *spectrum* of an observable it is understood the range of values of a physical quantity (the values that it can take in a certain state and would be the outcomes if the system be measured). The most important of the mentioned consequences is the following: if the observables defining a quantum system can display discontinuous properties as energy does, they cannot be represented by real variables and functions of real variables, as it occurs for CM. In other words, QM deals with a situation where physical quantities may have not only a continuous spectrum (as for classical systems), but also a discrete spectrum or even a combination of them. Mathematically, real variables are good for continuos spectra but are not a suitable tool for dealing with this problem. At the opposite, the spectrum of operators may have both a continuous and a discrete component. So, *operators* (in Sect. 1.1.3 I have introduced as first examples the energy and the unitary operators) are mathematical tools that allow to solve this problem. However, to introduce operators in a physical theory determines some further and not irrelevant consequences that we shall examine step by step.

Summarising, the quantisation principle can be formulated as follows: observables in QM are represented by operators on a Hilbert space, which we have defined as a linear vector space endowed with a scalar product and which is complete and separable. Hilbert spaces are the natural framework for state vectors in QM and can have a finite or infinite dimension (e.g. according to whether observables have discontinuous or continuous spectra).[63]

Matrices

Operators, for all what concerns us here, express possible operations on state vectors. For instance, the energy operator $\hat{H}$ of Eq. (1.10) determines the time evolution of $|\psi\rangle$ (a rotation in the Hilbert space, as we shall see). Finite-dimensional operators

[60] As pointed out in Auletta and Wang (2014, Sects. 4.1–4.3). This seems to have been also the central point of the so-called modal interpretation: see van Fraassen (1991), Dieks (1994).

[61] For what follows see also Jammer (1966, Chap. 5).

[62] Heisenberg (1925).

[63] For a technical account of this problem see Auletta et al. (2009, Sect. 1.3.2).

are represented by *matrices*, a mathematical generalisation of vectors: we have here in fact both columns and rows (they can be considered as the Cartesian product of a column vector and a row vector). Thus, they are represented by array of numbers. For instance, two arbitrary 2×2 matrices can be written as[64]

$$\hat{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad \text{and} \quad \hat{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}, \tag{1.74}$$

where the elements a's and b's are in general complex numbers. If $\alpha \in \mathbb{C}$, according to linearity, we have

$$\alpha \hat{A} = \begin{bmatrix} \alpha\, a_{11} & \alpha\, a_{12} \\ \alpha\, a_{21} & \alpha\, a_{22} \end{bmatrix} \tag{1.75a}$$

and

$$\hat{A} + \hat{B} = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \end{bmatrix}, \tag{1.75b}$$

Moreover, in the easiest case, matrix multiplication is an operation that takes a pair of square matrices (with the same numbers of rows and columns), and gives another square matrix as output. For instance, the product of the matrices (1.74), written as $\hat{A}\hat{B}$, is defined by

$$\hat{A}\hat{B} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix}. \tag{1.76}$$

Generalisation to higher dimensions is straightforward. If $\hat{A}$ is an $n \times m$ matrix and $\hat{B}$ is an $m \times p$ matrix, the product $\hat{A}\hat{B}$ of their multiplication is an $n \times p$ matrix but the product $\hat{B}\hat{A}$ is not defined. In other words, in order to perform the product we need that number of columns of the first matrix and the number of rows of the second matrix coincide. Matrix multiplication satisfies the following properties:

- $\hat{A}(\hat{B}\hat{C}) = (\hat{A}\hat{B})\hat{C}$, (1.77a)
- $\hat{A}(\hat{B} + \hat{C}) = \hat{A}\hat{B} + \hat{A}\hat{C}$, (1.77b)
- $(\hat{A} + \hat{B})\hat{C} = \hat{A}\hat{C} + \hat{B}\hat{C}$, (1.77c)
- $\alpha(\hat{A}\hat{B}) = (\alpha\hat{A})\hat{B} = \hat{A}(\alpha\hat{B})$, (1.77d)

where $\hat{A}$, $\hat{B}$, and $\hat{C}$ are matrices such that the above operations are defined and $\alpha \in \mathbb{C}$.

Note that the order of the matrices in a product does mathematically matter. Moreover, if matrices $\hat{A}$ and $\hat{B}$ represent some physical operations, the order in which we effect the product does matter also physically. The convention is that the matrix (or matrices) on the left of a matrix (or matrices) represent a subsequent

[64] See Byron and Fuller (1969–70, I, Sect. 3.7). I follow the short summary in Auletta and Wang (2014, Sect. 3.6).

operation relative to it (them) on its (their) right. The reason is that each previous operation needs to be in the scope of the next operation(s) if they are applied on the same state: thus, $\forall\, |\psi\rangle$

$$\hat{B}\hat{A}\,|\psi\rangle = \hat{B}\left(\hat{A}\,|\psi\rangle\right) = \hat{B}\left|\hat{A}\psi\right\rangle = \hat{B}\,|\psi'\rangle \tag{1.78}$$

means that we apply the subsequent operation $\hat{B}$ to some state $|\psi'\rangle = \left|\hat{A}\psi\right\rangle$ that is the initial state $|\psi\rangle$ already transformed by $\hat{A}$. Note that we can "absorb" an operator in a ket (and, as we shall see now, its adjoint in the relative bra).

Transposed Conjugate Matrix

I recall that, for a 2×2 matrix

$$\hat{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \tag{1.79}$$

its *adjoint* is a combination of complex conjugation and transposition. A transposed matrix is obtained from a matrix by a clockwise rotation of 90° of its rows as follows

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \longmapsto \begin{bmatrix} a & c \\ b & d \end{bmatrix}, \tag{1.80}$$

so that former rows become columns and vice versa. The symbol $\longmapsto$ means that the LHS maps to (or becomes, under a suitable operation) the RHS. When we perform the complex conjugation we finally get its adjoint:

$$\hat{A}^\dagger = \begin{bmatrix} a^* & c^* \\ b^* & d^* \end{bmatrix}. \tag{1.81}$$

The transposed conjugate satisfies the following properties:

- $(\hat{A} + \hat{B})^\dagger = \hat{A}^\dagger + \hat{B}^\dagger\ \forall$ matrices $\hat{A}$ and $\hat{B}$ of the same dimensions, (1.82a)
- $(\alpha\hat{A})^\dagger = \alpha^*\hat{A}^\dagger,\ \forall\alpha \in \mathbb{C}$ and $\forall$ matrix $\hat{A}$, (1.82b)
- Eigenvalues of $\hat{A}^\dagger$ are the complex conjugates of those of $\hat{A}$, (1.82c)
- $(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger\ \forall\ m \times n$ matrix $\hat{A}$ and $n \times p$ matrix $\hat{B}$, (1.82d)
- $(\hat{A}^\dagger)^\dagger = \hat{A}\ \forall$ matrix $\hat{A}$, (1.82e)
- If $\hat{A} = \hat{B}\hat{B}^\dagger$, we have that $\hat{A}^\dagger = (\hat{B}\hat{B}^\dagger)^\dagger = \hat{B}\hat{B}^\dagger = \hat{A}$. Similarly for $\hat{A} = \hat{B}^\dagger\hat{B}$. (1.82f)

Additionally, note that for all vectors $|\psi\rangle\ , |\varphi\rangle$ and operator $\hat{A}$, we have

$$\left\langle \hat{A}^{\dagger}\psi \mid \varphi \right\rangle = \left\langle \psi \mid \hat{A}\varphi \right\rangle, \quad \left\langle \hat{A}\psi \mid \varphi \right\rangle = \left\langle \psi \mid \hat{A}^{\dagger}\varphi \right\rangle. \tag{1.83}$$

In other words, a displacement of an operator from the left to the right (or vice versa) maps to its adjoint.

Projectors

As mentioned, a matrix can be understood as the Cartesian product of a column vector and a row vector. Confining us by now (and without loss of generality) to 2D Hilbert-spaces and considering the two state vectors given by the expressions (1.54), let us introduce the most elementary form of operators: *projectors*. The projectors $\hat{P}_{\psi_1}$ and $\hat{P}_{\psi_2}$ can be expressed in matrix form as[65]

$$\hat{P}_{\psi_1} = |\psi_1\rangle \langle\psi_1| = \begin{pmatrix} 1 \\ 0 \end{pmatrix} (1\ 0) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \tag{1.84a}$$

and

$$\hat{P}_{\psi_2} = |\psi_2\rangle \langle\psi_2| = \begin{pmatrix} 0 \\ 1 \end{pmatrix} (0\ 1) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}. \tag{1.84b}$$

Projectors are kinds of counterparts of scalar products when bra and ket coincide (they are the product of ket and a bra while the latter is a product of bra and a ket) and allow us to rewrite Eq. (1.59) as

$$\begin{aligned} |\psi\rangle &= |\psi_1\rangle \langle\psi_1 \mid \psi\rangle + |\psi_2\rangle \langle\psi_2 \mid \psi\rangle \\ &= \hat{P}_{\psi_1} |\psi\rangle + \hat{P}_{\psi_2} |\psi\rangle . \end{aligned} \tag{1.85}$$

As a consequence, $\forall j = 1, 2, \ldots$, we have

$$\hat{P}_{\psi_j} |\psi\rangle = c_j \left|\psi_j\right\rangle , \tag{1.86}$$

where I have made use of the properties of the scalar product:

$$\left\langle \psi_k \mid \psi_j \right\rangle = \delta_{jk}, \tag{1.87}$$

where δ_{jk} is the Kronecker delta (1.66). This clarifies that projectors are called in such a way because they select one of the components of the state in superposition. In particular, they project on their component (here $\left|\psi_j\right\rangle$), shortened by a factor given by the relative coefficient (here c_j), as displayed in Fig. 1.10.

Projectors have three properties: the first property is that they are positive semidefinite operators, which means that, $\forall |\psi\rangle$, we have

[65] I follow Auletta and Wang (2014, Sect. 3.7).

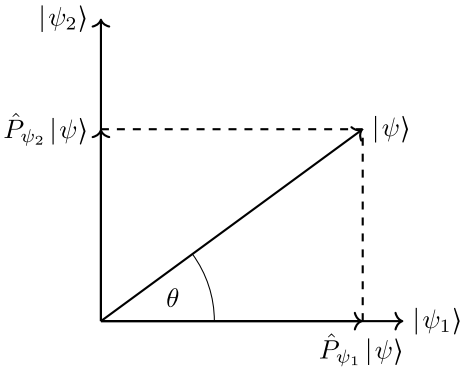


Fig. 1.10 Decomposition of an arbitrary vector $|\psi\rangle$. The pair $|\psi_1\rangle$ and $|\psi_2\rangle$ form an orthonormal basis. Note that the projection reduces the length of $|\psi_1\rangle$ and $|\psi_2\rangle$

$$\left\langle \psi \left| \hat{P}_{\psi_1} \right| \psi \right\rangle = \begin{pmatrix} c_1 & c_2 \end{pmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \geq 0, \tag{1.88a}$$

$$\left\langle \psi \left| \hat{P}_{\psi_2} \right| \psi \right\rangle = \begin{pmatrix} c_1 & c_2 \end{pmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \geq 0. \tag{1.88b}$$

The second property of the projectors is that the square of a projector is equal to the projector itself (this is called *idempotency*,) while the product of projectors onto orthogonal states vanishes:

$$\hat{P}_{\psi_1}^2 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \hat{P}_{\psi_1}, \tag{1.89a}$$

$$\hat{P}_{\psi_2}^2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \hat{P}_{\psi_2}, \tag{1.89b}$$

$$\hat{P}_{\psi_1} \hat{P}_{\psi_2} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = 0, \tag{1.89c}$$

$$\hat{P}_{\psi_2} \hat{P}_{\psi_1} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = 0, \tag{1.89d}$$

which can be expressed in a compact (and general) form as

$$\hat{P}_j \hat{P}_k = \delta_{jk} \hat{P}_k, \tag{1.90}$$

where $j, k = \psi_1, \psi_2$. The reason for the last two expressions in Eqs. (1.89) is that projectors onto orthogonal states represent mutually exclusive measurement outcomes. Mathematically, this is a consequence of the orthogonality of the corresponding state vectors. In general, a set of projectors satisfying Property (1.90) is said to be orthogonal.

Identity Operator

Since projectors represent mutually exclusive outcomes of a measurement, these outcomes occur with certain probabilities, and the sum of all probabilities of a set of mutually exclusive outcomes must be equal to 1, according to the normalisation condition. Consequently, all possible mutually exclusive outcomes of a measurement constitute a complete set of orthogonal states, and the third property of the projectors states that the sum of all projectors onto a complete set of orthogonal states is equal to the identity operator (which is the operatorial counterpart of the number 1 and was already met in Eq. (1.14)). In our 2D case, since the basis $\{|\psi_1\rangle\ ,\ |\psi_2\rangle\}$ is a complete orthonormal basis, we have

$$\hat{P}_{\psi_1} + \hat{P}_{\psi_2} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \hat{I}, \tag{1.91}$$

or equivalently,

$$\sum_{j=\psi_1,\psi_2} \hat{P}_j = \hat{I}, \tag{1.92}$$

where the identity operator $\hat{I}$ is defined as

$$\hat{I} := \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{1.93}$$

It is then clear that, for any Hilbert-space dimension n, the identity operator is represented by the identity matrix with all diagonal elements equal to 1 and all off-diagonal elements equal to 0, so that we have

$$\sum_{j=1}^{n} \hat{P}_j = \hat{I}. \tag{1.94}$$

The reason why it is called identity matrix is that, as already mentioned, it represents the operation that induces no change for all operators and states, i.e. we have

$$\hat{I}\hat{O} = \hat{O}\hat{I} = \hat{O} \quad \text{and} \quad \hat{I}\,|\psi\rangle = |\psi\rangle\ , \tag{1.95}$$

where $\hat{O}$ represents an arbitrary operator and $|\psi\rangle$ an arbitrary state. For this reason, the identity operator acting on an arbitrary state of system $\mathcal{S}$ is called *deterministic* since it gives with certainty (probability equal to 1) the input back. Equation (1.92) or (1.94) is also called the *completeness relation* or the *resolution of the identity*.

Spectral Decomposition

Let us now come back to the issue of the quantum-mechanical observables.[66] Since projectors express exclusive measurement outcomes, observables can be understood as sums of projectors taken together with the possible associated *eigenvalues*, i.e. the value that we would get when measuring that observables and obtaining the state described by the associated projector as outcome. In other words, an arbitrary (Hilbert-space) n-dimensional observable $\hat{O}$ can be written as

$$\hat{O} = o_1\hat{P}_1 + o_2\hat{P}_2 + \cdots + o_n\hat{P}_n = \sum_{j=1}^{n} o_j\hat{P}_j, \tag{1.96}$$

where o_j (with $1 \leq j \leq n$) is the eigenvalue associated with the projector $\hat{P}_j$. This is called the (discrete) *spectral decomposition* of the observable $\hat{O}$. Its continuous counterpart is given by the integral

$$\hat{O} = \int o\hat{P}(o)do, \tag{1.97}$$

where the quantity do represents infinitesimal increments or differences of the integration variable o. The fact that these decompositions exist for any observable $\hat{O}$ is called the *spectral theorem*.[67] In such a continuous case the completeness relation (1.94) takes the form

$$\int_{-\infty}^{+\infty} do\hat{P}(o) = \hat{I}. \tag{1.98}$$

Hermitian, Bounded and Symmetric Operators

Even though all the observables in QM should be represented by operators on a Hilbert space, not all mathematical operators on that same space, in fact, correspond to observables. Since the observables' values stand for physical quantities, they must be real numbers (we need to assume that any of such values can be, e.g. a number on a graduate scale of a measuring apparatus). Then, the operators associated to observables can only be Hermitian operators, after the name of the French mathematician Charles Hermite (1822–1901). A *Hermitian operator*, technically speaking, is that operator which coincides with its transposed conjugate or adjoint, i.e. $\hat{O} = \hat{O}^\dagger$ (see also Property (1.82f)), and is therefore also called self-adjoint operator (although,

[66] For the pure mathematical aspects see Byron and Fuller (1969–70, I, Sect. 3.10). For a technical account see Auletta et al. (2009, Chap. 2). For a simplified summary see Auletta and Wang (2014, Sect. 4.4).

[67] Prugovečki (1971, pp. 250–51). This is a very technical book but important for understanding the mathematics of QM.

as we shall, see not all self-adjoint operators are observables). In fact, a Hermitian operator is characterised by the fact that all its eigenvalues are real.[68]

If the dimensions are infinite (also in the discrete case), we can meet problems with the definition above if the operator is not bounded. A *bounded* operator is defined by the constraint

$$\frac{\left\langle \psi \left| \hat{O}^\dagger \hat{O} \right| \psi \right\rangle}{\langle \psi \mid \psi \rangle} \leq C^2, \tag{1.99}$$

where C is a real constant. The lowest possible value of it is the norm of $\hat{O}$. However, for our purposes, it suffices that, $\forall \, |\varphi\rangle \, , |\psi\rangle \in \mathcal{H}$, the operator is symmetric, i.e.

$$\left\langle \varphi \mid \hat{O}\psi \right\rangle = \left\langle \hat{O}\varphi \mid \psi \right\rangle, \tag{1.100}$$

and bounded from below.

Mean Value

If $\hat{O}$ is an observable with eigenvalues o_j's (which we take here as discrete for the sake of simplicity) the corresponding eigenstates will be denoted by $\left| o_j \right\rangle$'s (where j is some index labelling the eigenvalues). If the observable $\hat{O}$ is measured on a system in a state $|\psi\rangle$, and the measured result is one of its eigenvalues o_j (so that $\hat{P}_j = \left| o_j \right\rangle \left\langle o_j \right|$), the probability of measuring a given eigenvalue o_j is given by the generalisation of Eq. (1.57) to any nD Hilbert space:

$$\wp_j = \left\langle o_j \mid \psi \right\rangle \left\langle \psi \mid o_j \right\rangle = |\langle o_j | \psi \rangle|^2. \tag{1.101}$$

Thus, the *expectation value* of the observable $\hat{O}$ for a quantum system in the state $|\psi\rangle$, which is denoted by $\left\langle \hat{O} \right\rangle_\psi$, is given by the statistical mean

$$\begin{aligned} \left\langle \hat{O} \right\rangle_\psi &:= \langle \psi | \hat{O} | \psi \rangle \\ &= \sum_j o_j \left\langle \psi \mid o_j \right\rangle \left\langle o_j \mid \psi \right\rangle \\ &= \sum_j o_j |\langle o_j | \psi \rangle|^2, \end{aligned} \tag{1.102}$$

where use has been made of the spectral representation of the observable $\hat{O}$ given by Eq. (1.96).

[68] Byron and Fuller (1969–70, I, Sect. 4.4).

Superpositions and Change of Basis

As we have remarked, quantum systems can be in superposition. However, the introduction of operators for representing observables shows that superposition is a *relative* concept. In fact, a state of a quantum system can be considered to be a superposition according to a certain observable that we like to measure but be not in superposition according to another observable.[69] This means that the same state can be expanded in different bases (according to the different observables) and there are rules for passing from one basis to another. For instance, let us consider a photon that is in a superposition state of two components: $|h\rangle$ denoting horizontal polarisation (the direction of oscillation of the electromagnetic field, as shown in Fig. 1.2, Sect. 1.1.1) and $|v\rangle$ denoting vertical polarisation:

$$|\psi\rangle = \langle h|\psi\rangle\,|h\rangle + \langle v|\psi\rangle\,|v\rangle\,, \tag{1.103}$$

which can be understood as a variation of Eq. (1.59). However, note that the Hilbert-space geometry of state vectors is different relative to the Cartesian-space representation of the corresponding polarisation vectors. This is evident in the case of circular polarisation (at each point, the electric field of the wave has a constant magnitude but its direction rotates with time at a steady rate in a plane perpendicular to the direction of the wave). Of course, there is a right (CW) and left (CCW) circular polarisation, and the corresponding state vectors are

$$|r\rangle = \frac{1}{\sqrt{2}}\left(|h\rangle + i\,|v\rangle\right) \text{ and } |l\rangle = \frac{1}{\sqrt{2}}\left(|h\rangle - i\,|v\rangle\right), \tag{1.104}$$

respectively, but whose Cartesian representation is shown in Fig. 1.11.

Now, suppose two other polarisation states $|a\rangle$ and $|a'\rangle$ that are themselves superpositions of $|h\rangle$ and $|v\rangle$:

$$|a\rangle = \frac{1}{\sqrt{2}}\left(|h\rangle + |v\rangle\right) \quad \text{and} \quad |a'\rangle = \frac{1}{\sqrt{2}}\left(|v\rangle - |h\rangle\right), \tag{1.105}$$

as is evident by an inspection of Fig. 1.12. In other words, $|a\rangle$, $|a'\rangle$ represent polarisations at 45° and 135°, respectively. What is interesting is that we can write the state vector (1.103) in the new basis:

$$\begin{aligned} |\psi\rangle &= \langle a|\psi\rangle\,|a\rangle + \langle a'|\psi\rangle\,|a'\rangle \\ &= c_a\,|a\rangle + c_{a'}\,|a'\rangle\,. \end{aligned} \tag{1.106}$$

Now, the question is: how do we pass from one basis to another? Again, by means of unitary transformations. This is called *change of basis*.

[69] Here I follow Auletta and Wang (2014, Sects. 4.5–4.6).

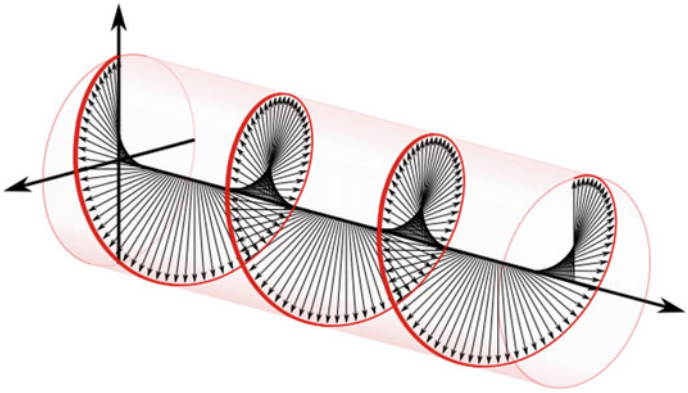


Fig. 1.11 Right-handed circular polarisation. Adapted from https://en.wikipedia.org/wiki/Circular_polarization

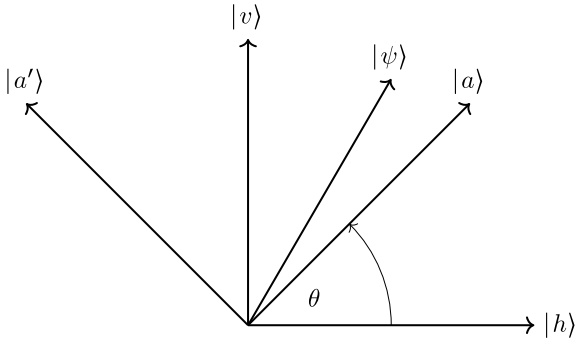


Fig. 1.12 Change of basis. It can be represented as a rotation of the basis. For instance, the basis $\{|a\rangle\ ,\ |a'\rangle\}$ is obtained from the original basis $\{|h\rangle\ ,\ |v\rangle\}$ by a counterclockwise rotation of 45°. Clearly, $|\psi\rangle$ can be expressed in both bases

I shall consider here the change of basis from the basis $\{|a\rangle\ ,\ |a'\rangle\}$ to the basis $\{|h\rangle\ ,\ |v\rangle\}$. The reverse can be obtained in a similar manner. Recalling the completeness relation of the kind (1.92) for the projectors $\hat{P}_a$ and $\hat{P}_{a'}$ as well as the second of Eq. (1.95), we have

$$\begin{aligned}\langle h\,|\,\psi\rangle &= \langle h|\hat{I}|\psi\rangle \\ &= \langle h\,|\,\big(|a\rangle\,\langle a\,|+|a'\rangle\langle a'\,|\big)\,|\,\psi\rangle\end{aligned} \tag{1.107a}$$

and

$$\begin{aligned}\langle v\,|\,\psi\rangle &= \langle v|\hat{I}|\psi\rangle \\ &= \langle v\,|\,\big(|a\rangle\,\langle a\,|+|a'\rangle\langle a'\,|\big)\,|\,\psi\rangle\ ,\end{aligned} \tag{1.107b}$$

which in turn allows us to formulate the relation between the coefficients $\{c_a, c_{a'}\}$ and $\{c_h, c_v\}$ as

$$\begin{aligned} c_h &= \langle h \mid \psi \rangle \\ &= \langle h \mid a \rangle \langle a \mid \psi \rangle + \langle h \mid a' \rangle \langle a' \mid \psi \rangle \\ &= \langle h \mid a \rangle c_a + \langle h \mid a' \rangle c_{a'} \end{aligned} \tag{1.108a}$$

and

$$\begin{aligned} c_v &= \langle v \mid \psi \rangle \\ &= \langle v \mid a \rangle \langle a \mid \psi \rangle + \langle v \mid a' \rangle \langle a' \mid \psi \rangle \\ &= \langle v \mid a \rangle c_a + \langle v \mid a' \rangle c_{a'}, \end{aligned} \tag{1.108b}$$

where the coefficients c_a, $c_{a'}$ are those present in Eq. (1.106). In matrix notation, the previous expressions can be put into a compact form as

$$\begin{pmatrix} c_h \\ c_v \end{pmatrix} = \begin{bmatrix} \langle h \mid a \rangle & \langle h \mid a' \rangle \\ \langle v \mid a \rangle & \langle v \mid a' \rangle \end{bmatrix} \begin{pmatrix} c_a \\ c_{a'} \end{pmatrix}, \tag{1.109a}$$

or equivalently,

$$\begin{pmatrix} \langle h \mid \psi \rangle \\ \langle v \mid \psi \rangle \end{pmatrix} = \begin{bmatrix} \langle h \mid a \rangle & \langle h \mid a' \rangle \\ \langle v \mid a \rangle & \langle v \mid a' \rangle \end{bmatrix} \begin{pmatrix} \langle a \mid \psi \rangle \\ \langle a' \mid \psi \rangle \end{pmatrix}. \tag{1.109b}$$

Let us denote the transformation matrix in Eqs. (1.109a) and (1.109b) by the change-of-basis unitary matrix $\hat{U}_B$, i.e.

$$\hat{U}_B = \begin{bmatrix} \langle h \mid a \rangle & \langle h \mid a' \rangle \\ \langle v \mid a \rangle & \langle v \mid a' \rangle \end{bmatrix}. \tag{1.110}$$

In fact, according to Eq. (1.14), we have $\hat{U}_B \hat{U}_B^\dagger = \hat{U}_B^\dagger \hat{U}_B = \hat{I}$, where, using the formulae (1.64a) and (1.81), we have the inverse transformation

$$\hat{U}_B^\dagger = \begin{bmatrix} \langle h \mid a \rangle^* & \langle v \mid a \rangle^* \\ \langle h \mid a' \rangle^* & \langle v \mid a' \rangle^* \end{bmatrix} = \begin{bmatrix} \langle a \mid h \rangle & \langle a \mid v \rangle \\ \langle a' \mid h \rangle & \langle a' \mid v \rangle \end{bmatrix}. \tag{1.111}$$

By expanding the basis vectors $|a\rangle$ and $|a'\rangle$ in the basis $\{|h\rangle, |v\rangle\}$ (using always the completeness condition for projectors) as

$$|a\rangle = (|h\rangle\langle h| + |v\rangle\langle v|)|a\rangle = \langle h \mid a \rangle |h\rangle + \langle v \mid a \rangle |v\rangle, \tag{1.112a}$$

$$|a'\rangle = (|h\rangle\langle h| + |v\rangle\langle v|)|a'\rangle = \langle h \mid a' \rangle |h\rangle + \langle v \mid a' \rangle |v\rangle, \tag{1.112b}$$

we find that the four matrix elements of $\hat{U}_B$ are just the expansion coefficients of the basis vectors $|a\rangle$ and $|a'\rangle$ in the basis $\{|v\rangle, |h\rangle\}$. Note that the basis used for expansion (here $\{|v\rangle, |h\rangle\}$) appear always as bras in the expansion coefficients.[70]

A generalisation can be provided as follows. Suppose the following expansion to arbitrary nD spaces (possibly infinite but discrete):

$$\begin{aligned} |\psi\rangle &= \sum_{j=1}^{n} c_j \, |o_j\rangle \\ &= \sum_{j=1}^{n} \langle o_j \mid \psi\rangle \, |o_j\rangle \ , \end{aligned} \tag{1.113}$$

where the $c_j = \langle o_j \mid \psi\rangle \in \mathbb{C}$ are again probability amplitudes. Recalling Eq. (1.67), we can pass from basis $|o_j\rangle$ to another basis $\{|b_j\rangle\}$ thanks to the following transformations:

$$\begin{aligned} \langle o_1 \mid \psi\rangle &= (\langle o_1 \mid b_1\rangle \, \langle b_1| + \langle o_1 \mid b_2\rangle \, \langle b_2| + \cdots + \langle o_1 \mid b_n\rangle \, \langle b_n|) \, |\psi\rangle \ , \\ \langle o_2 \mid \psi\rangle &= (\langle o_2 \mid b_1\rangle \, \langle b_1| + \langle o_2 \mid b_2\rangle \, \langle b_2| + \cdots + \langle o_2 \mid b_n\rangle \, \langle b_n|) \, |\psi\rangle \ , \\ &\cdots = \cdots \\ \langle o_n \mid \psi\rangle &= (\langle o_n \mid b_1\rangle \, \langle b_1| + \langle o_n \mid b_2\rangle \, \langle b_2| + \cdots + \langle o_n \mid b_n\rangle \, \langle b_n|) \, |\psi\rangle \ . \end{aligned} \tag{1.114}$$

These transformations can be written in a compact form as:

$$\begin{pmatrix} \langle o_1 \mid \psi\rangle \\ \langle o_2 \mid \psi\rangle \\ \cdots \\ \langle o_n \mid \psi\rangle \end{pmatrix} = \begin{bmatrix} \langle o_1 \mid b_1\rangle & \langle o_1 \mid b_2\rangle & \ldots & \langle o_1 \mid b_n\rangle \\ \langle o_2 \mid b_1\rangle & \langle o_2 \mid b_2\rangle & \ldots & \langle o_2 \mid b_n\rangle \\ \cdots & \cdots & \cdots & \cdots \\ \langle o_n \mid b_1\rangle & \langle o_n \mid b_2\rangle & \ldots & \langle o_n \mid b_n\rangle \end{bmatrix} \begin{pmatrix} \langle b_1 \mid \psi\rangle \\ \langle b_2 \mid \psi\rangle \\ \cdots \\ \langle b_n \mid \psi\rangle \end{pmatrix}, \tag{1.115}$$

where the matrix

$$\hat{U}_B = \begin{bmatrix} \langle o_1 \mid b_1\rangle & \langle o_1 \mid b_2\rangle & \ldots & \langle o_1 \mid b_n\rangle \\ \langle o_2 \mid b_1\rangle & \langle o_2 \mid b_2\rangle & \ldots & \langle o_2 \mid b_n\rangle \\ \cdots & \cdots & \cdots & \cdots \\ \langle o_n \mid b_1\rangle & \langle o_n \mid b_2\rangle & \ldots & \langle o_n \mid b_n\rangle \end{bmatrix} \tag{1.116}$$

represents again a *unitary* operator. Note that if we write the state vector $|\psi\rangle$ as a dependence on the two different bases, we have that the transformation

$$\hat{U}_B \, |\psi(\{o_k\})\rangle = |\psi(\{b_j\})\rangle \tag{1.117a}$$

is equivalent to

[70] By using unitary operators of this kind we can always diagonalize a matrix (Byron and Fuller 1969–70, I, Sect. 4.7).

$$
\begin{aligned}
|\psi(\{o_k\})\rangle &= \sum_j \hat{P}_{b_j} |\psi(\{o_k\})\rangle = \sum_j |b_j\rangle\langle b_j| \sum_k \langle o_k \mid \psi\rangle |o_k\rangle \\
&= \sum_j \sum_k |o_k\rangle \langle o_k| \langle b_j \mid \psi\rangle |b_j\rangle \\
&= \sum_k \hat{P}_{o_k} |\psi(\{b_j\})\rangle \\
&= |\psi(\{b_j\})\rangle , \qquad (1.117b)
\end{aligned}
$$

where I have used the completeness condition two times and the forelast step is allowed because the k's do not depend on the j's. In general, there is full freedom in displacing kets and bras in a chain, provided that we observe such dependencies and the involved operators (here projectors) span the same Hilbert space.

Eigenvalue Equation

Since superposition is relative concept, it can happen that the state $|\psi\rangle$ given by Eq. (1.113) is a superposition relatively to an arbitrary observable $\hat{O}$. Nevertheless, it can also happen that the components $|o_j\rangle$'s of that superposition are eigenstates of that observable. We call indeed *eigenstates* one of those state vectors that is represented by a ket in a spectral expansion like (1.96). They are the states that would be the experimental outcome if we in fact decided to measure $\hat{O}$. In such a situation, we have what is called the *eigenvalue equation*:

$$
\hat{O} |o_j\rangle = o_j |oj\rangle , \qquad (1.118)
$$

that tells us that the observable $\hat{O}$ acts on any of its j's eigenstates as a simple numerical multiplier, and each of those numbers $o_j \in \mathbb{R}$ are one of the eigenvalues of $\hat{O}$ (i.e. the value that we experimentally get when the outcome is, in fact $|o_j\rangle$).

Note that projectors (being operators and even observables) also have their eigenvalue equations. For instance, projector (1.84a) acts on orthogonal states $|\psi_1\rangle$, $|\psi_2\rangle$ as

$$
\hat{P}_{\psi_1} |\psi_1\rangle = |\psi_1\rangle , \quad \hat{P}_{\psi_1} |\psi_2\rangle = 0, \qquad (1.119)
$$

that is, they have values 1 or 0 according to whether they project on their component or on the orthogonal state vector.

1.2.3 Physical Observables and Different Representations

Position Observable

Up to now, I have introduced generic observables. However, in order to build a physical theory, we need to deal with specific ones that can describe the system at hand. In particular, we are interested in seeing how the previous two principles apply to more concrete physical situations. Let us first consider position and momentum observables (which, as in CM, are canonically conjugate). In the case of (a 1D) free particle (i.e. a particle moving on a line not subjected to fields or forces) we need to substitute the sum (over a determinate orthonormal basis: see e.g. Eq. (1.113)) with its continuous counterpart (see the beginning of Sect. 1.2.2). When dealing with integrals, we need often to define the interval of integration. Here, the interval of integration is, in fact, the whole interval of real numbers from minus infinity to plus infinity. In other words, a 1D free particle can be everywhere on an infinite line[71]:

$$|\psi\rangle = \int_{-\infty}^{+\infty} |x\rangle \langle x \mid \psi\rangle \, dx, \tag{1.120}$$

where $|x\rangle$ is the position eigenstate, and I have used the following counterpart of the completeness relation (1.98) for the (continuous) position eigenstates

$$\int_{-\infty}^{+\infty} |x\rangle \langle x| \, dx = \hat{I}. \tag{1.121}$$

Note that we have here a system that is 1D in Cartesian coordinates but whose states span a infinite-dimensional Hilbert space. Thus, the spectral decomposition of the 1D *position operator* $\hat{x}$ is given by an instance of Eq. (1.97):

$$\hat{x} = \int_{-\infty}^{+\infty} x \, |x\rangle \langle x| \, dx, \tag{1.122}$$

where $x \in \mathbb{R}$ is the eigenvalue corresponding to the ket $|x\rangle$. Note that, $\forall |x\rangle$ in Eq. (1.120), the scalar product $\langle x \mid \psi\rangle \in \mathbb{C}$ is the continuous expansion coefficient of the state vector $|\psi\rangle$ in the position operator eigenbasis $\{|x\rangle\}$, and hence, for a given state vector $|\psi\rangle$, can be understood as a continuous function of the position operator eigenvalue x, and therefore it pertains to the space $L^2(x)$ of linear functionals over that operator[72]: a linear functional on the Hilbert space $\mathcal{H}$ maps vectors to complex numbers and satisfies the linearity conditions (1.13). Thus, it is a linear map from a vector space to its field of scalars. Indeed, the probability amplitude $\langle x \mid \psi\rangle$ is known as the *wave function in position space*. Here, by position space I mean, as usually, the space in which ordinary objects are e.g. located at our physical scale. In other

[71] Here I follow Auletta and Wang (2014, Sect. 6.3).

[72] Byron and Fuller (1969–70, I, Sect. 5.1).

words, $\langle x \mid \psi \rangle$ describes the state of a quantum system as a function of the position and therefore can be written in functional form as

$$\psi(x) := \langle x \mid \psi \rangle \,. \tag{1.123}$$

In terms of the wave function $\psi(x)$, Eq. (1.120) can be rewritten as

$$|\,\psi\rangle = \int_{-\infty}^{+\infty} \psi(x)\,|\,x\rangle\; dx. \tag{1.124}$$

Note that, unlike the discrete case, the square modulus $|\,\langle x \mid \psi \rangle\,|^2$ of the probability amplitude $\langle x \mid \psi \rangle$ does *not* express the probability of finding a quantum system *at* the position x: in the continuum limit such a probability is identically zero because it requires infinite precision to exactly identify a real number (we encounter here again a limitation to the assumption of error-free measurement by CM as discussed in Sect. 1.1.2). Thus, we can only speak of the probability of finding a quantum system *in proximity of* the position x, or more precisely between x and $x + dx$, where I recall that dx, the differential of x, denotes a vanishing quantity. Such a probability is given by the square modulus of the probability amplitude multiplied by this small quantity, i.e.

$$\wp\,(x) = |\psi(x)|^2 dx, \tag{1.125}$$

while the square modulus $|\psi(x)|^2$ of the wave function gives what is called the *probability density* (describing the relative likelihood for a continuous random variable to take on a given value) instead of a true probability. Note that here (and for similar expressions) we must have $dx > 0$ since the probability is always non-negative. This allows us to write the continuous counterpart of the expectation value (1.102)

$$\left\langle \hat{x} \right\rangle_{\psi} = \int x\,|\psi(x)|^2\,dx. \tag{1.126}$$

Position Eigenfunctions

The action of the position operator on one of its eigenstates gives the eigenvalue equation for this observable as a particular case of Eq. (1.118):

$$\hat{x}\,|\,x\rangle = x\,|\,x\rangle \;, \tag{1.127a}$$

always with $x \in \mathbb{R}$; the bra version of the same equation is

$$\langle x\,|\,\hat{x} = x\;\langle x\,|\,, \tag{1.127b}$$

I recall that an eigenstate is an opposite concept relative to superposition: in fact, a superposition relative to $\hat{x}$ is a weighted sum (or integral) of eigenstates of $\hat{x}$, as it is evident by looking at Eq. (1.124). The mathematical framework in which the position observable acts as a numerical multiplier according to the previous equation is called *position representation*. In terms of wave function we have a similar eigenvalue equation:

$$\hat{x}\psi(x) = x\psi(x). \tag{1.127c}$$

Note that the scalar product $\langle x \mid x_0 \rangle$ for a fixed eigenvalue x_0 is called the *position eigenfunction*, that is, the particular position eigenstate $|x_0\rangle$ expressed in the position representation. In other words, it is the functional (continuous) counterpart of an eigenstate. Writing the position eigenfunction as

$$\varphi_{x_0}(x) := \langle x \mid x_0 \rangle \in \mathbb{C}, \tag{1.128}$$

from the orthonormal conditions (1.65a) we get

$$\varphi_{x_0}(x) = \delta(x - x_0), \tag{1.129}$$

where $\delta(x - x_0)$ represents a Dirac delta function (although it is rather a distribution),[73] after the name of P. Dirac, which, being the continuous counterpart of the Kronecker delta (1.66), in a continuum selects a specific point (x_0):

$$\forall x \neq x_0,\ \delta(x - x_0) = 0, \text{ but } \forall x = x_0,\ \delta(x - x_0) = \infty. \tag{1.130}$$

We can conceive a delta function as the limit of squeezing more and more a probability distribution around a value (here, x_0), which is typically a bell-shaped curve (a Gaussian), up to the point that it becomes a (vertical) straight line passing for that value. For the continuous case, we expect that the eigenstates of $\hat{x}$ be indeed delta functions. The corresponding eigenvalue equation for $\varphi_{x_0}(x)$ is given by

$$\hat{x}\varphi_{x_0}(x) = x_0\varphi_{x_0}(x). \tag{1.131}$$

Generalisation to the Cartesian 3D case is provided below.

Spatial Translations

The momentum of a system describes its spatial translations (changes of position), and therefore can be thought of as being strictly related to the generator of space translations, that is, as the dynamical parameter or operator that is associated with changes in the kinematic parameter represented by the position. To be concrete, let

[73] Byron and Fuller (1969–70, I, Sect. 5.3).

us consider an *infinitesimal* space translation in the x direction given by[74]

$$x \xrightarrow{\varepsilon} x'_{\varepsilon} = x + \varepsilon, \tag{1.132}$$

where ε is an infinitesimal distance. The translated position x' can be expressed in terms of position x as

$$x'_{\varepsilon} = \left(1 + \varepsilon \hat{D}_x\right) x, \tag{1.133}$$

where $\hat{D}_x = \frac{d}{dx}$, the derivative operator with respect to x, is the generator of space translations in the x direction. To pass from an infinitesimal to a *finite* space translation a in the x direction

$$x \xrightarrow{a} x' = x + a, \tag{1.134}$$

we may think of the latter as a result of an infinite number n of repeated applications of the infinitesimal translation (1.133). This is achieved by taking the limit for n approaching infinity. Since the effect of each infinitesimal translation is given by Eq. (1.133) and all of them together constitute the finite displacement a, we can multiply these infinitesimal displacement operations and express x' in terms of x as

$$\begin{aligned} x' &= \lim_{n\to\infty} \overbrace{\left(1 + \frac{a}{n}\hat{D}_x\right)\cdots\left(1 + \frac{a}{n}\hat{D}_x\right)}^{\text{total of } n \text{ factors}} x \\ &= \lim_{n\to\infty} \left(1 + \frac{a}{n}\hat{D}_x\right)^n x \\ &= e^{a\hat{D}_x} x, \end{aligned} \tag{1.135}$$

where in the last step use has been made of the mathematical formula

$$\lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n = e^x. \tag{1.136}$$

This formula can be derived from the Taylor expansion (1.18). Indeed, for fixed x, the ratio $\frac{x}{n}$ becomes exceedingly small as n approaches infinity, and, given the approximation (1.19), Eq. (1.136) follows since we have $\left(e^{\frac{x}{n}}\right)^n = e^x$. The expression $e^{a\hat{D}_x}$ denotes the exponential of the operator $a\hat{D}_x$, and we shall discover that it is itself a (unitary) operator (see Eq. (1.16)), called the space translation operator by a distance a in the x direction. Having established a connection with the Taylor expansion, I note that, in all its generality, the exponential of an operator $\hat{O}$ is defined in analogy with Eq. (1.18) as

$$e^{\hat{O}} = \hat{I} + \hat{O} + \frac{\hat{O}^2}{2!} + \frac{\hat{O}^3}{3!} + \cdots = \sum_{n=0}^{\infty} \frac{\hat{O}^n}{n!}, \tag{1.137}$$

[74] I follow Auletta and Wang (2014, Sect. 6.6).

allowing us to write

$$e^{a\hat{D}_x} = \hat{I} + a\hat{D}_x + \frac{a^2}{2}\hat{D}_x^2 + \cdots ,$$
$$= 1 + a\frac{d}{dx} + \frac{a^2}{2}\frac{d^2}{dx^2} + \cdots . \tag{1.138}$$

It is easy to check that when the operator $e^{a\hat{D}_x}$ is applied to x, i.e. to the LHS of Eq. (1.134), we get its RHS because only the constant term and the first derivative term in the expansion (1.138) are different from zero. This explains why $\hat{D}_x$ is referred to as the *generator* of space translations in the x direction.

Momentum Operator

Thus, the momentum operator in QM is closely related to the derivative operator $\hat{D}_x$. In the position representation, the corresponding 1D momentum operator takes the quantised form

$$\hat{p}_x = -i\hbar\frac{\partial}{\partial x}, \tag{1.139}$$

acting on kets and bras as follows (see Eqs. (1.127)),

$$\hat{p}_x\,|x\rangle = -i\hbar\frac{\partial}{\partial x}\,|x\rangle \quad \text{and} \quad \langle x\,|\,\hat{p}_x = -i\hbar\frac{\partial}{\partial x}\,\langle x\,|\,, \tag{1.140}$$

which precisely express the required variation of the position x determined by the momentum observable (and similarly if we use the wave function instead of the kets $|x\rangle$). Note that the states $|x\rangle$ are not eigenstates of the momentum (they represents superpositions relative to the latter) and therefore that observable acts like a true operator and not simply as a multiplier, as it is the case of $\hat{x}$ in Eqs. (1.127).

Note that I have used here the partial derivative $\partial/\partial x$ instead of the ordinary derivative d/dx. In fact, the 1D case can be considered as an extrapolation out of the 3D case. Moreover, when dealing with wave functions, we have both time and space dependency (although the former is often not explicitly written). This means that, to be exact, an arbitrary state vector $|\psi\rangle$ (Eq. (1.124)) should be expanded as $|\psi(t)\rangle = \int \psi(x,t)\,|x(t)\rangle\,dx$, although, in the following, I shall sometimes omit the dependence on time of the state vector when it is not strictly necessary. For a 1D particle in the position representation we get:

$$i\hbar\frac{\partial\psi(x,t)}{\partial t} = \left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x,t)\right]\psi(x,t)\ , \tag{1.141}$$

where I have used the explicit expression (1.12) for the Hamiltonian, while, for a free particle (when $V(x,t)=0$), we have[75]

$$\frac{\partial\psi(x,t)}{\partial t}=\frac{i\hbar}{2m}\frac{\partial^2\psi(x,t)}{\partial x^2}\ . \tag{1.142}$$

Analogously to what happens for the position operator, there will also be eigenvalues $p_x\in\mathbb{R}$ (where $-\infty<p_x<\infty$) and the corresponding eigenstates $|\,p_x\rangle$ of the 1D momentum operator $\hat{p}_x$ according to the eigenvalue equation

$$\hat{p}_x\,|\,p_x\rangle\ =p_x\,|\,p_x\rangle\ , \tag{1.143a}$$

or, in the bra formulation,

$$\langle p_x\,|\,\hat{p}_x=p_x\ \langle p_x\,|\,, \tag{1.143b}$$

which can be compared with Eq. (1.127b) for position. In other words, we are introducing here the *momentum representation*, in which the momentum operator $\hat{p}_x$ is simply represented by the real number p_x. In analogy with Eq. (1.129), the orthonormal conditions of the momentum eigenstates $|\,p_x\rangle$ are given by

$$\left\langle p_x\mid p_x'\right\rangle=\delta(p_x-p_x'), \tag{1.144}$$

where the Dirac delta function selects the value p_x'. Moreover, the completeness relation for the momentum eigenstates is given by the analogous of Eq. (1.122):

$$\int_{-\infty}^{+\infty}|\,p_x\rangle\,\langle p_x\,|\,dp_x=\hat{I}. \tag{1.145}$$

Momentum Eigenfunctions

To find the *momentum eigenfunction*

$$\varphi_p(x):=\langle x\mid p_x\rangle\in\mathbb{C}, \tag{1.146}$$

namely the counterpart for the momentum operator of what the position eigenfunction is for the position operator (see Eqs. (1.129)–(1.131)), we shall solve the eigenvalue equation (1.143a) in the position representation. Left multiplying both sides of Eq. (1.143a) by $\langle x\,|$, we obtain

$$\hat{p}_x\varphi_p(x)=p_x\varphi_p(x), \tag{1.147a}$$

[75] Note that Eq. (1.142) is formally identical to a classical diffusion equation.

or equivalently, using Eq. (1.140) for the explicit form of the momentum operator in the position representation,

$$-i\hbar\frac{\partial}{\partial x}\varphi_p(x) = p_x\varphi_p(x). \tag{1.147b}$$

The momentum eigenfunction, i.e. the solution to the above differential equation, is given by

$$\varphi_p(x) = \langle x \mid p_x\rangle = \frac{1}{\sqrt{2\pi\hbar}}\, e^{\frac{i}{\hbar}p_x x} \quad (-\infty < p_x < \infty), \tag{1.148}$$

where the normalisation constant $\frac{1}{\sqrt{2\pi\hbar}}$ is there to fulfil the orthonormal conditions (1.144). The validity of the solution can be verified by direct substitution of this expression into the eigenvalue equation (1.147b) and using the rule for the derivation of the exponential displayed in Table 1.1, Sect. 1.1.3. Like the position eigenstates, the momentum eigenstates also constitute a complete orthonormal basis, which allows us to write any state vector $|\psi\rangle$ as (see Eq. (1.120))

$$|\psi\rangle = \int_{-\infty}^{+\infty} |p_x\rangle\, \langle p_x \mid \psi\rangle\, dp_x, \tag{1.149}$$

where I have used the completeness relation (1.145). The scalar product $\langle p_x \mid \psi\rangle$ in the previous equation is the state vector $|\psi\rangle$ expressed in the momentum representation, and can be understood as a continuous function of the momentum eigenvalue p_x, that is,

$$\widetilde{\psi}(p_x) := \langle p_x \mid \psi\rangle \in \mathbb{C}, \tag{1.150}$$

which is called the *wave function in momentum space*. It has became convention to use the notation $\widetilde{\psi}(p_x)$, rather than the somewhat confusing one $\psi(p_x)$, in order to highlight the fact that, for a given state vector $|\psi\rangle$, the wave function in position space $\psi(x) = \langle x \mid \psi\rangle$ and the corresponding wave function in momentum space $\widetilde{\psi}(p_x) = \langle p_x \mid \psi\rangle$ in general have different functional forms. Analogous to the wave function in position space $\psi(x)$ (Eq. (1.123)), the wave function in momentum space $\widetilde{\psi}(p_x)$ is the probability amplitude whose square modulus $|\widetilde{\psi}(p_x)|^2$ gives probability density. In other words, the probability of finding a quantum system with momentum between p_x and $p_x + dp_x$ is given by $|\widetilde{\psi}(p_x)|^2 dp_x$.

Three-Dimensional Case

Generalisation to the Cartesian 3D case is straightforward:

$$\hat{\mathbf{p}} = -i\hbar\boldsymbol{\nabla} = -i\hbar\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) = \left(\hat{p}_x, \hat{p}_y, \hat{p}_z\right), \tag{1.151a}$$

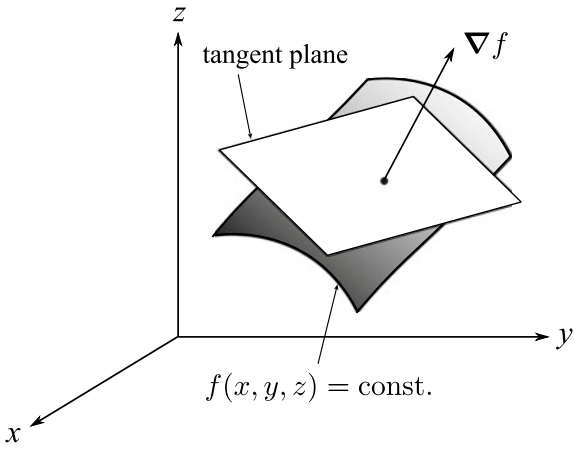


Fig. 1.13 The gradient of a function $f(x, y, z)$ at a certain point can be visualised as the vector normal to the plane tangent at that point to the surface $f(x, y, z) = \text{const.}$ (represented here in grey scale). The gradient points in the direction of the maximum rate of increase of $f(x, y, z)$ at that point, and whose magnitude is that rate of increase. Adapted from Auletta and Wang (2014, p. 135)

which is a row-vector formulation used also in CM, or also

$$\hat{\mathbf{p}}\,|\,\mathbf{r}\rangle = -i\hbar\nabla\,|\,\mathbf{r}\rangle\,, \tag{1.151b}$$

where

$$|\,\mathbf{r}\rangle = |\,x, y, z\rangle \tag{1.152}$$

is the 3D position eigenstate, so that the 3D position operator (always in row-vector representation)

$$\hat{\mathbf{r}} = (\hat{x}, \hat{y}, \hat{z}) \tag{1.153}$$

acts on that state as

$$\hat{\mathbf{r}}\,|\,\mathbf{r}\rangle = r\,|\,\mathbf{r}\rangle\,. \tag{1.154}$$

With the symbol ∇ we denote the nabla or del operator

$$\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right). \tag{1.155}$$

Its action on the function f expresses the *gradient*, which is a vector whose Cartesian components are the partial derivatives of $f(x, y, z)$, that is the row-vector

$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right). \tag{1.156}$$

The gradient of a function (or of a scalar field) at a point is the vector that points in the direction of maximum rate of increase of the function at that point, and whose

magnitude is precisely that rate of increase (Fig. 1.13).[76] Examples are represented by the gravitational potential (in this case the negative of the gradient points towards the gravitational field) or the concentration of a chemical (here the gradient points towards the maximum concentration). For this reason, the gradient of a scalar field is a *vector* field. An example of vector field is the velocity at each point in a moving fluid, like a hurricane. The gravitational field is both a scalar and a vector field: being the force field conservative, there is a scalar potential energy. Other examples of scalar fields are represented by temperature or density distribution of an object, or, for QM, the Higgs field. Using the expression

$$\psi(\mathbf{r}) = \langle \mathbf{r} \mid \psi \rangle \tag{1.157}$$

for the 3D wave function, the Schrödinger equation (1.141) in the Cartesian 3D case reads

$$i\hbar \frac{\partial \psi(\mathbf{r}, t)}{\partial t} = \left[-\frac{\hbar^2}{2m} \boldsymbol{\Delta} + V(\mathbf{r}, t) \right] \psi(\mathbf{r}, t) \ , \tag{1.158}$$

where

$$\boldsymbol{\Delta} = \boldsymbol{\nabla} \cdot \boldsymbol{\nabla} = \boldsymbol{\nabla}^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \tag{1.159}$$

is the *Laplacian* operator in Cartesian coordinates (after the name of the French physicist and mathematician Pierre–Simon de Laplace (1749–1827)), defined as the divergence of a gradient: the divergence is the quantity of a vector field's source at each point, or, more precisely, the volume density of the outward flux of a vector field from an infinitesimal volume around a given point.

The 3D momentum eigenstate is now denoted by

$$|\,\mathbf{p}\rangle = \left|\, p_x, p_y, p_z \right\rangle , \tag{1.160}$$

which satisfies the eigenvalue equations

$$\hat{\mathbf{p}}\,|\,\mathbf{p}\rangle = p\,|\,\mathbf{p}\rangle \ , \tag{1.161}$$

with $-\infty < p_x, p_y, p_z < \infty$. The completeness relation and orthonormal conditions are respectively given by

$$\int d^3 p\,|\,\mathbf{p}\rangle\,\langle\,\mathbf{p}\,| = \hat{I} \quad \text{and} \quad \left\langle \mathbf{p} \mid \mathbf{p}' \right\rangle = \delta^{(3)}(\mathbf{p} - \mathbf{p}'), \tag{1.162}$$

where

$$\int d^3 p = \int_{-\infty}^{+\infty} dp_x \int_{-\infty}^{+\infty} dp_y \int_{-\infty}^{+\infty} dp_z \tag{1.163}$$

[76] Byron and Fuller (1969–70, I, Sect. 1.7).

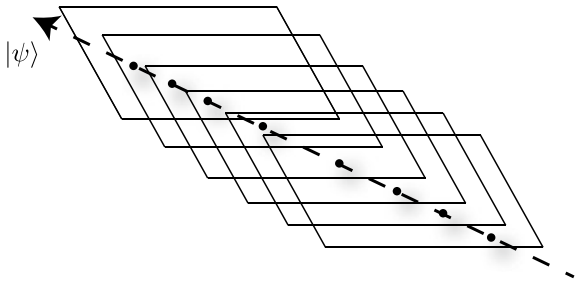


Fig. 1.14 De Broglie's waves as 1-form. The vector $|\psi\rangle$ pierces the different planes so that the surface pattern is given by the 1-form or covector $\langle \mathbf{k} |$ and their scalar product gives the different pierced surfaces

means the integral is taken over the whole 3D momentum space. Similarly for the 3D position operator. Moreover,

$$\widetilde{\psi}(\mathbf{p}) = \langle \mathbf{p} \mid \psi \rangle \tag{1.164}$$

is the Cartesian 3D counterpart of the wave function (1.150) as well as the 3D momentum eigenfunction (see Eq. (1.148)) is given by

$$\varphi_{\mathbf{p}}(\mathbf{r}) = \langle \mathbf{p} \mid \mathbf{r} \rangle = \frac{1}{(2\pi\hbar)^{3/2}} e^{\frac{i}{\hbar}\mathbf{p}\cdot\mathbf{r}}, \tag{1.165}$$

where $\mathbf{p} \cdot \mathbf{r} = p_x x + p_y y + p_z z$ is the dot product (1.62) of the vectors $\mathbf{p}$ and $\mathbf{r}$.

1-Form and Higher Forms

Let us now introduce the propagation vector

$$\hat{\mathbf{k}} = \frac{\hat{\mathbf{p}}}{\hbar} \tag{1.166}$$

which allows us to reformulate Eq. (1.164) as

$$\widetilde{\psi}(\mathbf{k}) = \langle \mathbf{k} \mid \psi \rangle \,, \tag{1.167}$$

as well as Eqs. (1.148) and (1.165) as

$$\varphi_k(x) = \frac{1}{\sqrt{2\pi\hbar}} e^{\frac{i}{\hbar}k_x x} \quad \text{and} \quad \varphi_{\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{8\pi^3\hbar^3}} e^{i\mathbf{k}\cdot\mathbf{r}} \,, \tag{1.168}$$

respectively. They are known as *plane* and *spherical waves*, respectively. Then, we can build surfaces with the same, integral phase spaced by (Fig. 1.14)

$$\frac{2\pi}{k} = \frac{hc}{E} = \frac{c}{\nu} = \lambda, \tag{1.169}$$

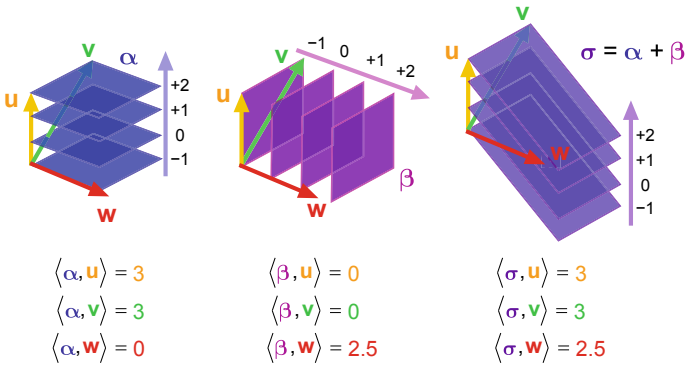


Fig. 1.15 1-forms $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ and their sum $\boldsymbol{\sigma}$ and vectors u, v, w, in 3D Euclidean space. The number of 1-form hyperplanes intersected by a vector equals the inner product (here, dealing with arbitrary vectors, in the form $\langle \cdot, \cdot \rangle$). Note that the inner product of orthogonal directions gives 0. Originally drawn in Misner et al. (1970, p. 57) and then remade at https://en.wikipedia.org/wiki/One-form

where I have used the magnitude of the propagation vector and Eqs. (1.1) and (1.3) with $n = 1$. This pattern of surfaces, determined by the scalar products with covectors, is called *1-form*.[77] It is also evident that 1-forms have directions (Fig. 1.15). Note that also the gradient can be understood as a (differential) 1-form. In fact, it represents the levels of a scalar field (Fig. 1.16, Left panel). This means that if this scalar field is denoted by f and $\boldsymbol{\xi}$ is any column vector, we have the scalar product

$$\nabla f \cdot \boldsymbol{\xi} = \boldsymbol{\xi}(f), \tag{1.170}$$

which parallels e.g. the wave function (1.150). Also higher forms can be built. A *2-form* is given by the following product (also called bivector) between 1-forms $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$:

$$\boldsymbol{\alpha} \wedge \boldsymbol{\beta}, \tag{1.171}$$

whose components are computed according to (Fig. 1.16, Right panel)

$$\alpha_{[j}\beta_{k]} = \frac{1}{2}(\alpha_j \beta_k - \alpha_k \beta_j). \tag{1.172}$$

In other words, a 2-form is a (commutator) combination of 1-forms.[78] The concept of commutator will be introduced below. Note that these elements are antisymmetric, i.e. $\alpha_{jk} = -\alpha_{kj}$. The operation can also be reiterated to get a 3-form and so on.

[77] Misner et al. (1970, pp. 53–59).

[78] Misner et al. (1970, p. 83).

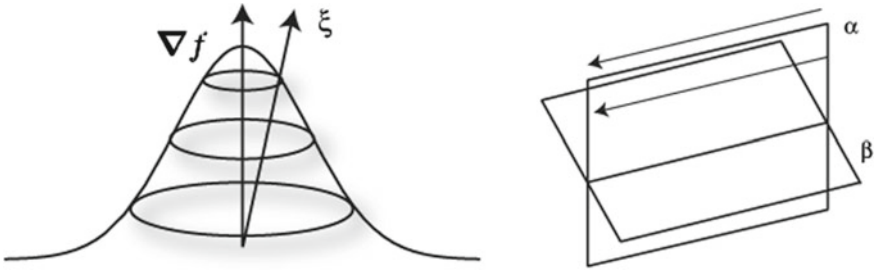


Fig. 1.16 Left panel: gradient of a scalar field as a 1-form: it displays the levels of the scalar field f (represented here by a Gaussian, after the name of the German mathematician and physicist Carl Friedrich Gauss (1777–1855)) while the vector $\boldsymbol{\xi}$ pierces the levels. The scalar product between ∇f and $\boldsymbol{\xi}$ gives the overlap between the direction of the 1-form and the latter vector (field). Right panel: an element, which is the intersection of two plane elements specified by the 1-forms $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, is described by the 2-form $\boldsymbol{\alpha} \wedge \boldsymbol{\beta}$. Note that the two 1-forms need to point into the same direction

Fourier Transform

Having a momentum representation, one may ask what is the form of the position observable in such a representation. As expected, it turns out to be an operator that can be shown to act on the wave function in momentum space as follows

$$\hat{x}\widetilde{\psi}(p_x) = i\hbar \frac{\partial}{\partial p_x}\widetilde{\psi}(p_x) \; . \tag{1.173}$$

We now have two representations (the position or the momentum representation) that we may use to express observables and states, and they are equivalent. In fact, the choice between these two representations is only a matter of convenience (for instance, determined by the physical context in which we describe a system), and the wave functions $\psi(x)$ and $\widetilde{\psi}(p_x)$ are related by a change of basis (Sect. 1.2.2), but using integrals, due to the continuous nature of the observables involved. To find the connection between $\psi(x)$ and $\widetilde{\psi}(p_x)$, given by Eqs. (1.123) and (1.150), and thus the connection between these two representations, we left multiply Eq. (1.149) by $\langle x \mid$ as well as Eq. (1.120) by $\langle p_x \mid$, so to obtain

$$\psi(x) = \langle x \mid \psi \rangle = \int_{-\infty}^{+\infty} \langle x \mid p_x \rangle \langle p_x \mid \psi \rangle \, dp_x = \int_{-\infty}^{+\infty} \langle x \mid p_x \rangle \, \widetilde{\psi}(p_x) dp_x \tag{1.174a}$$

and

$$\widetilde{\psi}(p_x) = \langle p_x \mid \psi \rangle = \int_{-\infty}^{+\infty} \langle p_x \mid x \rangle \langle x \mid \psi \rangle \, dx = \int_{-\infty}^{+\infty} \langle p_x \mid x \rangle \, \psi(x) dx, \tag{1.174b}$$

respectively. Now, $\langle x \mid p_x \rangle$ is the momentum eigenfunction $\varphi_p(x)$ given by Eq. (1.148), so that, using the property (1.64a), which here reads $\langle p_x \mid x \rangle = \langle x \mid p_x \rangle^*$,

we can rewrite the above equations as

$$\psi(x) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \widetilde{\psi}(p_x)\, e^{\frac{i}{\hbar} p_x x}\, dp_x \tag{1.175a}$$

and

$$\widetilde{\psi}(p_x) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \psi(x)\, e^{-\frac{i}{\hbar} p_x x}\, dx. \tag{1.175b}$$

These equations represent the connection between the position and momentum representations, and can be considered as the continuous counterpart of one of Eq. (1.114) and its inverse. Thus, the operators performing such transformations are unitary and are called the *Fourier transform* and the *inverse Fourier transform* (after the name of the French mathematician and physicist J.-B. Joseph Fourier (1768–1830))[79]

$$\hat{U}_{FT} := \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} e^{-\frac{i}{\hbar} p_x x} dx = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \langle p_x \mid x \rangle\, dx, \tag{1.176a}$$

$$\hat{U}_{FT}^{\dagger} := \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} e^{+\frac{i}{\hbar} p_x x} dp_x = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} \langle x \mid p_x \rangle\, dp_x, \tag{1.176b}$$

respectively, so that Eqs. (1.175) can be simply rephrased as

$$\hat{U}_{FT}\ \psi(x) = \widetilde{\psi}(p_x) \quad \text{and} \quad \hat{U}_{FT}^{\dagger}\ \widetilde{\psi}(p_x) = \psi(x). \tag{1.177}$$

Note that if

$$\int_{-\infty}^{+\infty} dx\, |\psi(x)|^2 = 1, \ \text{ then also } \int_{-\infty}^{+\infty} dp_x \left|\widetilde{\psi}(x)\right|^2 = 1. \tag{1.178}$$

This is called the Bessel–Parseval relation, after the names of the German astronomer, mathematician, physicist, and geodesist Friedrich Wilhelm Bessel (1784–1846) and the French mathematician Marc–Antoine Parseval des Chênes (1755–1836), a relation that contributed to lead to understanding the Fourier transform as an unitary operation.

Finally, the momentum eigenfunction (1.165) allows us to write the following 3D Fourier transforms between the wave functions $\psi(\mathbf{r})$ and $\widetilde{\psi}(\mathbf{p})$:

$$\psi(\mathbf{r}) = \frac{1}{(2\pi\hbar)^{3/2}} \int d^3p\, \widetilde{\psi}(\mathbf{p})\, e^{\frac{i}{\hbar}\mathbf{p}\cdot\mathbf{r}} \tag{1.179a}$$

and

$$\widetilde{\psi}(\mathbf{p}) = \frac{1}{(2\pi\hbar)^{3/2}} \int d^3r\, \psi(\mathbf{r})\, e^{-\frac{i}{\hbar}\mathbf{p}\cdot\mathbf{r}}. \tag{1.179b}$$

[79] Byron and Fuller (1969–70, I, Sect. 5.7).

Fourier's Analysis

In general, quantities like position and momentum which are related to each other by the Fourier transforms are called canonical conjugate quantities. Note that according to what is known as *Fourier's analysis*, any periodic function $f(x)$ can be decomposed in pure tones (essentially sine and cosine functions) oscillating n times in the recurrent period of length λ of $f(x)$, according to

$$\sin\left(\frac{2\pi nx}{\lambda}\right) \quad \text{and} \quad \cos\left(\frac{2\pi nx}{\lambda}\right), \tag{1.180}$$

which, thanks to the Euler formulas, after the name of the mentioned Swiss mathematician and physicist Leonhard Euler:

$$e^{i\phi} = \cos\phi + i\sin\phi \quad \text{and} \quad e^{-i\phi} = \cos\phi - i\sin\phi, \tag{1.181}$$

allow us to expand $f(x)$ in the series

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{\omega i n x}, \tag{1.182}$$

where $\omega = 2\pi/\lambda$ is the angular frequency and

$$c_n = \frac{1}{\lambda} F(\xi_n), \tag{1.183}$$

where $\xi_n = n/\lambda$ for some function $F(\xi_n)$. In the limit in which $\lambda \to \infty$ (and we replace sums by integrals), we get the classical Fourier transforms and its inverse:

$$F(x) = \int_{-\infty}^{+\infty} F(\xi) e^{\omega i n \xi x} d\xi, \quad F(\xi) = \int_{-\infty}^{+\infty} F(x) e^{-\omega i n x \xi} dx. \tag{1.184}$$

Normalisation with Continuous Functions

Let us consider a 1D system and an observable $\hat{O}$ with a continuous spectrum (1.97).[80] The state vector of the system can be expanded as

$$|\psi\rangle = \int do\, c(o)\, |o\rangle \ , \tag{1.185}$$

[80] Auletta et al. (2009, Sect. 2.2.2).

where the vectors $|o\rangle$ are the eigenkets of the observable $\hat{O}$, $c(o) = \langle o \mid \psi \rangle$, and $|c(o)|^2\, do$ represent the probability that the value of $\hat{O}$ can be found in the value interval $(o, o + do)$. Of course, we also have

$$\psi(x) = \langle x \mid \psi \rangle = \int do\, c(o)\, \langle x \mid o \rangle\,, \quad \psi^\star(x) = \langle \psi \mid x \rangle = \int do\, c^\star(o)\, \langle o \mid x \rangle\,, \tag{1.186}$$

which allow us to write

$$\begin{aligned} \int dx\, |\psi(x)|^2 &= \int dx \int doc(o)c^\star(o)\, \langle o \mid x \rangle\, \langle x \mid o \rangle \\ &= \int do\, |c(o)|^2\ . \end{aligned} \tag{1.187}$$

Substituting $\psi(x)$ as expanded in Eq. (1.186) into the RHS of the first row of Eq. (1.187), we obtain

$$\int do\, c^\star(o)c(o) = \int do\, c^\star(o) \left[\int dx \psi(x) \varphi_o^\star(x) \right]\,, \tag{1.188}$$

where, in analogy with Eq. (1.146), $\varphi_o(x) = \langle x \mid o \rangle$ is an eigenfunction of $\hat{O}$, which yields

$$c(o) = \int dx \psi(x) \varphi_o^\star(x)\ . \tag{1.189}$$

Back-substituting $\psi(x)$ as expanded in Eq. (1.186) into Eq. (1.189), we obtain

$$c(o) = \int do'\, c(o') \left[\int dx \varphi_{o'}(x) \varphi_o^\star(x) \right]\,, \tag{1.190}$$

from which we must conclude that

$$\begin{aligned} \int dx \varphi_{o'}(x) \varphi_o^*(x) &= \int dx \left\langle x \mid o' \right\rangle \langle o \mid x \rangle \\ &= \left| o' \right\rangle \langle o \right| \\ &= \delta(o - o')\,, \end{aligned} \tag{1.191}$$

where I have used the completeness relation (1.121). In other words, we get a Dirac delta function and, due to the properties (1.130) of the Dirac delta function, we have

$$\int dx\, |\varphi_o(x)|^2 = \infty, \tag{1.192}$$

and conclude that the eigenfunctions of an observable with a continuous spectrum are not normalisable. We can cast this in terms of Fourier transforms and rewrite Eq. (1.191) as

$$\hat{U}_{FT}\varphi_{o'}(x) = \int_{-\infty}^{+\infty} dx e^{-i2\pi ox} e^{i2\pi o'x} = \int_{-\infty}^{+\infty} dx e^{-i2\pi(o-o')x} = \delta(o-o'). \tag{1.193}$$

Energy Observable

Let us now consider the energy observable.[81] When the potential does not explicitly depend on time, the formal solution of the Schrödinger equation is given by Eq. (1.15):

$$|\psi(t)\rangle = e^{-\frac{i}{\hbar}\hat{H}t} |\psi(0)\rangle \ , \tag{1.194}$$

where $|\psi(0)\rangle$ is the state vector at time $t_0 = 0$. In general, however, it is not trivial to determine the action of the (unitary) operator $e^{-\frac{i}{\hbar}\hat{H}t}$ onto the state vector $|\psi(0)\rangle$. Of course, the eigenvectors and eigenvalues of $\hat{H}$ play a fundamental role in the determination of the time evolution of a quantum system. In fact, let us assume that the initial state vector of the system be an eigenstate of the Hamiltonian operator, i.e.

$$\hat{H} |\psi(0)\rangle = E |\psi(0)\rangle \ , \tag{1.195}$$

where $E \in \mathbb{R}$ is the corresponding eigenvalue of $\hat{H}$ for that state. Then, the action of the operator $e^{-\frac{i}{\hbar}\hat{H}t}$ onto the state vector $|\psi(0)\rangle$ results in a multiplication by the phase factor $e^{-\frac{i}{\hbar}Et}$, so that the state vector at time t can be written as

$$|\psi(t)\rangle = e^{-\frac{i}{\hbar}Et} |\psi(0)\rangle \ . \tag{1.196}$$

As a consequence, an energy eigenstate *does not evolve with time* because, due to orthonormality, a global (not relative) phase factor is irrelevant for the determination of a state (what counts are phase *differences* between components).

The situation is different when the initial state $|\psi(0)\rangle$ is not an energy eigenstate. In such a case, one has first to solve the eigenvalue equation for the Hamiltonian. For a discrete spectrum, one has

$$\hat{H} |\psi_n\rangle = E_n |\psi_n\rangle \ , \tag{1.197}$$

where the energy eigenvectors $|\psi_n\rangle$ are called *stationary states* and the $E_n \in \mathbb{R}$ are the corresponding eigenvalues, namely the energy levels of the system. In a second step, one has to expand the initial state vector $|\psi(0)\rangle$ onto the basis $\{|\psi_n\rangle\}$:

$$|\psi(0)\rangle = \sum_n c_n^{(0)} |\psi_n\rangle \tag{1.198}$$

[81] On this subject see Auletta et al. (2009, Sect. 3.1.3).

to determine the complex coefficients $c_n^{(0)}$. Thus, one is finally able to explicitly evaluate the RHS of Eq. (1.194) as

$$|\psi(t)\rangle = \sum_n e^{-\frac{i}{\hbar}E_n t} c_n^{(0)} |\psi_n\rangle \ , \tag{1.199}$$

which represents the time evolution of *any* initial state given its expansion in the basis of energy eigenvectors.

In the case of a continuous spectrum, the sums in Eqs. (1.197)–(1.199) have to be replaced by integration. Therefore, Eq. (1.197) becomes

$$\hat{H} \left| \psi_E \right\rangle = E \left| \psi_E \right\rangle \ , \tag{1.200}$$

while the expansion (1.198) can be written as

$$|\psi(0)\rangle = \int dE\, c^{(0)}(E) \left| \psi_E \right\rangle \ . \tag{1.201}$$

Finally, the time evolution of any initial state $|\psi(0)\rangle$ for the continuous case can be formulated as

$$|\psi(t)\rangle = \int dE\, c^{(0)}(E) e^{-\frac{i}{\hbar}Et} \left| \psi_E \right\rangle \ . \tag{1.202}$$

It is evident from the procedure above that the solution of the energy eigenvalue equation is a *necessary* step for the determination of the time evolution of any system. This is why Eq. (1.197) or Eq. (1.200) is time-independent and is often called *stationary Schrödinger equation*, while Eq. (1.10) is called *time-dependent Schrödinger equation*.

Note that the time-independent Schrödinger equation in the position representation is given by the following reformulation of Eq. (1.141):

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi_E(x) + V(x)\psi_E(x) = E\psi_E(x), \tag{1.203}$$

where $\psi_E(x) = \langle x \mid E \rangle$ is the energy eigenfunction. In three dimensions, the time-independent Schrödinger equation takes the form

$$-\frac{\hbar^2}{2m}\boldsymbol{\nabla}^2\psi_E(\mathbf{r}) + V(\mathbf{r})\psi_E(\mathbf{r}) = E\psi_E(\mathbf{r}), \tag{1.204}$$

where $\psi_E(\mathbf{r}) = \langle \mathbf{r} \mid E \rangle$ is the corresponding energy eigenfunction.

We have seen that introducing continuity in QM gives rise to three major and interdependent problems: (i) the problem of bounded or symmetric operators, (ii) the use of probability densities instead of probabilities, and (iii) the impossibility to normalise eigenfunctions. Let us discuss these aspects in terms of the energy observable. The energy spectrum of a system is given by the set of all possible

energy eigenvalues and stationary states correspond to these eigenvalues. As we know, such a spectrum can be continuous, discrete or a combination of both these possibilities. In particular, if the Hamiltonian is bounded from below (as in physically interesting cases), there is a state corresponding to the minimum eigenvalue that is called *ground state*, like, as mentioned in Sect. 1.1.2, the lowest energy level of an electron in an atom. At the opposite, we have a continuous spectrum when the system extends to infinity: this is typical for a free particle (like those that escape the atomic binding force as a consequence of the photoelectric effect). As we have seen, such a difference is in close relationship with the normalisation of the wave function describing the system: eigenfunctions corresponding to discrete eigenvalues are normalisable whereas eigenfunctions corresponding to continuous eigenvalues are not. In the former case $\int dx\, |\psi_E(x)|^2 < \infty$, which means that $\psi_E(x)$ tends to zero sufficiently fast for $x \to \pm\infty$ so that the integral converges. As a consequence, the probability of finding the particle at large distances goes rapidly to zero. This kind of states are called *bound states* (as those defining electrons trapped at certain orbital levels in an atom). Instead, if we consider a non-normalised wave function, i.e. corresponding to an eigenvalue belonging to the continuous spectrum, we have that $\int dx\, |\psi_E(x)|^2 = \infty$ (Eq. (1.192)). In this case, the eigenfunctions may not vanish at infinity. For instance, when $x \to \infty$, their absolute value may oscillate indefinitely. These states are called *unbound*. We could reject continuity on the basis of its ideal character (Sects. 1.1.2 and 1.2.2) and of the consequent impossibility to have true free systems. This is also true. Nevertheless, in several physical situations, it is of great help to assume continuity. Moreover, as we shall see, the continuity of free systems manifests a fundamental character of quantum systems. Thus, we need to accept the existence of such problems (and similar ones are also true for quantum fields) and explore whether there are better ways to deal with them.

1.2.4 Commutativity and Uncertainty Relations

Operations Often Do Not Commute

As mentioned in Sect. 1.2.2, an immediate consequence of the use of operators to describe quantum-mechanical observables (which is, in turn, a consequence of quantisation) is that some of them do not commute with each other. Elementary algebra tells us that all (real and imaginary) numbers commute, i.e. for arbitrary two numbers (or functions) a and b, we have $ab = ba$ (the order of factors in a product is irrelevant). This is not true of operators in general. Since, we have suggested to interpret operators as describing operations, we would like to introduce a very simple (classical) example of operations for understanding this property. Suppose that we are driving and executing the following subsequent operations. First, we turn left and then turn right. A quick reflection will show that inverting the order of the operations (that is, turning first right and then left) will bring us to a different location (we would, therefore, obtain a different result). We can mathematically represent that by

writing $\hat{R}\hat{L} \neq \hat{L}\hat{R}$, where I recall that we follow the usual convention to describe a succession of operations by use of multiplication (where the second operation is written to the left of the first one, and similarly for any subsequent operation).

Derivation of Commutation Relations

These concepts can also be applied to operators representing quantum observables. To prove that position and momentum (as well as other observables some of which correspond to the classical canonical conjugate pairs) do not commute was the contribution of the great German physicist Werner Heisenberg, and this turns out to be the cornerstone on which the new physical theory eventually called QM could be established.[82] To see that the position and momentum observables do not commute, let us consider the following expression (in the position representation):

$$\begin{aligned}\left(\hat{x}\,\hat{p}_x - \hat{p}_x\hat{x}\right)\psi(x) &= x\left(-i\hbar\frac{\partial}{\partial x}\right)\psi(x) + i\hbar\frac{\partial}{\partial x}\left[x\psi(x)\right] \\ &= -i\hbar x\frac{\partial}{\partial x}\psi(x) + i\hbar x\frac{\partial}{\partial x}\psi(x) + i\hbar\psi(x) \\ &= i\hbar\psi(x), \end{aligned} \tag{1.205}$$

where I have made use of the fact that the wave function $\psi(x)$ is a function of the position eigenvalue x as displayed by Eq. (1.127c) while the momentum acts on it as the operator given by Eq. (1.139). Note that in the second line use has been made of *Leibniz's product rule* (which is valid for both total and partial derivation), after the name of G. W. Leibniz,

$$\frac{d(fg)}{dx} = \frac{df}{dx}g + f\frac{dg}{dx}, \tag{1.206}$$

where $f(x)$ and $g(x)$ are differentiable functions.[83] Note that the product rule for exponential and logarithmic functions gives

$$\frac{d}{dx}e^{f(x)} = f'(x)e^{f(x)} \quad \text{and} \quad \frac{d}{dx}\ln(f(x)) = \frac{1}{f(x)}f'(x), \tag{1.207}$$

respectively. Since we have made no specific assumption on the wave function $\psi(x)$, we are allowed to say that the previous result is valid for an arbitrary quantum state: it holds true in the momentum representation as well and in fact it is representation independent. Thus, we can use the momentum representation, such that the position acts as a true operator as shown in Eq. (1.173) while momentum is a numerical

[82] Heisenberg (1925). For historical reconstruction see Jammer (1966, Sects. 5.1–5.2).

[83] In fact, the commutation relation can be considered as a particular case of Leibniz's rule (Penrose 2004, pp. 494–95).

multiplier as displayed by its eigenvalue equation (1.143a), and obtain an analogous result:

$$\begin{aligned}\left(\hat{x}\hat{p}_x - \hat{p}_x\hat{x}\right)\widetilde{\psi}(p_x) &= i\hbar\frac{\partial}{\partial p_x}\left[p_x\widetilde{\psi}(p_x)\right] - p_x i\hbar\frac{\partial}{\partial p_x}\widetilde{\psi}(p_x)\\ &= i\hbar\widetilde{\psi}(p_x) + i\hbar p_x\frac{\partial}{\partial p_x}\widetilde{\psi}(p_x) - i\hbar p_x\frac{\partial}{\partial p_x}\widetilde{\psi}(p_x)\\ &= i\hbar\widetilde{\psi}(p_x). \end{aligned} \tag{1.208}$$

Commutators

Then, we can affirm that, for any state, we always have the relation

$$[\hat{x}, \hat{p}_x] = i\hbar\hat{I}, \tag{1.209}$$

where the expression

$$[\hat{O}, \hat{O}'] := \hat{O}\hat{O}' - \hat{O}'\hat{O} \tag{1.210}$$

is called the *commutator* (or also Lie bracket, after the name of the Norwegian mathematician Sophus Lie (1842–1899)) of observables $\hat{O}$ and $\hat{O}'$. A relation between operators like that given by Eq. (1.209) is called the *commutation relation*. Similar commutation relations hold true for the other Cartesian components of the two observables[84]:

$$[\hat{x}, \hat{p}_x] = [\hat{y}, \hat{p}_y] = [\hat{z}, \hat{p}_z] = i\hbar\hat{I} \tag{1.211a}$$

while different Cartesian components always commute

$$[\hat{x}, \hat{p}_y] = [\hat{x}, \hat{p}_z] = [\hat{y}, \hat{p}_x] = [\hat{y}, \hat{p}_z] = [\hat{z}, \hat{p}_x] = [\hat{z}, \hat{p}_y] = 0. \tag{1.211b}$$

Because position and momentum are canonical conjugate quantities, the commutation relations of position and momentum (1.211) are also known as the canonical commutation relations.

What the commutation relation, for example, between x and p_x tells us is that the x Cartesian components of position and momentum are not jointly measurable. In fact, the necessary and sufficient condition for two observables to be jointly measured is that they commute (as it happens for all observables of CM). Note that when two quantum observables commute, then there exists a complete orthonormal set of common eigenvectors.[85] The example in Fig. 1.17 shows that also projectors, as any other operator (apart from identity), may not commute with other projectors. The

[84] Here I follow Auletta and Wang (2014, Sect. 6.8).

[85] Byron and Fuller (1969–70, I, Sect. 4.7).

lesson is that in order to get a certain outcome at all, the output state must be some superposition involving the input state as a component.

Note that commutators satisfy the following properties (for arbitrary operators $\hat{A}, \hat{B}, \hat{C}$):

- $[\hat{A}, \hat{B}] = -[\hat{B}, \hat{A}]$, (1.212a)
- $[\alpha\hat{A}, \hat{B}] = \alpha[\hat{A}, \hat{B}]$, (1.212b)
- $[\alpha\hat{A} + \beta\hat{B}, \hat{C}] = \alpha[\hat{A}, \hat{C}] + \beta[\hat{B}, \hat{C}]$, (1.212c)
- $[\hat{A}\hat{B}, \hat{C}] = \hat{A}[\hat{B}, \hat{C}] + [\hat{B}, \hat{C}]\hat{A}$, (1.212d)
- $[\hat{A}, [\hat{B}, \hat{C}]] + [\hat{C}, [\hat{A}, \hat{B}]] + [\hat{B}, [\hat{C}, \hat{A}]] = 0$, (1.212e)

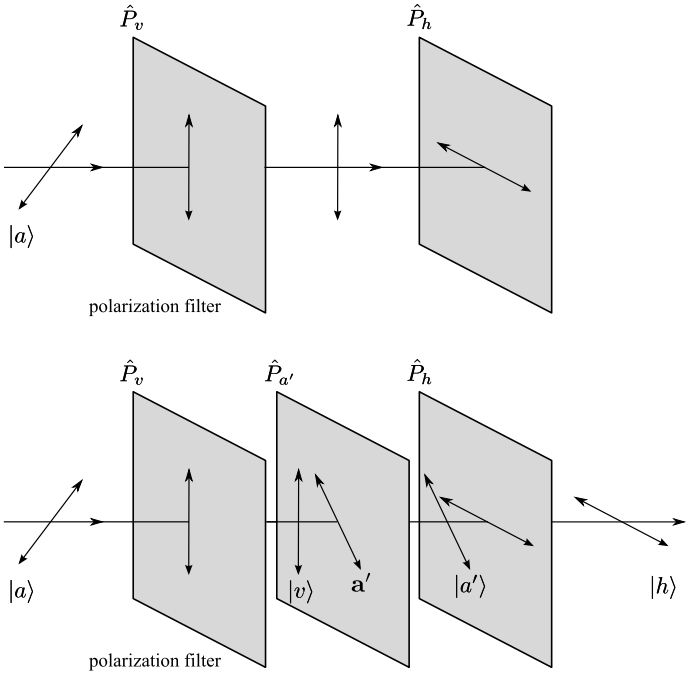


Fig. 1.17 Top: A light beam in an arbitrary polarisation state $|\,a\rangle$ along an arbitrary polarisation direction **a** is sent through the filter represented by projector $\hat{P}_v$ with vertical polarisation direction. After passing through the filter, the light beam is in the state $|\,v\rangle$ (with a reduced intensity, since some photons have been blocked). Then, no photon can pass through the filter $\hat{P}_h$ whose horizontal polarisation direction is orthogonal to that of the photons in state $|\,v\rangle$. In other words, any measurement outcome cannot be orthogonal to the input state (this is the only knowledge that we can have about a single measurement when the input state is unknown). Bottom: If a third polarisation filter $\hat{P}_{a'}$ is placed between the first and the last one, with an arbitrary polarisation direction $\mathbf{a}'$ (different from **a**), then some photons will pass through the last filter. This also shows that changing the order of projectors we get a different result, so that projectors may do not commute. Adapted from Auletta et al. (2009, p. 62)

where $\alpha, \beta \in \mathbb{C}$ and equality (1.212e) is known as the Jacobi identity, after the name of the mathematician Carl Jacobi.

Poisson's Brackets

Quantum commutators are the counterpart of the classical *Poisson's brackets*, after the name of the French mathematician and physicist Siméon Poisson.[86] In fact, the Hamilton equations (1.45) can be written in terms of the Poisson brackets as

$$\dot{q_j} = \{q_j, H\}, \quad \dot{p_j} = \{p_j, H\} \,, \tag{1.213}$$

where the Poisson brackets for two arbitrary functions f and g are defined as

$$\{f, g\} := \sum_j \left(\frac{\partial f}{\partial q_j} \frac{\partial g}{\partial p_j} - \frac{\partial f}{\partial p_j} \frac{\partial g}{\partial q_j} \right) \,. \tag{1.214}$$

They display the following properties:

$$\{f, g\} = -\{g, f\} \,, \tag{1.215a}$$
$$\{f, C\} = 0 \,, \tag{1.215b}$$
$$\{Cf + C'g, h\} = C\{f, h\} + C'\{g, h\} \,, \tag{1.215c}$$
$$0 = \{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} \,, \tag{1.215d}$$
$$\frac{\partial}{\partial t} \{f, g\} = \left\{ \frac{\partial f}{\partial t}, g \right\} + \left\{ f, \frac{\partial g}{\partial t} \right\} \,, \tag{1.215e}$$

where C, C' are constants and h is a third function. Equation (1.215d) is the classical Jacobi identity, and thus the counterpart of (1.212e). The advantage of this notation is that, for any physical quantity that is a function f of q and p, we can write its time evolution as

$$\frac{d}{dt} f = \{f, H\} \,, \tag{1.216}$$

which is zero when f represents a conserved quantity.

Orbital Angular Momentum

Commutation relations can be applied also to other canonical conjugate observables. The most important ones, together with those of position and momentum, are those among different components of the *orbital angular momentum*, the dynamical observable that rules the changes in position of a rotating body. It is necessarily

[86] Landau and Lifshitz (1976, Sect. 42). See also Landsman (2017, Sect. 3.2).

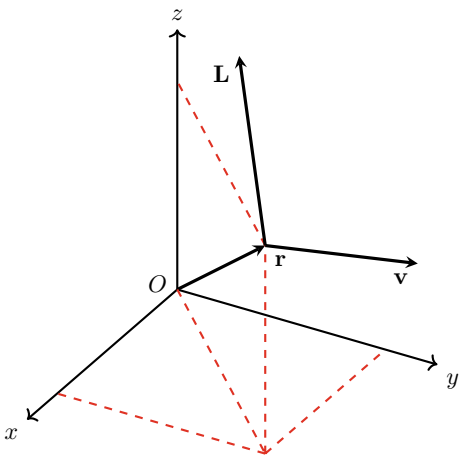


Fig. 1.18 A classical particle with mass m, velocity $\mathbf{v}$, and position $\mathbf{r}$ with respect to a certain reference frame $Oxyz$ has an angular momentum $\mathbf{L} = \mathbf{r} \times \mathbf{p} = m\mathbf{r} \times \mathbf{v}$

a Cartesian 3D observable that classically is given by the vector product between position and momentum[87]:

$$\mathbf{L} := \mathbf{r} \times \mathbf{p} \,, \tag{1.217}$$

where $\mathbf{r}$ is the position and $\mathbf{p}$ the linear momentum with respect to a given Cartesian reference frame $Oxyz$ (Fig. 1.18); thus, angular momentum has the physical dimension $[\mathrm{M}][\mathrm{L}^2][\mathrm{T}^{-1}]$. In mathematics, the *cross product* (or vector product) of two Euclidean vectors $\mathbf{a}$ and $\mathbf{b}$ is a vector that is orthogonal to both $\mathbf{a}$ and $\mathbf{b}$ (i.e. normal to the plane containing them), with a direction provided by the so-called right-hand rule and a magnitude equal to the area of the parallelogram that the vectors span. The direction of the cross product $\mathbf{a} \times \mathbf{b}$ is determined by placing $\mathbf{a}$ and $\mathbf{b}$ *tail-to-tail*, flattening the right hand, extending it in the direction of $\mathbf{a}$, and then curling the fingers in the direction that the angle $\mathbf{b}$ makes with $\mathbf{a}$. The thumb then points in the direction of $\mathbf{a} \times \mathbf{b}$ (Fig. 1.19). Note that the resultant of the cross product is an *axial vector* (or pseudovector). Pseudovectors are quantities that transform like vectors under a proper rotation, but do not change sign under inversion.[88] Geometrically it is the opposite (i.e. of equal magnitude but in the opposite direction), of its mirror image. Instead of, a true (or *polar*) vector, like $\mathbf{r}$ or $\mathbf{p}$, changes sign under inversion.

Summarising, while the scalar product is defined by Eq. (1.60) and is a number, the cross product $\mathbf{a} \times \mathbf{b}$ is the axial vector defined by

$$\mathbf{a} \times \mathbf{b} := \|\mathbf{a}\| \, \|\mathbf{b}\| \, \sin\theta \, \mathbf{e}_n, \tag{1.218}$$

where $\|\mathbf{a}\|$ and $\|\mathbf{b}\|$ are the respectively magnitudes of vectors $\mathbf{a}$, $\mathbf{b}$, $0 \leq \theta \leq \pi$ is the angle between $\mathbf{a}$ and $\mathbf{b}$, and $\mathbf{e}_n$ is a unit vector perpendicular to both $\mathbf{a}$ and $\mathbf{b}$. The

[87] On this subject see the extensive treatment in Auletta et al. (2009, Sect. 6.1) but also Auletta and Wang (2014, Sects. 8.1–8.4) for a simplified exposition.

[88] Byron and Fuller (1969–70, I, p. 17).

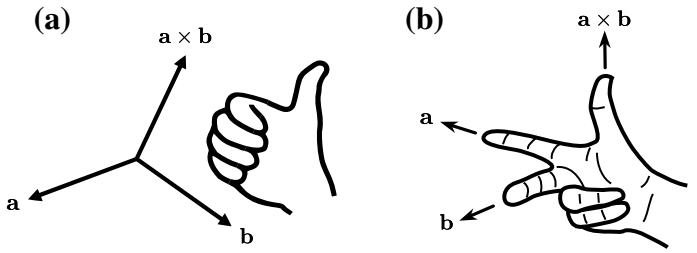


Fig. 1.19 The right-hand rule. The rule as defined in the text **(a)**. The right-hand rule for establishing a system of Cartesian coordinates, here the vectors **a** and **b** need to be orthogonal, and could be thought of as x and y, respectively, so that their vector product would be z **(b)**. Adapted from http://en.wikipedia.org/wiki/File:Right-hand_grip_rule.svg and http://en.wikipedia.org/wiki/File:Right_hand_rule_cross_product.svg

cross product satisfies the following properties:

- $\mathbf{a} \times \mathbf{a} = 0,$ (1.219a)
- $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a},$ (1.219b)
- $\mathbf{a} \times (\mathbf{b} + \mathbf{c}) = (\mathbf{a} \times \mathbf{b}) + (\mathbf{a} \times \mathbf{c}),$ (1.219c)
- $(k\mathbf{a}) \times \mathbf{b} = \mathbf{a} \times (k\mathbf{b}) = k(\mathbf{a} \times \mathbf{b})$ (1.219d)
- $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) + \mathbf{b} \times (\mathbf{c} \times \mathbf{a}) + \mathbf{c} \times (\mathbf{a} \times \mathbf{b}) = 0,$ (1.219e)

where $k \in \mathbb{R}$. Using Cartesian versors, the vector product of the vectors $\mathbf{a} = a_x\mathbf{e}_x + a_y\mathbf{e}_y + a_z\mathbf{e}_z$ and $\mathbf{b} = b_x\mathbf{e}_x + b_y\mathbf{e}_y + b_z\mathbf{e}_z$ is in component form[89]

$$\mathbf{a} \times \mathbf{b} = (a_y b_z - a_z b_y)\mathbf{e}_x + (a_z b_x - a_x b_z)\mathbf{e}_y + (a_x b_y - a_y b_x)\mathbf{e}_z. \tag{1.220}$$

Generator of Rotations

We consider now how a generic ket $|\psi\rangle$ is modified by the action of the unitary rotation operator, that is,

$$\hat{U}_{\mathbf{R}}(\theta) = e^{\frac{i}{\hbar}\theta\mathbf{n}\cdot\hat{\mathbf{L}}} = e^{i\theta\mathbf{n}\cdot\hat{\mathbf{I}}}\,, \tag{1.221}$$

which describes a rotation by an angle θ about the direction given by the versor $\mathbf{n}$, and where $\hat{\mathbf{L}}$ is the quantum angular momentum (that has the same general form of its classical counterpart) while

$$\hat{\mathbf{I}} = \frac{\hat{\mathbf{L}}}{\hbar} = \hat{\mathbf{r}} \times \hat{\mathbf{k}} \tag{1.222}$$

[89] This result can also be written in terms of determinants Byron and Fuller (1969–70, I, Sect. 3.8).

is its reduced form, $\hat{\mathbf{k}}$ being the propagation vector (1.166). Without loss of generality we may consider infinitesimal rotations, i.e.rotations by an angle $\delta\theta \to 0$. In such a case, in analogy with the procedure (1.137), by Taylor-expanding the unitary operator (1.221) to the first order we get

$$\hat{U}_{\mathbf{R}}(\delta\theta) \simeq 1 + i\delta\theta\mathbf{n}\cdot\hat{\mathbf{I}} = 1 + i\delta\theta\hat{\mathbf{R}}\ , \tag{1.223}$$

where $\simeq$ means "asymptotically equal to", while $\hat{\mathbf{R}} = \mathbf{n}\cdot\hat{\mathbf{I}}$ is called the *generator of the rotation* about the direction given by the vector $\mathbf{n}$ (in other words, it is the projection of $\hat{\mathbf{L}}$ along the direction $\mathbf{n}$). The unitary operator (1.223) induces, on a generic operator $\hat{O}$, the transformation

$$\begin{aligned}\hat{O} \mapsto \hat{O}' \simeq \hat{U}_{\mathbf{R}}(\delta\theta)\hat{O}\hat{U}^{\dagger}_{\mathbf{R}}(\delta\theta) &= \left(1 + i\delta\theta\hat{\mathbf{R}}\right)\hat{O}\left(1 - i\delta\theta\hat{\mathbf{R}}\right) \\ &= \hat{O} + \delta\hat{O}\ ,\end{aligned} \tag{1.224}$$

where, to the first order in $\delta\theta$,

$$\delta\hat{O} \simeq i\delta\theta\left[\hat{\mathbf{R}}, \hat{O}\right]\ . \tag{1.225}$$

As a consequence, if $\hat{O}$ is a scalar for rotations, it has to commute with the generator $\hat{\mathbf{R}}$ of the rotation, and, in such a case, $\delta\hat{O} = 0$. This is the reason why any component of the quantum orbital angular momentum $\hat{\mathbf{L}}$ commutes with $\hat{\mathbf{L}}^2$, $\hat{\mathbf{R}}$, and $\hat{\mathbf{p}}^2$, all of which are scalars for the rotations.

Let us write the explicit form of the generator of rotations.[90] For the 2D case we have

$$\hat{\mathbf{R}}(\phi) = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}\ , \tag{1.226}$$

and for the general three-dimensional case the matrix which describes a rotation about the axes x, y, z (by the Euler angles β, ϕ, and θ, respectively) is[91]

$$\hat{\mathbf{R}}(\beta,\phi,\theta) = \begin{bmatrix} \cos\beta\cos\phi\cos\theta - \sin\beta\sin\theta & \sin\beta\cos\phi\cos\theta + \cos\beta\sin\theta & -\sin\phi\cos\theta \\ -\cos\beta\cos\phi\sin\theta - \sin\beta\cos\theta & -\sin\beta\cos\phi\sin\theta + \cos\beta\cos\theta & \sin\phi\sin\theta \\ \cos\beta\sin\phi & \sin\beta\sin\phi & \cos\phi \end{bmatrix}\ . \tag{1.227}$$

Cartesian Components of the Angular Momentum

The three Cartesian components of the quantum orbital angular momentum,

[90] Byron and Fuller (1969–70, I, Sect. 1.4).

[91] I follow for the Euler angles the $x - y - z$-convention.

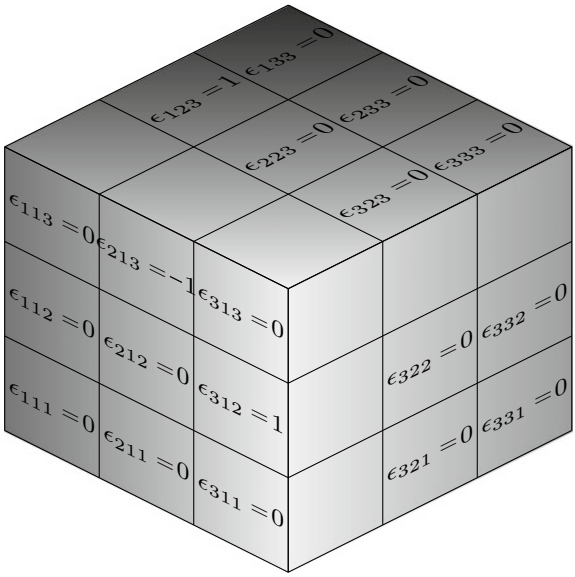


Fig. 1.20 Pictorial representation of the Levi–Civita tensor. A cube is segmented into 27 little boxes, according to the values (1, 2 or 3) of the indices j, k, n. Each little box then represents one component of the tensor. Only the values of the visible boxes are shown. Of the 27 components of the Levi–Civita tensor only six (those for which the indices are different from each other) are different from zero: they can have the values $+1$ or -1 depending on whether the sequence of the indices can be obtained from the sequence 123 with an even or an odd number of permutations, respectively. Adapted from Auletta et al. (2009, p. 195)

$$\hat{\mathbf{L}} = \hat{L}_x + \hat{L}_y + \hat{L}_z, \tag{1.228}$$

in the position representation are derived from Eqs. (1.217) and (1.220) but making use of the commutation relations (1.205):

$$\hat{L}_x = \hat{y}\hat{p}_z - \hat{z}\hat{p}_y = -i\hbar\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right), \tag{1.229a}$$

$$\hat{L}_y = \hat{z}\hat{p}_x - \hat{x}\hat{p}_z = -i\hbar\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right), \tag{1.229b}$$

$$\hat{L}_z = \hat{x}\hat{p}_y - \hat{y}\hat{p}_x = -i\hbar\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right), \tag{1.229c}$$

or, in more compact form,

$$\hat{L}_j = \epsilon_{jkn}\hat{r}_k\hat{p}_n, \tag{1.230}$$

11	12	13
21	22	23
31	32	33

Fig. 1.21 Graphical representation of a matrix. Here it is represented a 3×3 matrix. Each component of the matrix is represented by a small square. There are 9 combinations. Note that the 3 rows and the 3 columns represent 3 row and 3 columnar vectors, respectively

where j, k, n are indices which may assume the values $1, 2, 3$, $\hat{r}_1 = \hat{x}$, $\hat{r}_2 = \hat{y}$, and $\hat{r}_3 = \hat{z}$, and similarly $\hat{L}_1 = \hat{L}_x$, $\hat{L}_2 = \hat{L}_y$, and $\hat{L}_3 = \hat{L}_z$. The symbol ϵ refers to the Levi–Civita tensor (after the name of the Italian mathematician Tullio Levi–Civita (1873–1941)), which is antisymmetric (it changes sign by exchange of two indices):

$$\epsilon_{ijk} = -\epsilon_{jik} = \epsilon_{ikj}, \tag{1.231}$$

and is zero when two indices are equal, while $\epsilon_{123} = 1$ (Fig. 1.20). In Eq. (1.230) and in the following of this chapter, a summation over repeated indices is understood. Note that the Levi–Civita tensor is particularly interesting when expressing volumes that are the result of a mix of scalar and cross products:

$$\epsilon_{jkn} a_j b_k c_n = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}). \tag{1.232}$$

While matrices can be graphically represented as in Fig. 1.21 (the elements of which can be thought of as connections between two indices), a tensor can be represented as in Fig. 1.22 (whose elements can be thought of as connections between three indices). Sometimes the term *tensor* is used also for denoting matrices. Indeed, scalars are considered zeroth rank tensors, vectors as first rank tensors, and matrices as second rank ones.

A lengthy derivation finally shows that the commutation relations among the different Cartesian components of the orbital angular momentum are given by

$$\left[\hat{L}_x, \hat{L}_y\right] = i\hbar\hat{L}_z \ , \ \left[\hat{L}_y, \hat{L}_z\right] = i\hbar\hat{L}_x \ , \ \left[\hat{L}_z, \hat{L}_x\right] = i\hbar\hat{L}_y \ , \tag{1.233}$$

which can be written using Levi–Civita tensor as

$$\left[\hat{L}_j, \hat{L}_k\right] = i\hbar\epsilon_{jkn}\hat{L}_n \ . \tag{1.234}$$

In other words, the three components of the angular momentum are not simultaneously measurable. Instead, as mentioned, any component is measurable jointly with

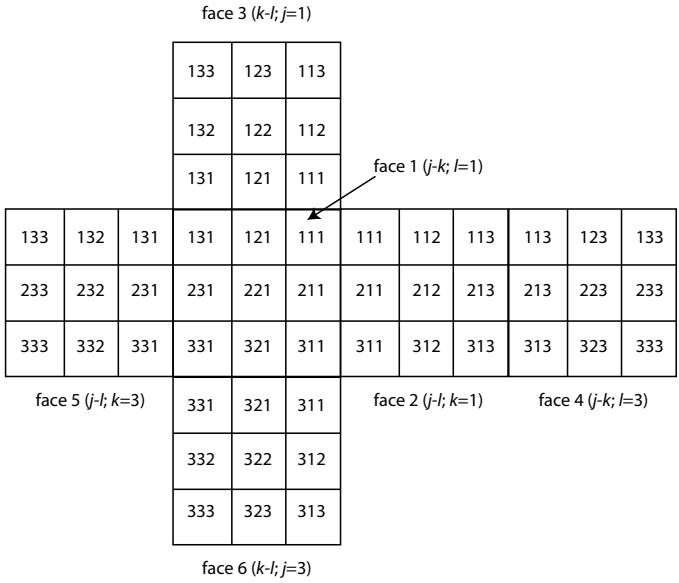


Fig. 1.22 Graphical representation of a tensor as an "opened" cube. A $3 \times 3 \times 3$ tensor is represented here. Each component of the tensor is represented by a small cube. There are 27 combinations but the central one (222) cannot be seen. Note that the six surfaces have two variables and a third constant (that can assume 1 or 3 as values)

the square of the total orbital angular momentum since commutes with the latter:

$$\left[\hat{L}_j, \hat{\mathbf{L}}^2\right] = 0 \,, \tag{1.235}$$

with $j = x, y, z$.

I incidentally observe that further commutation relations are possible: between time and energy, between angular momentum and angle, and so on.[92]

Quantisation of the Orbital Angular Momentum

In QM, angular momentum is quantised, that is, it cannot vary continuously but can only take discrete values. Quantisation of angular momentum was first postulated by Niels Bohr in his model of the atom (Sect. 1.1.2)[93]: starting from the angular momentum quantisation, Bohr was able to calculate the energies of the allowed orbits of the hydrogen atom and other hydrogen-like atoms and ions. We shall now find the eigenvalues and eigenstates of the quantised angular momentum operator. As

[92] On these subjects see Auletta et al. (2009, Sects. 3.9, 6.5).

[93] Bohr (1913).

mentioned, only the magnitude squared of the angular momentum and one, and only one, of its components are simultaneously measurable. The convention is to choose the latter as the z component of the angular momentum operator, $\hat{L}_z$. Note that, since it is up to us which space direction we denote as the z axis, the eigenvalues of the component of the angular momentum in an arbitrary direction are always quantised as in the following.

The commutation relation $[\hat{L}_z, \hat{\mathbf{L}}^2] = 0$ means that the operators $\hat{L}_z$ and $\hat{\mathbf{L}}^2$ have common eigenstates. Let us denote these common eigenstates by $|l, m\rangle$, where l and m are (discrete) quantum numbers that are respectively related to the eigenvalues of the operators $\hat{\mathbf{L}}^2$ and $\hat{L}_z$. Let us postulate that the eigenvalue equation for $\hat{L}_z$, in agreement with Eq. (1.118), has the following form:

$$\hat{L}_z\,|l, m\rangle = m\hbar\,|l, m\rangle\,, \tag{1.236}$$

that is, the states $|l, m\rangle$ are eigenstates of $\hat{L}_z$ and the respective eigenvalues are $m\hbar$. The number m is called the *magnetic quantum number*, whose value will be determined below. Since the expectation value of the square of an Hermitian operator is non-negative, we must have

$$m^2\hbar^2 = \langle l, m|\hat{L}_z^2|l, m\rangle = \langle l, m|(\hat{\mathbf{L}}^2 - \hat{L}_x^2 - \hat{L}_y^2)|l, m\rangle \le \langle l, m|\hat{\mathbf{L}}^2|l, m\rangle, \tag{1.237}$$

where I have used the obvious identity

$$\hat{\mathbf{L}}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2, \tag{1.238}$$

and the fact that the eigenstates $|l, m\rangle$ are normalised. Note that the orthonormal conditions and completeness relation for the angular momentum eigenstates $|l, m\rangle$ as

$$\langle l, m \mid l', m'\rangle = \delta_{ll'}\delta_{mm'} \quad \text{and} \quad \sum_{l=0}^{\infty}\sum_{m=-l}^{l} |l, m\rangle\,\langle l, m| = \hat{I}, \tag{1.239}$$

respectively. Inequality (1.237) shows that m^2 is bounded by the eigenvalues of $\hat{\mathbf{L}}^2$, hence we have

$$-l \le m \le l, \tag{1.240}$$

where the eigenvalue l is the maximum value of m and is called the *azimuthal quantum number*.

Our next goal is to find the possible values of m and l. Since we have a range of value of m for each fixed l, it is convenient to use the reduced form (1.222) and to introduce the raising and lowering operators

$$\hat{l}_+ = \hat{l}_x + i\hat{l}_y, \quad \hat{l}_- = \hat{l}_x - i\hat{l}_y, \tag{1.241}$$

which satisfy

$$\hat{l}_+ = \hat{l}_-^\dagger, \quad \hat{l}_- = \hat{l}_+^\dagger. \tag{1.242}$$

Thus, the operators $\hat{l}_+$ and $\hat{l}_-$ are not observables since, being each the Hermitian conjugate of the other, are not Hermitian operators. Let us now consider the action of $\hat{L}_z$ on the state vector $\hat{l}_+ |l, m\rangle$

$$\begin{aligned}
\hat{L}_z \hat{l}_+ |l, m\rangle &= \left(\hat{L}_z \hat{l}_+ - \hat{l}_+ \hat{L}_z + \hat{l}_+ \hat{L}_z\right) |l, m\rangle \\
&= \left(\left[\hat{L}_z, \hat{l}_+\right] + \hat{l}_+ \hat{L}_z\right) |l, m\rangle \\
&= \hbar \hat{l}_+ |l, m\rangle + \hat{l}_+ \hat{L}_z |l, m\rangle \\
&= \hbar \hat{l}_+ |l, m\rangle + m\hbar \hat{l}_+ |l, m\rangle \\
&= (m+1)\hbar \hat{l}_+ |l, m\rangle \,,
\end{aligned} \tag{1.243}$$

where I have made use of Eq. (1.236), the following commutation relations Eq. (1.233)

$$\begin{aligned}
\left[\hat{L}_z, \hat{l}_+\right] &= \left[\hat{L}_z, \frac{\hat{L}_x + i\hat{L}_y}{\hbar}\right] \\
&= \frac{1}{\hbar}\left[\hat{L}_z, \hat{L}_x\right] + \frac{i}{\hbar}\left[\hat{L}_z, \hat{L}_y\right] \\
&= i\hat{L}_y + \hat{L}_x \\
&= +\hbar \hat{l}_+,
\end{aligned} \tag{1.244}$$

and property (1.212a). From Eq. (1.243) it follows that $\hat{l}_+ |l, m\rangle$ is an eigenstate of $\hat{L}_z$ with eigenvalue $(m+1)\hbar$. Since the eigenstates $\hat{l}_+ |l, m\rangle$ and $|l, m+1\rangle$ correspond to the same eigenvalue of $\hat{L}_z$, they have to be proportional to each other. Thus, we can write

$$\hat{l}_+ |l, m\rangle = c_{lm}^+ |l, m+1\rangle \,, \tag{1.245}$$

where c_{lm}^+ is a proportional constant. A similar derivation shows that the state vector $\hat{l}_- |l, m\rangle$ is an eigenstate of $\hat{L}_z$ with eigenvalue $(m-1)\hbar$, and therefore

$$\hat{l}_- |l, m\rangle = c_{lm}^- |l, m-1\rangle \,, \tag{1.246}$$

where c_{lm}^- is again a proportional constant. Now, since according to Eq. (1.240) l and $-l$ are the maximum and minimum eigenvalues of the magnetic quantum number m, respectively, we must also have

$$\hat{l}_+ |l, l\rangle = 0, \quad \hat{l}_- |l, -l\rangle = 0. \tag{1.247}$$

Therefore, we can start with the highest eigenstate $|l, l\rangle$ of $\hat{L}_z$ and apply recursively the lowering operator $\hat{l}_-$ on it in order to obtain the other eigenstates $|l, l-1\rangle$, $|l, l-2\rangle$, ... of $\hat{L}_z$, where the process has to stop after finite steps because $\hat{l}_- |l, -l\rangle = 0$. Thus, it is possible to order the eigenstates of $\hat{L}_z$ in descending order of m as

$$|l, l\rangle \ , \ |l, l-1\rangle \ , \ \ldots, \ |l, -l+1\rangle \ , \ |l, -l\rangle \ . \tag{1.248}$$

Of course, the same result can be obtained by starting with the lowest eigenstate $|l, -l\rangle$ of $\hat{L}_z$ and applying recursively the raising operator $\hat{l}_+$ on that eigenstate. As a consequence, for each possible value of l, there are $2l+1$ possible eigenstates of $\hat{L}_z$. Moreover, since $2l+1$ must be a positive integer, the value of l should be either a non-negative integer or a positive half-integer (i.e. a number of the form $n + \frac{1}{2}$, where n is a non-negative integer). As we will see below, for the orbital angular momentum currently under discussion, the azimuthal quantum number l can only take integer values, while the half-integer values characterise the spin quantum number that will be discussed below. The set of the $2l+1$ states in Eq. (1.248) that are related by the raising and lowering operators $\hat{l}_+$ and $\hat{l}_-$ are called a multiplet.

In order to find the eigenvalues of $\hat{\mathbf{L}}^2$, we first write the product of the raising and lowering operators (1.241):

$$\begin{aligned} \hat{l}_- \hat{l}_+ &= \left(\hat{l}_x - i\hat{l}_y\right)\left(\hat{l}_x + i\hat{l}_y\right) \\ &= \hat{l}_x^2 + \hat{l}_y^2 + i\left[\hat{l}_x, \hat{l}_y\right] \\ &= \hat{\mathbf{l}}^2 - \hat{l}_z^2 - \hat{l}_z . \end{aligned} \tag{1.249}$$

Since (i) from the first identity in Eq. (1.247) and the previous result we have

$$\left[\hat{\mathbf{l}}^2 - \hat{l}_z\left(\hat{l}_z + 1\right)\right] |l, l\rangle = \hat{l}_- \hat{l}_+ |l, l\rangle = 0, \tag{1.250}$$

and (ii) from Eq. (1.236) it follows that

$$\hat{L}_z\left(\hat{L}_z + 1\right) |l, l\rangle = l(l+1)\hbar^2 |l, l\rangle \ , \tag{1.251}$$

where we do not need to consider global phase factor (allowing the back-translation from lower-case operators to upper case ones), we can finally obtain

$$\left[\hat{\mathbf{L}}^2 - l(l+1)\hbar^2\right] |l, l\rangle = 0 \text{ or } \hat{\mathbf{L}}^2 |l, l\rangle = l(l+1)\hbar^2 |l, l\rangle \ . \tag{1.252}$$

Using the fact that $\hat{\mathbf{L}}^2$ commutes with $\hat{l}_-$, we can generalise the above result to other values of m and obtain

$$\hat{\mathbf{L}}^2 |l, m\rangle = l(l+1)\hbar^2 |l, m\rangle \ , \tag{1.253}$$

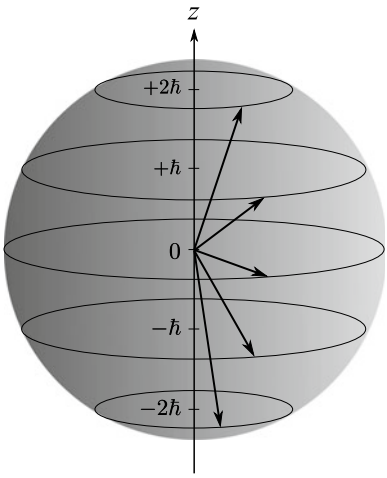


Fig. 1.23 Quantisation of angular momentum for the case of $l = 2$. In this case, there are five angular momentum eigenstates in the multiplet. The vectors from top to bottom illustrate respectively the angular momentum with $L_z = +2\hbar,\ +\hbar,\ 0,\ -\hbar,\ -2\hbar$. The magnitude of the angular momentum is given by $L = \sqrt{2(2+1)}\hbar = \sqrt{6}\hbar$. Adapted from Auletta and Wang (2014, p. 215)

which, in agreement with Eq. (1.118), is the eigenvalue equation for the operator $\hat{\mathbf{L}}^2$ that we were looking for, and the eigenvalue of the operator $\hat{\mathbf{L}}^2$ turns out to be $l(l+1)\hbar^2$ instead of $l^2\hbar^2$ as one would have expected. It is a peculiarity of QM that the eigenvalue of $\hat{\mathbf{L}}^2$ is not the square of the maximum eigenvalue of $\hat{L}_z$, and this is a direct consequence of the fact that components of the angular momentum operator do not commute with one another (Eq. (1.249)). In other words, the angular momentum and its z component can never be aligned. Indeed, if they would commute, the last term in the parentheses on the RHS of Eq. (1.250) were absent and the eigenvalue of $\hat{\mathbf{L}}^2$ be equal to $l^2\hbar^2$ (see Fig. 1.23, displaying the case $l = 2$).

Eigenfunctions of the Orbital Angular Momentum

We have not yet derived the specific forms of the eigenstates according to particular values of l and m. To this purpose, a more concrete approach is suitable to the same problem, by considering eigenfunctions of the angular momentum operator. However, a full treatment of the angular momentum eigenfunctions is a mathematically challenging subject to be dealt with here.[94] In the present context, I shall limit the exposition to a very sketchy presentation bypassing most of the details.

The 3D angular momentum eigenfunctions $\psi_{lm}(\mathbf{r})$ are the angular momentum eigenstates $|l, m\rangle$ represented in the 3D position eigenbasis $\{|\mathbf{r}\rangle\}$ (Eq. (1.152)), that is,

[94]The interested reader may have a look at Auletta et al. (2009, Chap. 6).

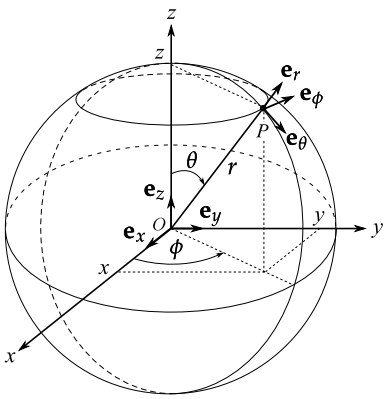


Fig. 1.24 Relation between Cartesian (rectangular) coordinates (x, y, z) and spherical coordinates (r, θ, ϕ) of a point P in three-dimensional space. The versor $\mathbf{e}_r$ makes an angle θ with the Cartesian versor $\mathbf{e}_z$. Adapted from Auletta and Wang (2014, p. 217)

$$\psi_{lm}(\mathbf{r}) = \langle \mathbf{r} \mid l, m \rangle \,. \tag{1.254}$$

Due to the rotational character of the angular momentum, it is convenient to use spherical coordinates instead of the usual Cartesian (rectangular) coordinates. In spherical coordinates the position of a point P in 3D space is specified by three numbers (r, θ, ϕ), where r (with $0 \leq r < \infty$) is the radial distance from the origin O to P, θ (with $0 \leq \theta \leq \pi$) is the polar angle of the line segment OP measured from the z axis, and ϕ (with $0 \leq \phi < 2\pi$) is the azimuthal angle of the line segment OP measured from the x axis to its orthogonal projection on the xy plane (Fig. 1.24). (Variations of θ individuate parallels while variations of ϕ individuate meridians.) The spherical coordinates (r, θ, ϕ) of a point are related to the Cartesian coordinates (x, y, z) by[95]:

$$r = \sqrt{x^2 + y^2 + z^2}, \quad \theta = \arctan \frac{\sqrt{x^2 + y^2}}{z}, \quad \phi = \arctan \frac{y}{x}, \tag{1.255a}$$

where the first one expresses Pitagora's theorem in 3D, arctan is the inverse of the tangent trigonometric function given by $\tan\theta = \sin\theta / \cos\theta$. In other words, the function $y = \arctan x$ is defined such that $\tan y = x$. Conversely, the Cartesian coordinates (x, y, z) can be obtained from the spherical coordinates (r, θ, ϕ) by

$$x = r \, \sin\theta \, \cos\phi, \quad y = r \, \sin\theta \, \sin\phi, \quad z = r \, \cos\theta. \tag{1.255b}$$

I note that relations of this kind that describe conversions between two coordinate systems of the same space are called coordinate transformation. In particular, the spherical basis $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_\phi\}$ can be expressed in terms of the Cartesian basis $\{\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z\}$ as (see Fig. 1.24)

[95] See Auletta and Wang (2014, Sect. 8.4) for details about the transformation from Cartesian into spherical coordinates and vice versa.

$$\mathbf{e}_r = \sin\theta\cos\phi\,\mathbf{e}_x + \sin\theta\sin\phi\,\mathbf{e}_y + \cos\theta\,\mathbf{e}_z, \tag{1.256a}$$
$$\mathbf{e}_\theta = \cos\theta\cos\phi\,\mathbf{e}_x + \cos\theta\sin\phi\,\mathbf{e}_y - \sin\theta\,\mathbf{e}_z, \tag{1.256b}$$
$$\mathbf{e}_\phi = -\sin\phi\,\mathbf{e}_x + \cos\phi\,\mathbf{e}_y. \tag{1.256c}$$

Of course, $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_\phi\}$ is an orthonormal basis. The inverse of the above transformations can be written as

$$\mathbf{e}_x = \sin\theta\cos\phi\,\mathbf{e}_r + \cos\theta\cos\phi\,\mathbf{e}_\theta - \sin\phi\,\mathbf{e}_\phi, \tag{1.257a}$$
$$\mathbf{e}_y = \sin\theta\sin\phi\,\mathbf{e}_r + \cos\theta\sin\phi\,\mathbf{e}_\theta + \cos\phi\,\mathbf{e}_\phi, \tag{1.257b}$$
$$\mathbf{e}_z = \cos\theta\,\mathbf{e}_r - \sin\theta\,\mathbf{e}_\theta. \tag{1.257c}$$

From transformation (1.255a), we obtain the three Cartesian partial derivatives expressed in spherical coordinates

$$\frac{\partial r}{\partial x} = \sin\theta\cos\phi, \quad \frac{\partial\theta}{\partial x} = \frac{1}{r}\cos\theta\cos\phi, \quad \frac{\partial\phi}{\partial x} = -\frac{1}{r}\frac{\sin\phi}{\sin\theta}, \tag{1.258a}$$
$$\frac{\partial r}{\partial y} = \sin\theta\sin\phi, \quad \frac{\partial\theta}{\partial y} = \frac{1}{r}\cos\theta\sin\phi, \quad \frac{\partial\phi}{\partial y} = \frac{1}{r}\frac{\cos\phi}{\sin\theta}, \tag{1.258b}$$
$$\frac{\partial r}{\partial z} = \cos\theta, \quad \frac{\partial\theta}{\partial z} = -\frac{1}{r}\sin\theta, \quad \frac{\partial\phi}{\partial z} = 0, \tag{1.258c}$$

where I have used the mathematical formula

$$\frac{d}{dx}\arctan x = \frac{1}{1+x^2} \tag{1.259}$$

as well as the inverse coordinate transformations (1.255b).

In spherical coordinates (which is still a kind of position representation) the angular momentum eigenfunctions are defined as

$$\psi_{lm}(r,\theta,\phi) := \langle r,\theta,\phi \,|\, l,m\rangle\,, \tag{1.260}$$

where the state vector $|r,\theta,\phi\rangle$ in spherical coordinates may be compared with its Cartesian counterpart (1.152). In order to find $\psi_{lm}(r,\theta,\phi)$ we need to solve the eigenvalue equations (1.236) and (1.253) in spherical coordinates. From the expressions (1.229) for the angular momentum components $\hat{L}_x$, $\hat{L}_y$, and $\hat{L}_z$ in Cartesian coordinates and the relation (1.255a), after some calculations (that are unnecessary here) we can rewrite $\hat{L}_z$ and $\hat{\mathbf{L}}^2$ in spherical coordinates as

$$\hat{L}_z = -i\hbar\frac{\partial}{\partial\phi}, \tag{1.261a}$$

$$\hat{\mathbf{L}}^2 = -\hbar^2\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right], \tag{1.261b}$$

where, for their derivation, extensive use has been made of the chain rule of derivatives

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx}, \tag{1.262}$$

where $y = f(u)$ and $u = g(x)$ are differentiable functions. Note that the above expressions only depend on the angles θ and ϕ. Indeed, the use of spherical coordinates is advantageous, allowing the factorisation of the angular momentum eigenfunction $\psi_{lm}(r, \theta, \phi)$ into the radial (depending on r) and angular (depending on both θ and ϕ) parts as

$$\psi_{lm}(r, \theta, \phi) = f(r)Y_{lm}(\theta, \phi), \tag{1.263}$$

allowing us to dismiss $f(r)$ and to write the eigenfunctions as

$$Y_{lm}(\phi, \theta) = \langle\theta, \phi \,|\, l, m\rangle\,. \tag{1.264}$$

Collecting the results (1.261) and (1.263), the eigenvalue equations (1.236) and (1.253) in spherical coordinates reduce to

$$\hat{L}_z Y_{lm}(\phi, \theta) = m\hbar Y_{lm}(\phi, \theta), \tag{1.265a}$$

$$\hat{\mathbf{L}}^2 Y_{lm}(\phi, \theta) = l(l+1)\hbar^2 Y_{lm}(\phi, \theta). \tag{1.265b}$$

The eigenfunctions $Y_{lm}(\phi, \theta)$ represent the probability amplitudes whose square moduli yield the angular probability distributions of finding the system under consideration, what means that $Y_{lm}(\phi, \theta)$ be single-valued functions of θ and ϕ. Mathematically, this is equivalent to the requirement that l and m are integers, as already anticipated. In mathematical terms the angular momentum eigenfunctions $Y_{lm}(\phi, \theta)$ are known as the *spherical harmonics*, which in general are defined as

$$Y_{lm}(\theta, \phi) = \sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}}\, P_l^m(\cos\theta)\, e^{im\phi}, \tag{1.266}$$

where the $P_l(x)$'s are called the *Legendre polynomials*, after the name of the French mathematician Adrien–Marie Legendre (1752–1833). The first explicit expressions are[96]

[96] Byron and Fuller (1969–70, I, Sects. 1.8 and 5.5).

$$P_0(\cos\theta) = 1, \tag{1.267a}$$
$$P_1(\cos\theta) = \cos\theta, \tag{1.267b}$$
$$P_2(\cos\theta) = \frac{1}{2}\left(3\cos^2\theta - 1\right), \tag{1.267c}$$
$$P_3(\cos\theta) = \frac{1}{2}\left(5\cos^3\theta - 3\cos\theta\right), \tag{1.267d}$$

where the $P_l^m(x)$ are the *associated Legendre polynomials*

$$P_l^m(x) = (-1)^m(1-x^2)^{m/2}\frac{d^m}{dx^m}P_l(x), \text{ with } P_l(x) = \frac{1}{2^l l!}\frac{d^l}{dx^l}(x^2-1)^l. \tag{1.268}$$

Spherical harmonics are quite general tools used for representing functions defined on the surface of a sphere just as circular functions like sine and cosine are used for representing functions on a circle through Fourier series (1.182). I summarise here some of the basic properties satisfied by the spherical harmonics:

- $Y_{lm}(\pi-\theta, \pi+\phi) = (-1)^l Y_{lm}(\theta,\phi)$, (1.269a)
- $\int_0^{2\pi}\int_0^{\pi} Y^*_{lm}(\theta,\phi)Y_{lm}(\theta,\phi)\sin\theta d\theta d\phi = \delta_{ll'}\delta_{mm'}$, (1.269b)
- $\sum_{l=0}^{\infty}\sum_{m=l}^{l} Y_{lm}(\theta,\phi)Y^*_{lm}(\theta',\phi') = \delta(\cos\theta-\cos\theta')\delta(\phi-\phi')$. (1.269c)

The last two of the previous equations are a reformulation of the orthonormal conditions and completeness relation for the angular momentum eigenstates, respectively (see Eqs. (1.239)).

From definition (1.266), it is possible to derive the first few spherical harmonics.[97] For $l = 0$ we have

$$Y_{0,0}(\theta,\phi) = \frac{1}{\sqrt{4\pi}}. \tag{1.270}$$

For $l = 1$ we have

$$Y_{1,0}(\theta,\phi) = \sqrt{\frac{3}{4\pi}}\cos\theta, \quad Y_{1,\pm1}(\theta,\phi) = \mp\sqrt{\frac{3}{8\pi}}\sin\theta\, e^{\pm i\phi}. \tag{1.271}$$

Finally, for $l = 2$ we have

[97] Byron and Fuller (1969–70, I, Sect. 5.8).

$$Y_{2,0}(\theta,\phi) = \sqrt{\frac{5}{16\pi}}\,(3\cos^2\theta - 1), \tag{1.272a}$$

$$Y_{2,\pm1}(\theta,\phi) = \mp\sqrt{\frac{15}{8\pi}}\,\sin\theta\cos\theta\, e^{\pm i\phi}, \tag{1.272b}$$

$$Y_{2,\pm2}(\theta,\phi) = \sqrt{\frac{15}{32\pi}}\,\sin^2\theta\, e^{\pm 2i\phi}. \tag{1.272c}$$

Note that the use of the plus-minus sign ($\pm$) and the minus-plus sign ($\mp$) in Eqs. (1.271) and (1.272) is a shorthand notation to present *two* equations in one expression, in which the upper and lower signs are interrupted separately and independently.

The names s-states (or s-waves), p-states (or p-waves), d-states (or d-waves), and f-states (or f-waves) are used to refer to angular momentum eigenstates with quantum number $l = 0, 1, 2$ and 3, respectively. The symbols s, p, d, and f stand for the terms *sharp*, *principal*, *diffuse*, and *fundamental*, respectively, which are derived from the characteristics of their spectroscopic lines. The rest angular momentum eigenstates with higher azimuthal quantum numbers are named in alphabetical order.

Uncertainty Relations

Given the result (1.205), it is evident that the state vector (or the wave function) cannot be understood as a classical state, that is, a state for which two observables of a canonical pair (like position and momentum) possess determined values at the same time; as Dirac himself stressed (see Sect. 1.2.1), it is impossible to define the state of a quantum system by ascribing determined values to all observables.[98]

In fact, from the commutation relations between position and momentum Heisenberg, in another historical paper, derived their uncertainty relation.[99] Let us start from a general case making use of the expectation or mean value (1.102). The *variance* of an observable is defined as the measure of the spread of the measurement results about its expectation value, taking account of the probabilities of all possible results. The variance of the observable $\hat{O}$ in the state $|\psi\rangle$ is then defined as

$$\begin{aligned}\Delta^2_\psi \hat{O} &:= \left\langle \left(\hat{O} - \left\langle \hat{O} \right\rangle_\psi\right)^2 \right\rangle_\psi \\ &= \left\langle \psi \left| \left(\hat{O} - \left\langle \psi \left| \hat{O} \right| \psi \right\rangle\right)^2 \right| \psi \right\rangle .\end{aligned} \tag{1.273}$$

[98] Auletta (2000, pp. 32, 35, 127, 216–17, 291–94).

[99] Heisenberg (1927). For historical reconstruction see Jammer (1966, Sect. 7.1). On this subject see also Auletta (2000, Chap. 7) and literature therein. For a rather technical account of this subject see Auletta et al. (2009, Sect. 2.3.2).

In words, it is the expectation value of the squared deviation of an observable from its own expectation value with both taken on the same state, here $|\psi\rangle$. I recall that $\hat{O}$ is a Hermitian operator and $\left\langle \hat{O} \right\rangle_\psi$ is by definition real. Using the norm (1.72) of a vector, we can reformulate the variance $\Delta^2_\psi \hat{O}$ as the norm squared of the following state

$$\left|\psi'\right\rangle = (\hat{O} - \langle \hat{O} \rangle_\psi)\,|\psi\rangle\ , \tag{1.274}$$

since, according to Eq. (1.78), if $|\psi\rangle$ is a state vector and $\hat{O}$ is an arbitrary observable, also $\hat{O}\,|\psi\rangle$ is a state vector. In other words, we get:

$$\Delta^2_\psi \hat{O} = \left\langle \psi' \mid \psi' \right\rangle = \| \left|\psi'\right\rangle \|^2 \geq 0, \tag{1.275}$$

where, as for all norms, equality holds iff (if and only if) $\left|\psi'\right\rangle = 0$. This means that the variance of an observable must be real and non-negative. The *uncertainty* of the observable $\hat{O}$ in the state $|\psi\rangle$ is then defined as the square root of its variance in that state (or as the norm of $\left|\psi'\right\rangle$), i.e.

$$\Delta_\psi \hat{O} := \sqrt{\left\langle \left(\hat{O} - \langle \hat{O} \rangle_\psi \right)^2 \right\rangle_\psi} = \| \left|\psi'\right\rangle \| \ . \tag{1.276}$$

This justifies the use of the above notation $\Delta^2_\psi \hat{O}$ for the variance of the observable $\hat{O}$ in the state $|\psi\rangle$.

We can now derive the *uncertainty relation* for non-commuting observables. Let us consider two (arbitrary and likely non-commuting) observables $\hat{O}$ and $\hat{O}'$ and, without loss of generality, assume that the expectation values $\left\langle \hat{O} \right\rangle_\psi$ and $\left\langle \hat{O}' \right\rangle_\psi$ of the two observables in the state $|\psi\rangle$ are both equal to zero (this, however, does not imply that the expectation values in all states are zero). With these assumptions, the uncertainties of the observable $\hat{O}$ and $\hat{O}'$ in the state $|\psi\rangle$ reduce to

$$\Delta_\psi \hat{O} = \sqrt{\langle \psi | \hat{O}^2 | \psi \rangle} \quad \text{and} \quad \Delta_\psi \hat{O}' = \sqrt{\langle \psi | (\hat{O}')^2 | \psi \rangle}. \tag{1.277}$$

Let us now consider the two states (see again Eq. (1.78))

$$|\varphi\rangle := \left| \hat{O}\psi \right\rangle \quad \text{and} \quad \left|\varphi'\right\rangle := \left| \hat{O}'\psi \right\rangle . \tag{1.278}$$

The latter equations allow us to rewrite the expressions (1.277) in terms of the norms of the vectors (1.278):

$$\Delta_\psi \hat{O} = \sqrt{\langle \varphi \mid \varphi \rangle} \quad \text{and} \quad \Delta_\psi \hat{O}' = \sqrt{\langle \varphi' \mid \varphi' \rangle}. \tag{1.279}$$

From a pure mathematical point of view, all vectors $|\varphi\rangle$, $\left|\varphi'\right\rangle$, satisfy the *Cauchy–Schwarz inequality* (after the names of the French mathematician Augustin–Louis

Cauchy (1789–1857) and the German mathematician Hermann A. Schwarz (1843–1921)):

$$\left|\langle\varphi \mid \varphi'\rangle\right| \leq \sqrt{\langle\varphi \mid \varphi\rangle}\sqrt{\langle\varphi' \mid \varphi'\rangle}. \tag{1.280}$$

This is obviously true for the case that we are considering since the scalar product of two arbitrary orthonormal vectors $|a\rangle$ and $|b\rangle$ satisfies $0 \leq |\langle a \mid b\rangle| \leq 1$ and we always have $\langle a \mid a\rangle = \langle b \mid b\rangle = 1$ (see Eq. (1.65a)). By back-substituting Eqs. (1.278) into the inequality (1.280), we obtain

$$\left|\left\langle\psi\left|\hat{O}\hat{O}'\right|\psi\right\rangle\right| \leq \sqrt{\langle\psi \mid \hat{O}^2 \mid \psi\rangle}\sqrt{\langle\psi \mid (\hat{O}')^2 \mid \psi\rangle} = \Delta_\psi \hat{O}\ \Delta_\psi \hat{O}', \tag{1.281a}$$

where I have used also Eqs. (1.277) and the fact that $\hat{O}$ and $\hat{O}'$ are Hermitian operators. Obviously, a similar result can be obtained if we interchange the positions of the two observables, that is,

$$\left|\left\langle\psi\left|\hat{O}'\hat{O}\right|\psi\right\rangle\right| \leq \sqrt{\langle\psi \mid \hat{O}^2 \mid \psi\rangle}\sqrt{\langle\psi \mid (\hat{O}')^2 \mid \psi\rangle} = \Delta_\psi \hat{O}\ \Delta_\psi \hat{O}'. \tag{1.281b}$$

Let us consider the last two equations. From the triangle inequality for arbitrary numbers $z_1, z_2 \in \mathbb{C}$

$$|z_1 + z_2| \leq |z_1| + |z_2|, \tag{1.282}$$

it follows that, if $a \in \mathbb{R}$ such that $|z_1|, |z_2| \leq a$, then we also have

$$|z_1 - z_2| \leq 2a. \tag{1.283}$$

This, together with the fact that the uncertainty of an observable is real and non-negative, allows us to derive

$$\left|\left\langle\psi\left|\hat{O}\hat{O}'\right|\psi\right\rangle - \left\langle\psi\left|\hat{O}'\hat{O}\right|\psi\right\rangle\right| \leq 2\,\Delta_\psi \hat{O}\ \Delta_\psi \hat{O}', \tag{1.284}$$

which, since

$$\left\langle\psi\left|\hat{O}\hat{O}' - \hat{O}'\hat{O}\right|\psi\right\rangle = \left\langle\left[\hat{O}, \hat{O}'\right]\right\rangle_\psi, \tag{1.285}$$

can also be expressed in the equivalent form

$$\Delta_\psi \hat{O}\ \Delta_\psi \hat{O}' \geq \frac{1}{2}\left|\left\langle\left[\hat{O}, \hat{O}'\right]\right\rangle_\psi\right|, \tag{1.286}$$

which is usually called the *generalised uncertainty relation*.[100] This result is particularly interesting because it shows that the uncertainty relation is a direct consequence of the non-commutativity between quantum observables. Moreover, it is a general

[100] Robertson (1929).

result which deals with any pair of arbitrary observables (not necessarily conjugate). Since this inequality is true for an arbitrary pair of observables (if they commute the RHS vanishes and the inequality is trivially satisfied), it must also be true for the position and momentum observables. By taking into account again the commutation relations (1.209), we finally obtain the uncertainty relation for position and momentum[101]

$$\Delta x \, \Delta p_x \geq \frac{\hbar}{2}, \tag{1.287}$$

and similarly for the other Cartesian directions. Here, I have dropped any reference to the state since this conclusion is of general validity and does not depend on the specific state we have chosen.

Time-Energy Uncertainty Relation

As for commutation relations, also in the case of uncertainty relations, we can establish similar relationships among several pairs of physical observables.[102] I shall examine here the case of energy-time in particular. Consider a wave packet (which, I recall, is a wave concentrated in a relatively small spatial region, as mentioned in Sects. 1.1.3–1.2.1) obtained by superposing a large number of 1D plane waves (see the first of Eq. (1.168)) in a small region of width $(\Delta k)^{-1}$ around its centre, where k is the magnitude of the propagation vector (1.166). In this case, we can describe the wave packet as having width Δx and group velocity v_g (with this expression is meant the speed of the centre of the wave packet). As a consequence, the exact time t at which the wave packet crosses a certain point is defined with an uncertainty

$$\Delta t \approx \frac{\Delta x}{v_g} \, , \tag{1.288}$$

where I recall that the symbol $\approx$ means "approximately equal to". On the other hand, thanks to Eq. (1.7), the wave packet has an energy uncertainty ΔE due to its spread in momentum space

$$\Delta E \approx \frac{\partial E}{\partial p_x} \Delta p_x = v_g \Delta p_x \, , \tag{1.289}$$

since, as for the classical case (1.45), we have

$$\frac{\partial E}{\partial p_x} = \frac{\partial}{\partial p_x} \left(\frac{p_x^2}{2m} \right) = \frac{p_x}{m} = v_g . \tag{1.290}$$

The product of Eqs. (1.288) and (1.289) yield

[101] Heisenberg (1927).

[102] See Auletta et al. (2009, Sects. 3.8, 4.4, 6.5, 13.3).

$$\Delta t \, \Delta E \approx \Delta x \, \Delta p_x \ . \tag{1.291}$$

By using the momentum–position uncertainty relation (1.287), we then derive the energy-time uncertainty relation

$$\Delta E \, \Delta t \geq \frac{\hbar}{2} \ , \tag{1.292}$$

which limits the product of the spread ΔE of the energy spectrum of the wave packet and the accuracy Δt of the prediction of the time of passage at a given point. Among the other uncertainty relations, I recall those between angular momentum and angle, creation and annihilation operators (ruling generation and absorption of photons, respectively), number operator (telling us how many photons there are in a given context) and phase of the electromagnetic field. Some of them will be dealt with later on.

General Significance of Uncertainty Relations

Let us make some general considerations on the uncertainty relations. Many speak of an uncertainty "principle" instead of "relations"; actually, is not exactly a principle as such, given that its formulation derives from the more general formalism of the commutation relations between observables, and the resulting non-commutativity for many pairs, in turn, be a consequence of the use of operators for representing physical quantities due to the quantisation principle (Sect. 1.2.2). This amounts to say that uncertainty relations are deduced from the quantisation principle (as a development of Bohr's 1913 quantum postulate (Sect. 1.1.2)).

Having established the status of the uncertainty relations, let us consider what they tell us. The uncertainty relations tell us that, for every pair of (conjugated) observables, the product between both uncertainties cannot be less than the value given by their respective relations of uncertainty (in the case of relations (1.287) and (1.292) it is $\hbar/2$). This result means that we cannot get a perfect determination of, e.g. *both* position and momentum at the same instant. Every time we try to reduce the uncertainty in position, this is done at the expenses of the increase of the uncertainty in momentum, and vice versa. In other words, uncertainty relations establish the maximum of certainty possible as regards two conjugated observables. Conjugated observables, such as position and momentum, or time and energy, are important for the knowledge of the total state of the system: in CM, the state of a system is perfectly described by knowing both momentum and position or both energy and time. This means that, if we represent such a product as a square (where x and p_x are, for instance, base and height), its area is not less than the reduced Planck constant $\hbar$ divided by 2 (Fig. 1.25). It is quite obvious that this condition sets limitations on any operation we can perform: if e.g. by measuring we increase the determination of the position, what means that we reduce the uncertainty Δx, then we must have a simultaneous spread of p_x (an increase of its uncertainty Δp_x).

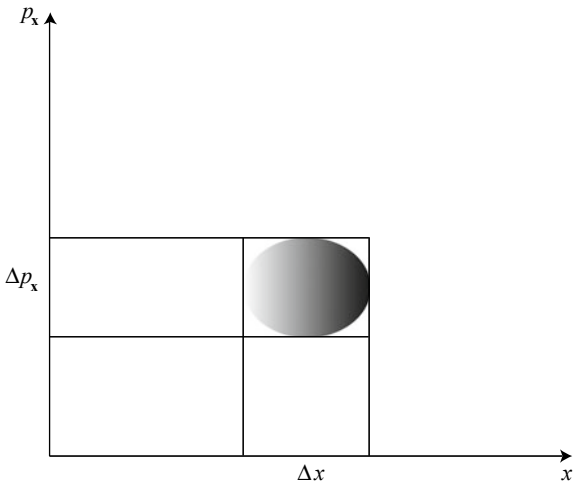


Fig. 1.25 Pictorial representation of the uncertainty relations position-momentum. The space whose axes are position and momentum is called *phase space*. The minimum of uncertainty can be approximated by an elliptical spot (whose gradation has no physical significance), whose minimal area $\pi(\Delta x/2) \cdot (\Delta p_x/2)$ is $h/16$. Adapted from Auletta et al. (2009, p. 88)

Heisenberg's Experiment

The so-called uncertainty principle was enunciated by Heisenberg in 1927 through a critical examination of the measurement operations on the observables of position and momentum.[103] For showing evidence for this uncertainty relation, Heisenberg designed an ideal experiment. Suppose we want to calculate the position of a small body, such as an electron revolving around the nucleus in a semi-classical model (that still dominated at that time). The light with which we try to illuminate the position of the body in order to detect it, will alter the quantity of motion of the latter in accordance with the quantisation principle. This was already proved by the photo-electric effect (Sect. 1.1.1) That is, being energy a discrete quantity, the alteration it will produce in the momentum of the body will be necessarily proportional to that energy and, therefore, there will always be a certain uncertainty in its momentum, as it is also displayed by Eq. (1.291). If we want to avoid the imprecision in the measurement of momentum, we will have to cast a light of a lesser frequency (energy) on the body, thereby obtaining an imprecision in the position due to the increase in the length of the wave (and so of the space interval that we are measuring). If we represent the degree of simultaneous determination of these two quantities by a product, it turns out that the latter can never be less than a quantity proportional to the Plank's constant or its reduced form, according to Eq. (1.287). In sum, it is not possible to determine with absolute precision the position and momentum of a physical corpuscle at one and the same time. From this example we can understand how the uncertainty relations are related to the quantisation principle, since it is the fact that the energy is quantised to determine uncertainty in the values of a pair of

[103] Heisenberg (1927). See also Cohen-Tannoudji (1991, pp. 60–61).

conjugated observables: as light is cast on a quantum system to be observed, we in fact alter it.

Its Interpretation

This raises the problem of the interpretation of the uncertainty relations, especially the question of whether the effects described are a consequence of our observation (ultimately a cognitive operation), or it is an ontological property of quantum systems (that is, a reality that is independent of observation). Heisenberg himself took the relations of uncertainty to be an experimental consequence of the process of measurement, what could be called an interactionist standpoint. In this sense, the uncertainty relations would be grounded on a type of disturbance which is brought about by the interaction between the measuring instrument and the system, interpretation which has been called "Heisenberg's assumption". This can be expressed by saying that the probabilistic or statistical character of QM comes from the observation and not from the definition provided by its basic equations.[104] This assumption has laid down the basis for the subsequent Kantian interpretations (after the name of the eighteenth century German philosopher and scientist Immanuel Kant (1724–1804)), like the Copenhagen interpretation, of QM.[105] In fact, from this assumption it follows that Heisenberg's relations would impose, at most, a reciprocal limitation on the precision with which we can establish the value of a pair of conjugated observables.

The opposite point of view is represented by those scholars who regarded this interpretation as too maximalist: they hypothesised that the uncertainty relations would by no means imply a basic limit to the precision of the measurement, but rather, Heisenberg's relations would simply describe the dispersion of the values obtained from these measurements. As for the values, the uncertainty relations would not hinder the simultaneous determination of the values of the conjugated observables, but, at most, it would point to a limitation in the possibility of a joint prediction of those values. These standpoints arose from a statistical approach to QM, according to which the object of this discipline would be statistical sets and not individual entities. Schrödinger proved that such a statistical interpretation is inadequate.[106] As a matter of fact, as we shall see, current experiments are done with single systems and nevertheless the uncertainty relations are always satisfied.

[104] Auletta (2000, pp. 124–27).

[105] It is interesting to point out the relationship that from the beginning was formed between QM and Kant's philosophy, highlighted in Heisenberg (1969, Chap. 10). On the more recent Kantian point of view see D'espagnat (1995).

[106] See Auletta (2000, pp. 107–115, 129–31) and quoted literature.

Interaction-Free Measurement

Also avoiding statistical interpretations, there is a problem in Heisenberg's assumption. If it was true, some phenomena would go unexplained, especially those that are related to the possibility of obtaining measurement results without interacting with the system, i.e. without directly affecting the system with the measuring instrument. This is the so-called *interaction-free measurement*, for the first time proposed by the German physicist Mauritius K. Renninger (1905–1987).[107] I shall come back later on the details of this experiment. By now, I wish to point out that this phenomenon could not be considered a disproof of Heisenberg's interpretation if it was not for the fact that, in spite of being an interaction-free measurement (a measurement without interaction), the relations of uncertainty remain still valid. In other words, the uncertainty relations hold true even when there is *no direct interaction* between the measuring instrument and the object system. With this state of things, the uncertainty relations cannot derive from a disturbance in the measurement, i.e. they cannot be an experimental consequence in general, but rather the formalism of the quantum theory states that a quantum system cannot possess precise definite values for e.g. its position and momentum simultaneously, and this *independently of any operation* that we perform or could perform on that system. Thus, they are the uncertainty relations to determine the value of the minimum disturbance acceptable and not vice versa. This is the first time that we meet a formal aspect of the theory that determines (set restrictions on) the way in which we make experience. Clearly, that formal aspect describes something (some kind of reality) that should be in fact responsible for that effect, although to find out what it is will occupy us across the next chapters.

Uncertainty Relations Affect the Definition of a System

In other words, uncertainty relations do not only affect a quantum-mechanical system's observation, but also its *definition*, in the sense that, at the opposite of the classical case, a full definition of the system in the sense of the *omnimoda determinatio* (perfect determination) is no longer possible, and, *as a consequence*, a precise measurement in the sense of the postulate to the reduction to zero of the measurement error is impossible as well (Sects. 1.1.2 and 1.2.1).[108] In fact, CM assumed that a state of a system is equivalent to a complete list of (determined) proprieties (the fact that these properties evolve also deterministically grounds the typical determinism of CM). Note that the violation of this in QM will bring us to the conclusion that we cannot take even a single property (associated with some determined eigenvalue of an observable) as a kind of ontological character of a physical system. We shall come back on this problem, but, from such an examination, we can already conclude that, given the system's intrinsic uncertainty, there cannot be any determined paths of the system neither in real space, nor in the phase space (the space whose coordinate axes

[107] Renninger (1960).

[108] The reader interested in this kind of problems may have a look at Auletta (2004).

are position and momentum). All we can draw is a set of spots which stand for the probability of presence, as displayed in Fig. 1.25.

Uncertainty and Superposition

A couple of additional words on the relation between the superposition and the quantisation principles (with its consequences, especially the uncertainty relations) can be appropriate. The superposition principle is a formal property of the system derived from the linearity of the Schrödinger equation. It is a property pervasive in the most specific aspects of the description of quantum systems, with immediate implications for the fully physical phenomena, even on a mesoscopic level. We shall consider this aspect below. The superposition principle tells us that if two states of a system are possible, any linear combination of both is also possible; the uncertainty relations, on the other hand, establish a minimum value of uncertainty which is not present in the first one at all. In fact, there are situations in which, despite the fact that there is no superposition, the relations of uncertainty still remain valid, such as the lowest level state of the harmonic oscillator. In this case, superposition would specify the possibility that a system is in any linear combination of its possible energy states; whereas the relations of uncertainty refer to each of the states in particular. They are, then, two independent formal tools as they have different contents and range of application.

1.2.5 Unitary Transformations and Symmetries

Unitary Transformations

I recall that position, time and angle are called kinematic observables, while momentum, energy, and angular momentum are dynamical observables acting 'upon' their kinematic counterpart. The *Noether's theorem* after the name of the German mathematician and physicist A. Emmy Noether (1882–1935), tells us in fact that changes of a kinematic parameter are ruled by a dynamical quantity that is the generator of that transformation.[109]

All of the physical transformations mentioned so far can be represented by unitary operators.[110] We have already found the form for the energy observable (Eq. (1.16)). Moreover, the space translation operator $e^{a\hat{D}_x}$ in Eq. (1.135) can be written in terms of the momentum operator (1.139) as

$$\hat{U}_x(a) = e^{-\frac{i}{\hbar}a\hat{p}_x}, \tag{1.293}$$

[109] Arnold (1978, pp. 88–89), Auletta et al. (2009, Chap. 8).

[110] On unitary operators see Auletta et al. (2009, Chaps. 3 and 8).

which is also a unitary operator. As a result, the momentum eigenstates $|\, p_x \rangle$ are also eigenstates of $\hat{U}_x(a)$, namely, we have

$$\hat{U}_x(a)\,|\, p_x \rangle = e^{-\frac{i}{\hbar} p_x a}\,|\, p_x \rangle \ . \tag{1.294}$$

Let us consider the action of $\hat{U}_x(a)$ on the position eigenstate $|\, x \rangle$. By inserting the identity operator (1.145) between $\hat{U}_x(a)$ and $|\, x \rangle$, we obtain

$$\begin{aligned}
\hat{U}_x(a)\,|\, x \rangle &= \int_{-\infty}^{+\infty} \hat{U}_x(a)\,|\, p_x \rangle\, \langle p_x \,|\, x \rangle\, dp_x \\
&= \int_{-\infty}^{+\infty} \langle p_x \,|\, x \rangle\, \hat{U}_x(a)\,|\, p_x \rangle\, dp_x \\
&= \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} e^{-\frac{i}{\hbar} p_x x}\, e^{-\frac{i}{\hbar} p_x a}\,|\, p_x \rangle\, dp_x \\
&= \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} e^{-\frac{i}{\hbar} p_x (x+a)}\,|\, p_x \rangle\, dp_x \\
&= \int_{-\infty}^{+\infty} |\, p_x \rangle\, \langle p_x \,|\, x + a \rangle\, dp_x \\
&= |\, x + a \rangle \ ,
\end{aligned} \tag{1.295}$$

where use has been made of Eqs. (1.148) and (1.294). This proves that, under a space translation by a distance a, the state $|\, x \rangle$ transforms into the state $|\, x + a \rangle$. The corresponding 3D translation transformation can be written as

$$\begin{aligned}
\hat{U}_{\mathbf{r}}(\mathbf{a}) &= \hat{U}(a_x, a_y, a_z) \\
&= e^{-\frac{i}{\hbar} \hat{p}_x a_x} e^{-\frac{i}{\hbar} \hat{p}_y a_y} e^{-\frac{i}{\hbar} \hat{p}_z a_z} \\
&= e^{-\frac{i}{\hbar} \hat{\mathbf{p}} \cdot \mathbf{a}} \ .
\end{aligned} \tag{1.296}$$

In a similar way, it can be shown that the 3D momentum translations by a vectorial amount $\mathbf{v}$ are described by the unitary transformation

$$\hat{U}_{\mathbf{p}}(\mathbf{v}) = e^{\frac{i}{\hbar} \mathbf{v} \cdot \hat{\boldsymbol{r}}} \tag{1.297}$$

or

$$\hat{U}_{\mathbf{p}}(\mathbf{v}) \hat{\mathbf{p}} \hat{U}_{\mathbf{p}}^{\dagger}(\mathbf{v}) = \hat{\mathbf{p}} - \mathbf{v} \ , \tag{1.298}$$

where $\mathbf{r} = (x, y, z)$ and $\mathbf{v} = v_x, v_y, v_z$ are the 3D position and velocity, respectively. Similarly, we can write the rotation unitary operator for the total angular momentum, in the case of a rotation $\hat{\mathbf{R}}$ of an angle θ about an axis $\mathbf{n}$,

$$\hat{U}_{\mathbf{R}}(\theta) = e^{-\frac{i}{\hbar} \boldsymbol{\theta} \mathbf{n} \cdot (\hat{\mathbf{L}} + \hat{\mathbf{S}})} = e^{-\frac{i}{\hbar} \boldsymbol{\theta} \mathbf{n} \cdot \hat{\mathbf{J}}} \ . \tag{1.299}$$

The total angular momentum $\hat{\mathbf{J}}$ has a orbital component $\hat{\mathbf{L}}$ and a spin component $\hat{\mathbf{S}}$ that will be introduced later.

Stone Theorem and Groups

Note that the previous equations as well as Eq. (1.16) are particular instances of the general equation for unitary operators[111]

$$\hat{U} = e^{ia\hat{O}} = 1 + ia\hat{O} + \frac{1}{2!}(ia\hat{O})^2 + \dots \,, \tag{1.300}$$

which is akin to Eq. (1.137), where $\hat{O}$ is a Hermitian operator and $0 \leq a \in \mathbb{R}$ is a parameter, satisfying the semigroup property

$$\hat{U}(a + a') = \hat{U}(a)\hat{U}(a') \,. \tag{1.301}$$

This is known as the *Stone's theorem*, after the name of the American mathematician Marshall H. Stone (1903–1989).[112] The semigroup property can be easily proved for spatial translations. It suffices to take the product $\hat{U}_x(a')\hat{U}_x(a)$ and first perform the derivation (1.295) with $\hat{U}_x(a)$ and reiterate the same procedure on state $|x + a\rangle$ with $\hat{U}_x(a')$, so to get $\left|x + a + a'\right\rangle$, what shows that

$$\hat{U}_x(a')\hat{U}_x(a)\,|x\rangle = \hat{U}_x(a + a')\,|x\rangle \,. \tag{1.302}$$

Being all the transformations introduced so far unitary, they preserve the scalar product and are reversible. In other words, they form groups. A *group* $\mathcal{G}$ is a set of elements such that there is an operation $\star$ (called group multiplication) satisfying following requirements[113]:

- For any elements $a, b \in \mathcal{G}$, also $a \star b \in \mathcal{G}$ (in other words $\mathcal{G}$ is closed under group multiplication),
- Associativity holds: $\forall a, b, c \in \mathcal{G}$, we have $a \star (b \star c) = (a \star b) \star c$,
- There is a unit element e such that $\forall a, e \star a = a \star e = a$,
- $\forall a$ there is an inverse a^{-1} such that $a \star a^{-1} = a^{-1}a = e$.

Clearly, unitary transformations involve symmetries. For instance, the rotational symmetry related to the generator of rotations. The notion of *symmetry* is thus strictly connected with that of invariance (under some transformation). However, the notions of symmetry and group are not identical. In fact, "for a classical dynamical system

[111] See Landsman (2017, Sect. 5.12).

[112] Also J. von Neumann contributed to prove this theorem.

[113] Tinkham (1964, p. 6). This is a good textbook for knowing more about group theory. See also Byron and Fuller (1969–70, II, Chap. 10). The reader interested in the notion of symmetry may also have a look at Fuchs and Schweigert (1997).

the symmetries can be viewed as acting on the points of the configuration space; hence they form a group, with the group multiplication provided by the composition of maps. In contrast, in quantum physics the configuration space is no longer present, so that this argument does not apply any more".[114] This character of QM is due to the non-commutativity of observables. Note that a *representation* of a group is embedding elements of the group into operators that act on a vector space, like the examples previously considered.

Note that dealing here with continuity, the group of quantum transformations is a *Lie algebra*, after the name of Sophus Lie, which provides us of a complete information about the local structure of the group thanks to *infinitesimal group elements*, although it may not provide the global structure of the group.[115] In particular, in the two previous subsections, I have in fact shown this already for momentum as generator of space translations and orbital angular momentum as generator of rotations.

Note also that some of the most important groups are the unitary group $\mathrm{U}(n)$ for $n \times n$ unitary matrices, of which we have considered some cases, the special unitary group $\mathrm{SU}(n)$ with matrices with determinant $= 1$, the permutation group (or Galois group, after the name of the French mathematician Évariste Galois (1811–1832)) $\mathrm{P}(n)$, the orthogonal group $\mathrm{O}(n)$ that is the group of distance-preserving transformations of a nD Euclidean space that preserve a fixed point (the matrices (1.226)–(1.227) constitute the generators of the groups O(2) and O(3), respectively), the special orthogonal or rotation group $\mathrm{SO}(n)$ with determinant $= 1$, of which SO(3) or, equivalently, SU(2), is the group of rotations in 3D Euclidean space, especially relevant for angular momentum.

Time Translation and Time Reversal

As seen, time translations are ruled by unitary operators of the form (1.16).[116] Introducing its adjoint

$$\hat{U}_t^\dagger(\tau) = \left(e^{-\frac{i}{\hbar}\hat{H}\tau}\right)^\dagger = e^{+\frac{i}{\hbar}\hat{H}\tau}, \tag{1.303}$$

where $\tau = t - t_0$ for some initial time t_0, we have that the scalar product is preserved:

$$\langle\psi(t) \mid \psi(t)\rangle = \left\langle\psi(t_0)\left|\hat{U}_t^\dagger(\tau)\hat{U}_t(\tau)\right|\psi(t_0)\right\rangle = \langle\psi(t_0) \mid \psi(t_0)\rangle = 1, \tag{1.304}$$

where I have used Eq. (1.15) and the unitarity property

$$\hat{U}_t^\dagger(\tau)\hat{U}_t(\tau) = \hat{U}_t(\tau)\hat{U}_t^\dagger(\tau) = \hat{I}. \tag{1.305}$$

---

[114] Fuchs and Schweigert (1997, p. 11).

[115] Penrose (2004, Sect. 13.5).

[116] On what follows see Auletta et al. (2009, Sects. 3.5.3 and 8.3).

Moreover, multiplying both sides of Eq. (1.15) by

$$\hat{U}_t^{\dagger}(\tau) := \hat{U}_{-t}(\tau) = \hat{U}_t(-\tau)\,, \tag{1.306}$$

and using the unitarity property of $\hat{U}_t(\tau)$, we obtain

$$\hat{U}_t(-\tau)\,|\,\psi(t)\rangle\, = |\,\psi(t_0)\rangle\,, \tag{1.307}$$

showing that time evolution in QM is in fact reversible. In other words, $\hat{U}_t(\tau)$ and $\hat{U}_t^{\dagger}(\tau)$ act as the forward and backward time evolution operators, respectively, bringing a system from its initial state to its final state and vice versa. This shows, as expected, that

$$\hat{U}_t(-\tau) = \hat{U}_t^{-1}(\tau), \tag{1.308}$$

i.e. $\hat{U}_t(-\tau)$ is the inverse of $\hat{U}_t(\tau)$.

Time reversal (i.e. the mapping $t \longmapsto -t$) is anti-unitary, where an anti-unitary mapping like $\tilde{\hat{U}}$ satisfies the following properties

$$\begin{aligned}
\left\langle \tilde{\hat{U}}\psi \,|\, \tilde{\hat{U}}\varphi \right\rangle &= \langle \varphi \,|\, \psi \rangle\,, \\
\tilde{\hat{U}}[\,|\,\psi\rangle + |\,\varphi\rangle\,] &= \tilde{\hat{U}}\,|\,\psi\rangle + \tilde{\hat{U}}\,|\,\varphi\rangle\,, \\
\tilde{\hat{U}}\,c\,|\,\psi\rangle &= c^*\tilde{\hat{U}}\,|\,\psi\rangle\,.
\end{aligned} \tag{1.309}$$

Note that the latter property together with the second one determines that this transformation is also anti-linear. Both unitary and anti-unitary transformations preserve the norm and the scalar product, as it is evident from Eq. (1.304). Thus, we have the anti-unitary mapping

$$\tilde{\hat{U}}_{-t}(t)\,|\,\psi(t)\rangle\, = |\,\psi(-t)\rangle\,, \tag{1.310}$$

while the temporal relation between two time-reversed states should be unitary according to

$$|\,\psi(-t_1)\rangle\, = \hat{U}_t(\tau)\,|\,\psi(-t_2)\rangle\,, \tag{1.311}$$

where $\tau = t_2 - t_1$, $t_2 > t_1$, and $|\,\psi(-t_1)\rangle\, = \tilde{\hat{U}}_{-t}(t_1)\,|\,\psi(t_1)\rangle\,$, $|\,\psi(-t_2)\rangle\, = \tilde{\hat{U}}_{-t}(t_2)\,|\,\psi(t_2)\rangle$. This allows us to write the scalar product

$$\langle \psi(t_1) \,|\, \psi(-t_1)\rangle = \left\langle \psi(t_2)\hat{U}_t^{\dagger}(\tau) \,|\, \hat{U}_t(\tau)\psi(-t_2) \right\rangle. \tag{1.312}$$

All quantum transformations that satisfy a symmetry are either unitary and linear (for continuous transformations) or anti-unitary and anti-linear (for discrete transformations), and all of them preserve the scalar product: this is the so-called *Wigner's*

theorem after the name of the Hungarian mathematician and physicist Eugene Wigner (1902–1995).[117]

Parity and CPT Theorem

Time reversal is a discrete symmetry, but there are also other examples, represented by space reflection or parity (which in fact inverts the sign of the argument of a wave function):

$$\hat{U}_R \psi(x) = \psi(-x), \tag{1.313}$$

where $\hat{U}_R$ is the space reflection unitary operator. In other words, parity refers to the symmetry between right and left. Note that two space reflections will bring back to the initial state. In the following, I shall denote the parity unitary operator for both position and momentum by $\hat{\Pi}$.

Also charge conjugation, which interchanges particles and antiparticles (in fact, the anti-electron, called positron, has positive charge), is unitary but antisymmetric. Note that there is a fundamental theorem of quantum field theory, called the *CPT theorem* that tells us that the product of charge conjugation, parity, and time reversal is always conserved although none of these may be individually conserved. A further discrete symmetry is permutation, on which we shall come back.

Degeneracy of Eigenvalues

Energy eigenvalues can be degenerate, i.e. it can be the case that two or more eigenvectors (or eigenfunctions) share the same eigenvalue.[118] This has a relevant physical meaning. We can establish a necessary and sufficient condition in order to have energy degeneracy: commutativity. For the sake of the argument, we assume that the observable $\hat{O}$ commutes with the Hamiltonian, i.e.$[\hat{H}, \hat{O}] = 0$, but is not a function of $\hat{H}$ only and is not the identity operator. Then, using Eqs. (1.78) and (1.197), we obtain

$$\hat{H}\hat{O}\left|\psi_j\right\rangle = \hat{O}\hat{H}\left|\psi_j\right\rangle = E_j\hat{O}\left|\psi_j\right\rangle = E_j\left|\hat{O}\psi_j\right\rangle \ , \tag{1.314}$$

which shows that the ket $\left|\hat{O}\psi_j\right\rangle$ is an eigenket of $\hat{H}$ with eigenvalue E_j just as $\left|\psi_j\right\rangle$ is. Moreover, the ket $\left|\hat{O}\psi_j\right\rangle$ cannot be proportional to $\left|\psi_j\right\rangle$, i.e.it cannot be written in the form $\hat{O}\left|\psi_j\right\rangle = f(E_j)\left|\psi_j\right\rangle$, where $f(E_j)$ is a function of the jth energy eigenvalue, since in such a case $\hat{O}$ would be equal to $f(\hat{H})$, contradicting one of the assumptions.

---

[117] Wigner (1959).

[118] On this subject see Auletta et al. (2009, Sect. 3.1.4).

Let us now prove the equivalence in the other sense, that is, that degeneracy is sufficient condition of commutativity. If we have degeneracy, we may always partition the whole Hilbert space $\mathcal{H}$ into subspaces $\mathcal{H}_1, \mathcal{H}_2, \ldots$, in each of which the energy eigenvalue is *constant*, i.e.

$$\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \ldots \ , \tag{1.315}$$

where the symbol $\oplus$ denotes the direct sum: e.g. $\mathbb{R} \oplus \mathbb{R} = \mathbb{R}^2$ (this sum represents the Cartesian plane). This is technically true only if the spectrum is discrete, while in the case of a continuous spectrum, a similar formula holds, but it is more complex. Limiting here our considerations to the discrete case, we have

$$\hat{H} \, | \psi \rangle \ = E_k \, | \psi \rangle \ \ , \tag{1.316}$$

$\forall \, | \psi \rangle \ \in \mathcal{H}_k$. Indeed, suppose that $| \psi \rangle$ is an arbitrary superposition of energy eigenstates $| \psi_n \rangle$. Then, we have

$$\hat{H} \, | \psi \rangle \ = \hat{H} \left(\sum_{n=1}^{k} c_n \, | \psi_n \rangle \right) = \sum_{n=1}^{k} c_n E_k \, | \psi_n \rangle \ = E_k \, | \psi \rangle \ . \tag{1.317}$$

This means that the Hamiltonian $\hat{H}_k$ in each subspace $\mathcal{H}_k$ is a multiple of the identity. Here, $\hat{H}_k$ is the block of $\hat{H}$ pertaining to $\mathcal{H}_k$, i.e.

$$\hat{H}_k = \begin{bmatrix} E_k & 0 & 0 & \ldots \\ 0 & E_k & 0 & \ldots \\ & & \ddots & \\ 0 & \ldots & \ldots & E_k \end{bmatrix} \ , \tag{1.318}$$

where the dimension of $\mathcal{H}_k$ is equal to the degree of degeneracy of the eigenvalue E_k. Now, if there are degeneracies, since the Hamiltonian, restricted to the space of the degenerate eigenvectors, will be a multiple of the identity, any operator in that space would commute with it. Thus, we can then build a Hermitian operator

$$\hat{O} = \hat{O}_1 \otimes \hat{O}_2 \otimes \ldots \ , \tag{1.319}$$

where $\otimes$ is called the *tensor* or *direct product* (on which we shall come back) between operators belonging to the different subspaces (in which the energy eigenvalue is a constant) of the total Hilbert space, and $\hat{O}_k$ is the operator that takes a vector on $\mathcal{H}_k$ into a vector of $\mathcal{H}_k$, and obviously commutes with $\hat{H}_k$. This is sufficient condition for $[\hat{O}, \hat{H}] = 0$.

Let us now formulate some general considerations about the degeneracy of the energy eigenvalues. In order to have energy degeneracy, there must be a conserved quantity in addition to energy, and all dynamic observables (energy, momentum,

angular momentum) so far considered, being associated with the generators of the respective group, must represent conserved quantities. They form continuous symmetries (the generators are represented by differential operators). This amounts to say that any of them must commute with the Hamiltonian (see also Eq. (1.3)). In other words, for any generator $\hat{O}_G$ of a quantum group, we have (see also Eq. (1.216))

$$\left[\hat{O}_G, \hat{H}\right] = 0. \tag{1.320}$$

This is also known as the already mentioned *Noether's theorem*.[119] A very interesting case is when the conserved quantity exists classically. In systems with a central potential (like the solar system or the hydrogen atom), the angular momentum is conserved and states with non-zero angular momentum form a multiplet that are degenerate in energy. That is, a state with angular momentum l belongs to a multiplet where there are $2l + 1$ states. More complex cases are possible if there are additional conserved observables. For a two-dimensional rotation, for instance, when there is an additional constant of motion, the classical orbits of the finite motion must be closed (e.g. ellipses, as in the Kepler's orbital problem, after the name of the German mathematician, astronomer, and astrologer Johannes Kepler (1571–1630)), as opposed to open orbits in absence of conservation (see also Fig. 1.5, Sect. 1.1.3).

1.3 Some Further Principles

1.3.1 Pauli's Exclusion Principle and Spin

Discovery of the Spin

Pauli's exclusion principle, after the name of the Austrian–Swiss physicist Wolfgang Pauli (1900–1958), was formulated in 1925 and so baptised in 1926 by P. Dirac. It has its roots in what would later be the discovery of the magnetic moment (called spin), a degree of freedom characterising quantum particles like electrons: it is an amount of magnetic force related to a sort of motion of the electron around its own axis (from here the term *spin*). Thus, spin has the same physical dimensions of angular momentum (Sect. 1.2.4), and in fact the sum of the two constitutes the total orbital momentum.

In the famous experiment performed in 1921–1922 by the German experimental physicists Walther Gerlach (1889–1979) and Otto Stern (1888–1969) it was found out, as identical silver atoms were tunnelled through an oriented magnetic field of a constant gradient, that those atoms took only two orientations of what was later called spin, i.e. *spin up* or *spin down* (Fig. 1.26).[120] Thus, another aspect of quanti-

[119] On the subject see Auletta et al. (2009, Sect. 8.1).

[120] Gerlach and Stern (1922a, b, c). For historical reconstruction see Jammer (1966, Sect. 3.4).

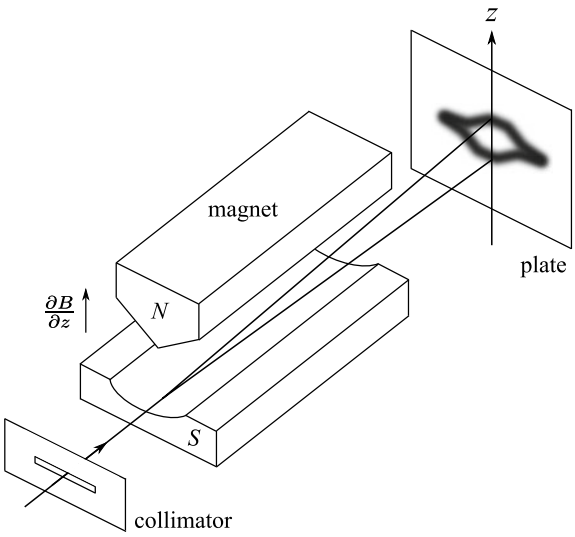


Fig. 1.26 Scheme of the Stern–Gerlach experiment. A beam of silver atoms is sent through a strongly inhomogeneous magnetic field **B** with the gradient $\partial \mathbf{B}/\partial z$ along the z direction, which is perpendicular to the beam axis. Due to quantisation of spin angular momentum and associated magnetic moment of the electron, the beam is separated into two distinct parts on the collector plate (note that there is some interference between the upper and lower parts). Adapted from Auletta and Wang (2014, p. 236)

sation arose. The presence of this intrinsic magnetic momentum makes the general description of the state of a particle as having to supply not only e.g. the probability of the different positions in the space of the particle, but also the probability of the different orientations of this spin, which means that this description should include a discrete variable, which would account for either orientation.

Eigenstates of the Spin

The spin is an important (intrinsic) degree of freedom of elementary particles and it enters in a huge number of experimental and theoretical problems.[121] It is characterised by a *spin quantum number* $s \in \mathbb{Q}^+$ (where $\mathbb{Q}^+$ is the set of non-negative rational numbers) that takes different values according to the involved particles. In the easiest case of fermions like electrons, with $s = \frac{1}{2}$, it is convenient to write the spin eigenstates along the z direction (which is customarily used) as $|\uparrow_z\rangle$ and $|\downarrow_z\rangle$, so that, by introducing the spin observable $\hat{S}_z$ along the z direction, we have, as an instance of the eigenvalue equation (1.118),

[121] I follow here Auletta and Wang (2014, Sect. 8.6), which gives a quite comprehensive account of this subject.

$$\hat{S}_z\,|\uparrow_z\rangle = \frac{\hbar}{2}\,|\uparrow_z\rangle\,, \quad \text{and} \quad \hat{S}_z\,|\downarrow_z\rangle = -\frac{\hbar}{2}\,|\downarrow_z\rangle\,. \tag{1.321}$$

Note that, as for the orbital angular momentum, the presence of the reduced Planck constant in the eigenvalue is due to quantisation. The projection of the spin along the z direction is called *spin magnetic quantum number*, is denoted by m_s, and, according to the result following Eq. (1.248), takes $2s + 1$ possible values (which explain the two values for $s = 1/2$). The above two eigenstates are usually referred to as the *spin up* and *spin down* states in the z direction, respectively. They also constitute an orthonormal basis of the 2D Hilbert space, called the spin space, and thus, in analogy with expressions (1.54), can be defined as

$$|\uparrow_z\rangle := \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad |\downarrow_z\rangle := \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{1.322}$$

In the following, I shall often omit the subscript referring to the direction whenever no confusion may arise. In the basis $\{|\uparrow_z\rangle\,, |\downarrow_z\rangle\}$, the operators $\hat{S}_x$, $\hat{S}_y$, and $\hat{S}_z$ in the three Cartesian directions are represented by

$$\hat{S}_x = \frac{\hbar}{2}\hat{\sigma}_x, \quad \hat{S}_y = \frac{\hbar}{2}\hat{\sigma}_y, \quad \hat{S}_z = \frac{\hbar}{2}\hat{\sigma}_z, \tag{1.323}$$

where the matrices $\hat{\sigma}_x, \hat{\sigma}_y$, and $\hat{\sigma}_z$ are called the *Pauli matrices* (or Pauli operators)[122]:

$$\hat{\sigma}_x := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \hat{\sigma}_y := \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \hat{\sigma}_z := \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{1.324}$$

They are both Hermitian and unitary, and therefore satisfy following properties:

- $\hat{\sigma}_x^2 = \hat{\sigma}_y^2 = \hat{\sigma}_z^2 = \hat{I}$, (1.325a)
- $\hat{\sigma}_j\hat{\sigma}_k + \hat{\sigma}_k\hat{\sigma}_j = 2\delta_{jk}\hat{I}$, (1.325b)
- $[\hat{\sigma}_x, \hat{\sigma}_y] = 2i\hat{\sigma}_z, \quad [\hat{\sigma}_y, \hat{\sigma}_z] = 2i\hat{\sigma}_x, \quad [\hat{\sigma}_z, \hat{\sigma}_x] = 2i\hat{\sigma}_y,$ (1.325c)

where $j, k = x, y, z$ and $\hat{I}$ is the 2×2 identity matrix (1.93). The latter row displays the spin commutation relations. Note that the Pauli spin matrices together with the identity operator constitute a complete set of observables so that any 2×2 matrix $\hat{O}$ can be expressed as a combination of them:

$$\hat{O} = c_0\hat{I} + c_1\hat{\sigma}_x + c_2\hat{\sigma}_y + c_3\hat{\sigma}_z = \begin{bmatrix} c_0 + c_3 & c_1 - ic_2 \\ c_1 + ic_2 & c_0 - c_3 \end{bmatrix}, \tag{1.326}$$

where the c_j's ($j = 0, 1, 2, 3$) are some coefficients $\in \mathbb{C}$. Thus, Pauli matrices constitute the group SU(2): in fact, all the traces (the sum of the diagonal elements of

[122] Pauli (1927).

Table 1.2 The actions of the Pauli matrices on the spin up and spin down states in the z direction

$\hat{\sigma}_x \left\vert\uparrow_z\right\rangle = \left\vert\downarrow_z\right\rangle$	$\hat{\sigma}_x \left\vert\downarrow_z\right\rangle = \left\vert\uparrow_z\right\rangle$
$\hat{\sigma}_y \left\vert\uparrow_z\right\rangle = i \left\vert\downarrow_z\right\rangle$	$\hat{\sigma}_y \left\vert\downarrow_z\right\rangle = -i \left\vert\uparrow_z\right\rangle$
$\hat{\sigma}_z \left\vert\uparrow_z\right\rangle = \left\vert\uparrow_z\right\rangle$	$\hat{\sigma}_z \left\vert\downarrow_z\right\rangle = - \left\vert\downarrow_z\right\rangle$

a matrix) are 0, which is equivalent to O(3) (Sect. 1.2.5), since, as we shall see, the Pauli matrices enable us to describe rotations about the x, y, z axes by π radians.

The action of the Pauli spin matrices on the eigenstates (1.322) can be computed by directly applying them to the latter as displayed in Table 1.2. The eigenvectors and eigenvalues of $\hat{\sigma}_x$ and $\hat{\sigma}_y$ are given by the following equations:

$$\hat{\sigma}_x \left\vert\uparrow_x\right\rangle = \left\vert\uparrow_x\right\rangle \ , \ \hat{\sigma}_x \left\vert\downarrow_x\right\rangle = - \left\vert\downarrow_x\right\rangle \ , \tag{1.327a}$$

$$\hat{\sigma}_y \left\vert\uparrow_y\right\rangle = \left\vert\uparrow_y\right\rangle \ , \ \hat{\sigma}_y \left\vert\downarrow\right\rangle_y = - \left\vert\downarrow\right\rangle_y \ , \tag{1.327b}$$

where

$$\left\vert\uparrow_x\right\rangle = \frac{1}{\sqrt{2}} \left(\left\vert\uparrow_z\right\rangle + \left\vert\downarrow_z\right\rangle \right) \ , \ \left\vert\downarrow_x\right\rangle = \frac{1}{\sqrt{2}} \left(\left\vert\uparrow_z\right\rangle - \left\vert\downarrow_z\right\rangle \right) \ , \tag{1.328a}$$

$$\left\vert\uparrow_y\right\rangle = \frac{1}{\sqrt{2}} \left(\left\vert\uparrow_z\right\rangle + i \left\vert\downarrow_z\right\rangle \right) \ , \ \left\vert\downarrow\right\rangle_y = \frac{1}{\sqrt{2}} \left(\left\vert\uparrow_z\right\rangle - i \left\vert\downarrow_z\right\rangle \right) \ . \tag{1.328b}$$

In other words, the eigenkets of the spin in the x and y directions are (different) superpositions of the eigenkets of the spin in the z direction. Note that in whatever direction, the Pauli matrices act on their own eigenstates in the same way, i.e. giving eigenvalues ± 1 for up and down states.

As mentioned, spin represents an additional and internal degree of freedom that is independent of the external degrees of freedom like position and time. Then, as mentioned, the 'usual' wave function has to be multiplied by a spin part, which, in the case of spin-1/2, has two components in order to account for the two possible values of the spin. The resulting function is called *spinor* (a concept on which I shall come back) and will be indicated by

$$\psi(\mathbf{r}, s; t) = \begin{pmatrix} \psi_\uparrow(\mathbf{r}; t) \\ \psi_\downarrow(\mathbf{r}; t) \end{pmatrix} \ . \tag{1.329}$$

In other words, $\psi_\uparrow(\mathbf{r})$ is the eigenfunction of the spin projection along z with eigenvalue $+1/2$ and $\psi_\downarrow(\mathbf{r})$ the eigenfunction corresponding to the eigenvalue $-1/2$. Then, the total wave function can be written as the "factorised superposition"

$$\psi(\mathbf{r}, s; t) = \begin{pmatrix} \psi_\uparrow(\mathbf{r}; t) \\ \psi_\downarrow(\mathbf{r}; t) \end{pmatrix} = \psi_\uparrow(\mathbf{r}; t) \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \psi_\downarrow(\mathbf{r}; t) \begin{pmatrix} 0 \\ 1 \end{pmatrix} \ . \tag{1.330}$$

For different spin numbers we clearly have different matrices defining the spin observable. I give here the example of spin number $s = 1$, with the following Hilbert-space 3D spin matrices:

$$\hat{S}_x = \frac{\hbar}{\sqrt{2}} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad \hat{S}_y = \frac{\hbar}{\sqrt{2}} \begin{bmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{bmatrix}, \quad \hat{S}_z = \hbar \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}. \tag{1.331}$$

They can be lead to the group SU(3) (all traces are in fact again 0).

Lie Algebra of Spin

Just as the orbital angular momentum is the generator of rotations in real space, the spin angular momentum is the generator of rotation in spin space. Taking into account Eq. (1.299), the rotation operator in spin space about a direction $\mathbf{n}$ (which is a unit vector) by an angle θ is given by the unitary operator

$$\hat{U}_{\mathbf{n}}^{(s)}(\theta) = e^{-\frac{i}{\hbar}\theta \mathbf{n}\cdot\hat{\mathbf{S}}}, \tag{1.332}$$

where the superscript s denotes the spin quantum number s and

$$\mathbf{n} \cdot \hat{\mathbf{S}} = n_x \hat{S}_x + n_y \hat{S}_y + n_z \hat{S}_z, \tag{1.333}$$

with $n_x^2 + n_y^2 + n_z^2 = 1$. Making use of Eq. (1.323), for the spin $\frac{1}{2}$ case we have

$$\begin{aligned} \hat{U}_{\mathbf{n}}^{(1/2)}(\theta) &= e^{-\frac{i}{2}\theta \mathbf{n}\cdot\hat{\boldsymbol{\sigma}}} \\ &= \cos\frac{\theta}{2}\hat{I} - i \sin\frac{\theta}{2}\mathbf{n} \cdot \hat{\boldsymbol{\sigma}}, \end{aligned} \tag{1.334}$$

where $\hat{\boldsymbol{\sigma}} = (\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z)$ and

$$\begin{aligned} \mathbf{n} \cdot \hat{\boldsymbol{\sigma}} &= n_x \hat{\sigma}_x + n_y \hat{\sigma}_y + n_z \hat{\sigma}_z \\ &= \begin{bmatrix} n_z & n_x - i n_y \\ n_x + i n_y & -n_z \end{bmatrix}. \end{aligned} \tag{1.335}$$

In obtaining the second equality in Eq. (1.334), use has been made of Euler formulas (1.181) and Eq. (1.326), with $c_0 = \cos(\theta/2)$ and $c_j = -i \sin(\theta/2) n_j$ $(j = x, y, z)$. In particular, for a rotation of angle θ about the z axis we have

$$\hat{U}_z^{(1/2)}(\theta) = \cos\frac{\theta}{2}\hat{I} - i\sin\frac{\theta}{2}\hat{\sigma}_z$$
$$= \begin{bmatrix} \cos\frac{\theta}{2} - i\sin\frac{\theta}{2} & 0 \\ 0 & \cos\frac{\theta}{2} + i\sin\frac{\theta}{2} \end{bmatrix}$$
$$= \begin{bmatrix} e^{-\frac{i}{2}\theta} & 0 \\ 0 & e^{\frac{i}{2}\theta} \end{bmatrix}, \tag{1.336}$$

while for rotations about the x and y axes we have

$$\hat{U}_x^{(1/2)}(\theta) = \cos\frac{\theta}{2}\hat{I} - i\sin\frac{\theta}{2}\hat{\sigma}_x = \begin{bmatrix} \cos\frac{\theta}{2} & -i\sin\frac{\theta}{2} \\ -i\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{bmatrix}, \tag{1.337a}$$

$$\hat{U}_y^{(1/2)}(\theta) = \cos\frac{\theta}{2}\hat{I} - i\sin\frac{\theta}{2}\hat{\sigma}_y = \begin{bmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{bmatrix}. \tag{1.337b}$$

Note that the latter matrix coincides with (1.226). Moreover, for $\theta = \pi$ we find from Eq. (1.334) that

$$\hat{U}_x^{(1/2)}(\pi) = -i\,\hat{\sigma}_x, \quad \hat{U}_y^{(1/2)}(\pi) = -i\,\hat{\sigma}_y, \quad \hat{U}_z^{(1/2)}(\pi) = -i\,\hat{\sigma}_z. \tag{1.338}$$

Therefore, as anticipated, for spin $\frac{1}{2}$, up to a global phase factor the Pauli matrices $\hat{\sigma}_x$, $\hat{\sigma}_y$, and $\hat{\sigma}_z$ are the respective rotation operators by an angle π about the x, y, and z axis.

Total Angular Momentum

As recalled, the spin operator $\hat{\mathbf{S}}$ and the orbital angular momentum $\hat{\mathbf{L}}$ (Eqs. (1.229)), having the same physical dimensions, give rise to a total angular momentum given by (Eq. (1.299))

$$\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{S}}\,. \tag{1.339}$$

which, as it happens for $\hat{\mathbf{S}}$, has the same commutation relations as $\hat{\mathbf{L}}$ (see Eq. (1.234)), that are

$$\left[\hat{S}_k, \hat{S}_n\right] = i\hbar\epsilon_{knr}\hat{S}_r, \quad \left[\hat{J}_k, \hat{J}_n\right] = i\hbar\epsilon_{knr}\hat{J}_r\,. \tag{1.340}$$

In fact, the total angular momentum's Hilbert–space 3D matrices are identical to the corresponding spin matrices. Note that the values j of the total angular momentum are a combination of orbital angular momentum l values and spin s values.[123]

[123] Auletta et al. (2009, Sect. 6.4.4).

Addition of Angular Momenta

Very often we need to consider the total orbital angular momentum $\hat{\mathbf{L}}_1 + \hat{\mathbf{L}}_2$ of a system of two particles (1 and 2), the total spin angular momentum $\hat{\mathbf{S}}_1 + \hat{\mathbf{S}}_2$ of two particles (1 and 2), or the total angular momentum $\hat{\mathbf{L}} + \hat{\mathbf{S}}$ of a particle with both spin and orbital angular momenta. In these situations we shall deal with the problem of *how to add two angular momenta*. The formalism of angular momentum addition is very general, so we will follow the usual convention to denote a generic angular momentum by $\hat{\mathbf{J}}$, which could be orbital ($\hat{\mathbf{L}}$), spin ($\hat{\mathbf{S}}$), or some combined quantity (e.g., $\hat{\mathbf{L}} + \hat{\mathbf{S}}$). Consider the addition of two angular momenta $\hat{\mathbf{J}}_1$ and $\hat{\mathbf{J}}_2$:

$$\hat{\mathbf{J}} = \hat{\mathbf{J}}_1 + \hat{\mathbf{J}}_2, \tag{1.341}$$

where the respective angular momentum quantum numbers (j_1, m_1) and (j_2, m_2) associated with $\hat{\mathbf{J}}_1$ and $\hat{\mathbf{J}}_2$ are known (they can, therefore, mean azimuthal quantum number and magnetic quantum number or spin quantum number and spin magnetic quantum number or also any combination pair by pair). The operators $\hat{\mathbf{J}}_1$ and $\hat{\mathbf{J}}_2$ commute as they refer to two independent particles or to different properties of the same particle. I recall in fact that the components of $\hat{\mathbf{J}}$ obey the commutation relations (1.340).

In particular, let us denote the quantum numbers associated with $\hat{\mathbf{J}}$ by (j, m), then our goal is to find the possible values of j and m in terms of j_1 and m_1 for subsystem 1, and j_2 and m_2 for system 2. I first note that there are common eigenstates of $\hat{\mathbf{J}}_1^2$, $\hat{J}_{1z}$, $\hat{\mathbf{J}}_2^2$, and $\hat{J}_{2z}$

$$|\, j_1, m_1; j_2, m_2 \rangle \ , \tag{1.342}$$

such that $\hat{\mathbf{J}}_1^2$ and $\hat{J}_{1z}$ act only on $|\, j_1, m_1 \rangle$ as well as $\hat{\mathbf{J}}_2^2$ and $\hat{J}_{2z}$ act only on $|\, j_2, m_2 \rangle$. From Eq. (1.341), the z component of $\hat{\mathbf{J}}$ is given by

$$\hat{J}_z = \hat{J}_{1z} + \hat{J}_{2z}, \tag{1.343}$$

which upon acting on $|\, j_1, m_1; j_2, m_2 \rangle$ yields

$$\begin{aligned} \hat{J}_z \,|\, j_1, m_1; j_2, m_2 \rangle &= (\hat{J}_{1z} + \hat{J}_{2z}) \,|\, j_1, m_1 \rangle \otimes |\, j_2, m_2 \rangle \\ &= (m_1 + m_2)\hbar \,|\, j_1, m_1 \rangle \otimes |\, j_2, m_2 \rangle \\ &= (m_1 + m_2)\hbar \,|\, j_1, m_1; j_2, m_2 \rangle \ , \end{aligned} \tag{1.344}$$

where the cross product $\otimes$ separates the states of the two subsystems. Hence, the state $|\, j_1, m_1; j_2, m_2 \rangle$ is an eigenstate of $\hat{J}_z$ with the eigenvalue given by $(m_1 + m_2)\hbar$. The generalisation of this examination is provided by the Clebsch–Gordan coefficients, after the names of the German mathematicians Alfred Clebsch (1833–1872) and Paul Gordan (1837–1912). Let us denote the common eigenstates of $\hat{\mathbf{J}}^2$ and $\hat{J}_z$ by

$|\, j, m\rangle$.[124] From Eq. (1.344) and the superposition principle we can express $|\, j, m\rangle$ as

$$|\, j, m\rangle = \sum_{\substack{-j_1 \le m_1 \le j_1 \\ -j_2 \le m_2 \le j_2}}{}' C^{j, j_1, j_2}_{m, m_1, m_2} \,|\, j_1, m_1;\, j_2, m_2\rangle \,, \tag{1.345}$$

where the constants $C^{j, j_1, j_2}_{m, m_1, m_2}$ are the *Clebsch–Gordan coefficients* and a prime on the sum symbol means the summation over m_1 and m_2 is subject to the constraint $m_1 + m_2 = m$, resulting from Eq. (1.344). This result means that the number of possible values of m is the product of those of m_1 and m_2, i.e $(2j_1 + 1)(2j_2 + 1)$, in agreement with the result of Eq. (1.248). Since j is defined to be the maximum of m, one would expect that $j = j_1 + j_2$. This however is not correct because if it were the case then the number of possible values of m would be $2j + 1 = 2(j_1 + j_2) + 1$, instead of the actual number $(2j_1 + 1)(2j_2 + 1)$. A detailed analysis shows that the possible values of j are given by

$$j = j_1 + j_2,\ \ j_1 + j_2 - 1,\ \ \ldots,\ \ j_1 - j_2 + 1,\ \ |j_1 - j_2|, \tag{1.346}$$

and to each value of j there are $2j + 1$ possible values of m given by

$$m = j,\ \ j - 1,\ \ \ldots,\ \ -j + 1,\ \ -j. \tag{1.347}$$

The total number of possible values of m is then

$$\sum_{j=|j_1 - j_2|}^{j_1 + j_2} (2j + 1) = (2j_1 + 1)(2j_2 + 1), \tag{1.348}$$

as expected from previous calculations.

Two Statistics

The discovery of spin led to the acknowledgment that quantum systems are ruled by two fgtypes of particle statistics[125]: the Bose–Einstein statistics (after the names of A. Einstein and the Indian physicist and mathematician Satyendra N. Bose (1894–1974)) or the statistics of *bosons*, and the Fermi–Dirac statistics (after the names of P. Dirac and the Italian physicist Enrico Fermi (1901–1954)) or the statistics of *fermions*. The former refer to particles which have an integral spin and which are described by symmetric wave functions (functions that remain invariable upon an interchange of coordinates). The known elementary particles thus include fermions

[124] More precisely, we should consider the common eigenstates $|\, j_1,\, j_2,\, j, m\rangle$ of $\hat{\mathbf{J}}_1^2$, $\hat{\mathbf{J}}_2^2$, $\hat{\mathbf{J}}^2$, and $\hat{J}_z$. This however does not change our discussion and conclusions. Therefore, $|\, j, m\rangle$ can be thought of as a shorthand for $|\, j_1,\, j_2,\, j, m\rangle$.

[125] On this subject see Auletta et al. (2009, Chap. 7).

Table 1.3 Apart from the photon and gluons (and Higgs bosons), all elementary particles have mass. All leptons and quarks are spin $\frac{1}{2}$ particles, and all elementary bosons are spin 1 particles. The electron, muon, and tau lepton all have a charge of -1 (in units of elementary charge), and all neutrinos are electrically neutral. Quarks have fractional electric charge. The elementary bosons are force carriers: photon carries electromagnetic force, W and Z bosons carry the weak force, and gluons carry the strong force, responsible for the stability of the atomic nucleus. The photons, Z boson, and gluons are electrically neutral, while the W^+ and W^- bosons have a charge of $+1$ and -1, respectively. For each elementary particles there is a corresponding antiparticle with the same mass and spin, but opposite electric charge (according to the CPT theorem (Sect. 1.2.5)). Electrically neutral bosons are their own antiparticles, and the W^+ and W^- bosons are particle and antiparticle of each other. Antiparticles of leptons are known as anti-leptons ($\bar{l}$); antiparticles of quarks are called antiquarks ($\bar{q}$). For instance, the antiparticle of the electron is the positron (e^+), the antiparticle of electron neutrino is the anti-electron neutrino ($\bar{\nu}_e$), the antiparticle of the up quark is the anti-up quark ($\bar{u}$). However, it is an open question whether neutrinos are their own antiparticles. Adapted from Auletta and Wang (2014, p. 251)

Elementary particles		
Fermions		Bosons
Leptons	Quarks	
Electron (e^-), electron neutrino (ν_e)	Up (u), down (d)	Photon (γ)
Muon (μ^-), muon neutrino (ν_μ)	Charm (c), strange (s)	W, Z bosons ($W^\pm$, Z^0)
Tau lepton (τ^-), tau neutrino (ν_τ)	top (t), bottom (b)	Gluons (eight types) (g)

(six quarks and six leptons) and bosons (three kinds apart from the Higgs boson that will occupy us later). The names and symbols of the observed elementary particles are listed in Table 1.3. In the case of composed particles, both two bosons and two fermions give rise to a boson while the composition of a fermion and a boson gives rise to a fermion. This is a consequence of their different parity.

Permutation Operator

Thus, symmetric and antisymmetric state vectors or wave functions refer to the action of the *permutation* unitary operator $\hat{U}_P$, which exchanges the coordinates of particles.[126] In particular, in the case of two particles, we have

$$\hat{U}_P \, |x_1\rangle \, |x_2\rangle = |x_2\rangle \, |x_1\rangle \tag{1.349}$$

for symmetric particles and

$$\hat{U}_P \, |x_1\rangle \, |x_2\rangle = - \, |x_2\rangle \, |x_1\rangle \tag{1.350}$$

[126]The reader interested to know more can have a look at Auletta et al. (2009, Sect. 7.2) in particular.

for antisymmetric particles. Permutations are an example of *multiplicative* quantum numbers (we have previously dealt with some additive quantum numbers). I recall that permutations constitute a group (Sect. 1.2.5) so that two subsequent permutations constitute themselves a permutation. If we ask about the number n of reflections for going back to the initial state, in the case of symmetric particles, i.e. bosons, since two reflections bring back to the initial state and here parity has to be ± 1, the group has $n = 2$. In the case of fermions, due to their antisymmetry, we have $n = 4$, which means that parity would have to be $+i$ or $-i$, and its *double* mirror reflection would have the effect of a 2π rotation.[127] This is characteristic of spinors. Note that the dynamical states of a system of N identical particles are either all symmetric or all antisymmetric with respect to the exchange of any two particles. I also recall that this symmetry is not always conserved in time, as it happens for weak interactions. In fact, as mentioned in Sect. 1.2.5, parity must be taken together with charge and time to give a quantity that is always conserved. Permutation is a wide concept and it can be shown that also e.g. time evolution unitary operators turn out to be a species of permutation operators.[128]

Bosons' Statistics

When we have a certain quantity of bosons, an exchange between two of them does not affect their statistical distribution: in this case all that matters is the number of particles occupying a certain energy level. This statistics can be expressed by means of a binomial coefficient in the following way: the number ζ_S (the subscript denotes symmetric distribution) of possible ways of distributing N bosons among k states, which could represent energy levels (and, for our purposes, could be represented by different boxes) is given by the combinations without repetition of $k - 1$ elements among $N + k - 1$, that is,

$$\zeta_S = \binom{k + N - 1}{k - 1} , \tag{1.351}$$

where a binomial coefficient, in its generality, is defined as

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}, \tag{1.352}$$

and I recall that $n! = 1 \cdot 2 \cdot 3 \cdots (n-1) \cdot n$ is the factorial. Therefore, what this statistics tells us is that bosons are *indistinguishable*. For instance, if we let cross two photons at a so-called beam splitter (a half-silvered mirror for merging and splitting beams of photons or neutrons) and let them emerge along two different paths, then we shall be unable to say which photons has taken which path even if the inputs came from two different LASERs (Light Amplification by Stimulated Emission of

[127] Penrose (2004, Sect. 5.5).

[128] 'T Hooft (2016, p. 20).

Radiation). This sets severe limitations in the way in which we can make ontological ascription to quantum systems: in fact, if they cannot be distinguished, how can we single out a system at all?

Indistinguishability may be considered a strangeness of QM. Nevertheless, there is an interesting thought experiment of the American physicist, chemist and mathematician Josiah Willard Gibbs (1839–1903) that can help us here[129]: suppose a box separated in two equal compartments filled by a gas like hydrogen with a shutter that at a certain time is opened allowing the two portions of gas to mix. Now, the properties of the gas in equilibrium are fully characterised by its volume, temperature, and density. If the atoms composing the gas would differ from each other, the entropy (measure of the disorder proportional to the different ways to distribute the molecules) resulting from mixing would be much more than that which is in fact observed and that depends finally on those parameters.

Fermions' Statistics

The second statistics refers to particles of half-integral spin which are described by antisymmetrical wave functions. These ones, called fermions, correspond to the particles that constitute the foundations of matter such as the electrons, protons and neutrons. For fermions, the individuality of each of them is to a certain extent important (although in most situations they cannot be distinguished). In fact, according to both Pauli's exclusion and quantisation principles, due to their fractional spin, electrons are distributed on various energetic levels of an atom such that at most a couple of them with *opposite* spin orientation (along a given direction) can occupy the same orbital level.[130] So the electrons conform to the statistical distribution of Fermi–Dirac. The Fermi–Dirac statistics (with antisymmetric distribution) in this case is determined by the number of ways one can choose N among k elements:

$$\zeta_A = \frac{k!}{N!(k-N)!} = \binom{k}{N} . \tag{1.353}$$

1.3.2 The Correspondence Principle

A Heuristic Principle

In the early stages of the quantum theory, and in the absence of an independent formalism, there was only a qualitative or semi-classical treatment of many

[129]Reported in Wilczek (2006). An accessible book (not a textbook!) to many problems of fundamental physics.

[130]Pauli (1925).

problems. Planck proposed, as early as 1906, the following Limiting assumption: "Classical mechanics is a limit case of microphysics when h is insignificant".[131] From this assumption Bohr, in 1920, following Einstein's work on absorbed and emitted radiation (including stimulated emission, crucial for lasers),[132] could formulate the Correspondence Principle between QM and classical physics: the elements (like the observables) of quantum theory have analogous elements (variables) in classical physics.[133] Historically, this principle allowed in fact the definition of some key operators, such as the Hamiltonian operator or, along the lines of its classical meaning, the momentum operator.

The reason for this principle is to be found in the fact that CM was well established at that time as a complete theory, able to describe (macroscopic) physical quantities very successfully (see Sect. 1.1.3), while the new mechanics, in its developmental stage (between 1900 and 1927), was nothing more than an ensemble of a small number of assumptions and experiments. Thus, to fill the gaps, recourse to classical concepts was necessary. This exigency was expressed in very cogent form by Dirac[134]:

> QM was built upon a foundation of analogy with the Hamiltonian theory of CM. This is because the classical notion of canonical coordinates and momenta was found to be one with a very simple quantum analogue, as a result of which the whole of the classical Hamiltonian theory, which is just a structure built up on this notion, could be taken over in all its details into QM.

The Problem of the Limit

From the inverse point of view, that from QM to CM, this principle is applicable within that physical scale in which the action (1.23) is far greater than Planck's constant. This is typically the macroscopic scale of our ordinary experience. At this scale, according to Planck, the value of this constant is insignificant and so no quantum effects can be detected from a technical point of view (at least currently: if they are present, they should be tiny). In other words, Planck's constant, in spite of being a constant, is irrelevant given its null *relative weight* regarding the action of the system.

It is important to remark that we cannot understand such a limit as a kind of reduction of QM to CM or vice versa. In the limit in which h tends to zero, we should recover a dimensionless quantity containing h such that we can perform a Taylor expansion in that limit (Sect. 1.1.3). This works well in some cases. Planck

[131] Planck (1906).

[132] Einstein (1916). See also Home and Whitaker (2007, pp. 26–27), Schwartz (2014, Sect. 2.1).

[133] A summary in Jammer (1966, Sect. 3.2). See also Auletta (2000, Sect. 3.2). In Bokulich (2008, Sect. 4.2) it is argued that Bohr's formulation of the principle was slightly different and concerned the correspondence between transition probabilities between stationary states like for different atomic levels, on the one hand, and amplitudes of the components of a classical harmonic oscillator. This should correspond to the early formulation of the principle but it is rather limited in scope.

[134] Dirac (1933, p. 111).

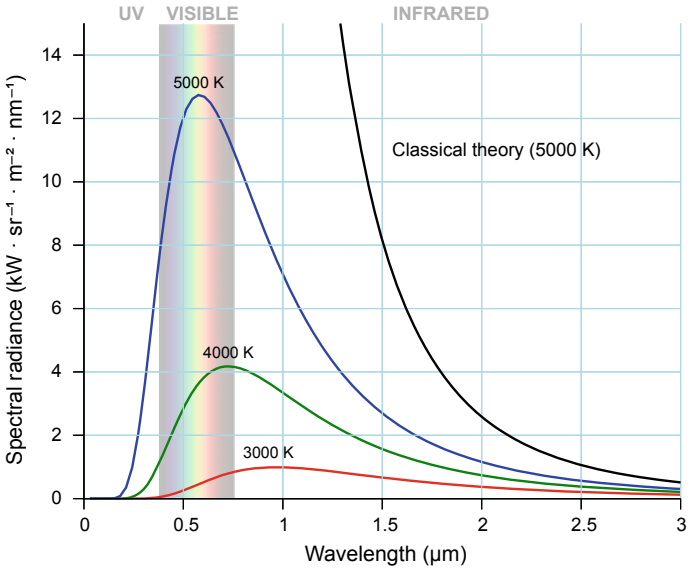


Fig. 1.27 Planck's radiation spectrum for different temperatures. The temperature is expressed in Kelvin

was able to derive the following formula for the spectrum of black body radiation (Fig. 1.27):

$$f(\nu) = \frac{8\pi\nu^2}{c^2} \frac{h\nu}{e^{h\nu\beta} - 1} \,, \tag{1.354}$$

with $\beta := (k_B T)^{-1}$, T being the temperature, and

$$k_B = 1.38064852(79) \times 10^{-23} \mathrm{J/K} = \mathrm{m^2 kgs^{-2} K^{-1}} \tag{1.355}$$

the Boltzmann constant (after the name of the Austrian physicist and philosopher Ludwig Boltzmann (1844–1906)), with physical dimension $[\mathrm{E}][\Theta^{-1}]$. Note that the proportionality factor $8\pi\nu^2/c^2$ is derived from Maxwell's equations and in Planck's formula (1.1) allows to convert the resonators' energy $h\nu$ to the field energy $f(\nu)$.[135] Now, this formula can be reduced to the classical Rayleigh–Jeans formula (after the names of the British physicists John W. S. Rayleigh (1842–1919) and James H. Jeans (1877–1946))

$$f(\nu) = \frac{8\pi\nu^2}{c^2} k_B T \,, \tag{1.356}$$

[135] See Kuhn (1978, p. 118).

by taking the dimensionless quantity $h\nu\beta$ and making a Taylor expansion of the exponential. However, the British physicist Michael Berry has pointed out that there are cases in which this is not possible because the limit is singular.[136] In fact, the issue here is not whether one quantity can be derived from the other but whether there is consistency at the limit, that is, that the implications of one theory (or set of equations) are congruent with the observational features of the other.[137]

1.3.3 The Complementarity Principle

Complementarity

The complementarity principle was formulated by Bohr in 1927 (after the discovery of uncertainty relations) as a first attempt at imposing a certain order and coherence to the quick, spot-wise, and surprising building of quantum theory. In particular, it was understood as a way to cope with the wave–particle duality (Sect. 1.1.3), one of the biggest mysteries brought by the theory.[138]

In order to understand the meaning of the principle, it should be noted that, according to Bohr, two types of description are said to be *complementary* only if they are at once (at least quantum-mechanically) *incompatible*, but both are *classically necessary* to describe a physical system in a complete way. According to this formulation, there must be two necessary conditions: the exclusiveness of each of the complementarity types and the necessity of both to attain a fuller (classical) description of the system. The exclusiveness of the two types of complementary description has been related also to the fact that it is impossible to determine two conjugated observables at the same time above a threshold fixed by the uncertainty relations (Sect. 1.2.4). As for necessity, in order to classically know a system's state, it is necessary to know the two conjugated variables. I remark that, since it is not possible to know with absolute precision two conjugated observables, QM could be branded incomplete as compared with CM. This will be the object of later examination, in particular,

[136]For all this issue, I recommend (Bokulich 2008, Sect. 1.4). Bokulich points out that also the Ehrenfest theorem (allowing to derive quantum analogues of the Hamilton's equations (1.45): on this see Auletta et al. (2009, Sect. 3.7) cannot be taken to adequately characterise the classical limit. In fact, the related equation only holds under highly restricted circumstances due to the fact that quantum-mechanical expectation values do not follow classical equations of motion. Thus, in that limit, one must be able to replace the mean values of functions (say, of position) with a function of the means (of position). However, "this substitution is legitimate only if the Hamiltonian of the system is a polynomial of second degree or less, such as in the case of linear or harmonic oscillator potentials."

[137]Home and Whitaker (2007, p. 54).

[138]It was made known to a large scientific audience in the historical paper (Bohr 1928). For historical reconstruction see Jammer (1966, Sect. 7.2). See also Auletta (2000, Chap. 8) for further examination and references. In fact, there is a huge amount of different interpretations of the meaning of this principle.

focusing on the experiment proposed by Einstein, Podolsky, Rosen for testing this point.

Wave–Particle Dualism

Bohr thought that the corpuscular-undulatory description represented the basic type of complementarity,[139] which does not coincide with the issue of the non-commutativity of quantum observables. This means, as stated by Bohr, that both the corpuscular and the undulatory behaviours are *classical descriptions* by means of which we approach quantum systems. Thus, we can remove the aforesaid dualism by attributing it to the different and complementary *ways* with which we describe, and interact with, the system. This sets the principle in epistemic terms. Both complementary descriptions would be an attempt at interpreting experimental evidence wherever our theory finds itself limited and insufficient. The Dutch physicist Gerard 'T Hooft calls this interpretation psi-epistemic.[140] This formulation of the Complementarity principle has been the basis of what was known later of as the *Copenhagen interpretation* (I recall that Bohr was in fact Danish).

Although Bohr stressed that this complementarity must be expressed in classical terms, I would like to point out from the start that classical depictions like "wave" and "particle", strictly speaking, do not fit the new situation described by QM. In fact, quantum systems present peculiarities that are unknown classically in both cases (when they are understood as particles and when they are understood as waves). Some of them have been already pointed out and some others shall be discovered in the course of our examination. Nevertheless, I provisionally use these terms and accept their usage taking into account that they do not exhibit a satisfactory physical interpretation. These last considerations will allow us to deal with an ulterior (and more general) formulation of the complementarity principle later on. As I shall argue in the following, the complementarity principle should be understood not in terms of sharp opposition between two alternative descriptions but in terms of a smooth complementarity between two behaviours expressed in ontological and not epistemic, terms: if we follow Bohr's own understanding of the principle, both particle and wave are concepts of a classical origin that in fact represent the two extremes of the range of behaviours typical of quantum systems. Which amounts to say that the generalised (and corrected) Complementarity principle states that quantum-mechanical entities 'oscillate' between these two extreme forms of behaviour and can be in any possible intermediate regime.

[139] A point of view still supported in Englert et al. (1994).

[140] 'T Hooft (2016, p. 30).

Relations Among the Principles

After having summarised the principles that govern QM, I would like to show some relationships among them. As we have seen, the uncertainty relations results from the quantisation principle, and somehow display a limiting nature, given that they exert constraints on the amount of knowledge we could draw from any given system and, likewise, its measurability. It is paramount to underline that it is the uncertainty relations that restrict both the intra and extra data of a system, since their very nature impedes the total knowledge of two observables that do not commute between each other. Or put in other terms, such a restriction is independent of the way which we can draw the data or of how we measure (Sect. 1.2.4).

The superposition principle is a kind of 'positive' principle as far as it enlarges enormously the number of states that were possible according to CM, since any appropriate combination (superposition) of states gives rise to another state in the same Hilbert space (Sect. 1.2.1). The complementarity principle, especially considering its reformulation, is an autonomous principle that, in spite of the efforts by the scientific community to make it derive from the uncertainty relations, retains its foundational character.[141]

While the correspondence principle has no foundational character, the issue of the limit between quantum and classical behaviour is a crucial question that needs to be discussed step by step in the following. Finally, Pauli's exclusion principle is in fact absorbed in the two different statistics. In other words, up to now, no fundamental reason has been provided of why there are bosons and fermions apart from the fact that two different forms (symmetric and antisymmetric) of wave function are possible, although the latter denotes a general character of QM.

1.4 Density Operator and Compound Systems

When we look around us we see everywhere compound systems, so that we are accustomed to deal with them in several sciences. However, being quantum systems the most elementary ones in Nature, it is not a surprise that their interconnections show several tricky aspects. Let us introduce here the most important ones.

1.4.1 Pure and Mixed States

Several Quantum Systems in the Same State

A very helpful formalism for treating many of the subsequent problems as well as some conceptual nodes is centred on the notion of *density matrix*, (for the finite

[141] Englert et al. (1994).

cases) which can be considered as a generalisation of the concept of projector.[142] The focus here is the description of *many* quantum systems (for instance, photons) prepared in some state (a population of quanta). Let us first consider the elementary example of photon polarisation. Suppose that we have prepared a large number of photons and all of the photons are in the same polarisation state. Making use of the basis horizontal–vertical polarisation that we express in analogy with Eqs. (1.54) as

$$|\,h\rangle := \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad |\,v\rangle := \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{1.357}$$

we assume that the polarisation state of the photons is the two-state superposition (1.103), which I rewrite here

$$|\,\psi\rangle = c_h\,|\,h\rangle + c_v\,|\,v\rangle\,, \tag{1.358}$$

where $c_h, c_v \in \mathbb{C}$ are probability amplitudes satisfying $|c_h|^2 + |c_v|^2 = 1$. The density matrix that describes the system, which in this case coincides with the corresponding projector, is given by

$$\begin{aligned} \hat{\rho} = \hat{P}_\psi &= |\,\psi\rangle\,\langle\psi\,| \\ &= (c_h\,|\,h\rangle + c_v\,|\,v\rangle)\,\left(c_h^*\,\langle h\,| + c_v^*\,\langle v\,|\right) \\ &= |c_h|^2\,|\,h\rangle\,\langle h\,| + |c_v|^2\,|\,v\rangle\,\langle v\,| + c_v^* c_h\,|\,h\rangle\,\langle v\,| + c_v c_h^*\,|\,v\rangle\,\langle h\,|\,. \end{aligned} \tag{1.359}$$

In the basis $\{|\,h\rangle\,, |\,v\rangle\,\}$, $\hat{\rho}$ takes the matrix form

$$\hat{\rho} = \begin{bmatrix} |c_h|^2 & c_h c_v^* \\ c_h^* c_v & |c_v|^2 \end{bmatrix}. \tag{1.360}$$

What these equations tell us is that *all* of the photons being described are in the superposition state (1.358). Note in particular the "interference" terms $c_h c_v^*$ and $c_h^* c_v$, which are not probabilities and depend from the self-interference of each photon in the state $|\,\psi\rangle$ (Sect. 1.2.1). Of course, the density matrix (1.360) is a projector that is not written in its eigenbasis. The latter would be represented by $|\,\psi\rangle\,, \left|\,\psi^\perp\right\rangle$, where $|\,\psi\rangle$ is given by Eq. (1.358) and its orthogonal $\left|\,\psi^\perp\right\rangle = c_h^\perp\,|\,h\rangle + c_v^\perp\,|\,v\rangle$, (with coefficients satisfying the orthonormality condition and $\left\langle h\left|c_h^\star c_h^\perp\right|\,h\right\rangle = -\left\langle v\left|c_v^\star c_v^\perp\right|\,v\right\rangle$ in order to satisfy $\left\langle\psi\,|\,\psi^\perp\right\rangle = 0$). In such eigenbasis, that density matrix takes the canonical form of projectors of the kind (1.84). Note that it is always possible to find an eigenbasis of a density matrix (as for any other observable) through a unitary transformation.[143] In our case, we have $\hat{U}^\dagger \hat{\rho} \hat{U}$, with

[142] Density matrices were first introduced in Landau (1927). I follow here Auletta and Wang (2014, Sects. 7.8–7.9). The reader interested in a more technical treatment can see Auletta et al. (2009, Chap. 5).

[143] Byron and Fuller (1969–70, I, p. 123).

$$\hat{U} = \frac{1}{\sqrt{2}} \begin{bmatrix} c_h & c_h^{\perp} \\ c_v & c_v^{\perp} \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} \langle h \mid \psi \rangle & \langle h \mid \psi^{\perp} \rangle \\ \langle v \mid \psi \rangle & \langle v \mid \psi^{\perp} \rangle \end{bmatrix}, \tag{1.361}$$

where the two columns are vectors $|\,\psi\rangle$, $\left|\,\psi^{\perp}\right\rangle$, respectively.

Quantum Systems in Different States

Consider now the following situation that is quite common in physics. Suppose that we have prepared a large number of photons, with some of the photons in horizontal polarisation and the others in vertical polarisation. The state of the photons, in this case, is totally different from the superposition state $|\,\psi\rangle$ given by Eq. (1.358) and we need to represent it with two distinct projectors, namely, one for horizontal polarisation and the other for vertical polarisation. The corresponding density matrix is given by

$$\hat{\hat{\rho}} = \wp_h \hat{P}_h + \wp_v \hat{P}_v, \tag{1.362}$$

where $\wp_h + \wp_v = 1$ with $\wp_h = |c_h|^2$ and $\wp_v = |c_v|^2$ being the respective probabilities for finding each photon either in the horizontal or the vertical polarisation. In this basis, the projectors $\hat{P}_h$ and $\hat{P}_v$ have the form of projectors (1.84), that is,

$$\hat{P}_h = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \hat{P}_v = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}. \tag{1.363}$$

We might note that the density matrix is no longer a projector, but rather a *weighted sum* of projectors (where $\wp_h$ and $\wp_v$ are those weights):

$$\hat{\hat{\rho}} = \begin{bmatrix} \wp_h & 0 \\ 0 & \wp_v \end{bmatrix}, \tag{1.364}$$

where only diagonal elements are non-zero and interference terms have thus disappeared.

Quantum states described by density matrices of the type in Eq. (1.359) are called *pure states*, while those described by density matrices of the form in Eq. (1.362) are called *mixed states* (or *mixtures*). Equation (1.362) can be easily generalised to

$$\hat{\hat{\rho}} = \sum_j \wp_j \hat{P}_j, \tag{1.365}$$

where I stress that the states $\hat{P}_j$ in which the system may be found do *not* have to be mutually orthogonal (as it happens for $\hat{P}_h$ and $\hat{P}_v$), while Eq. (1.359) to

$$\hat{\rho} = |\,\psi\rangle\,\langle\psi\,| = \sum_j |c_j|^2 |b_j\rangle\langle b_j| + \sum_{j \neq k} c_j c_k^* |b_j\rangle\langle b_k| \tag{1.366}$$

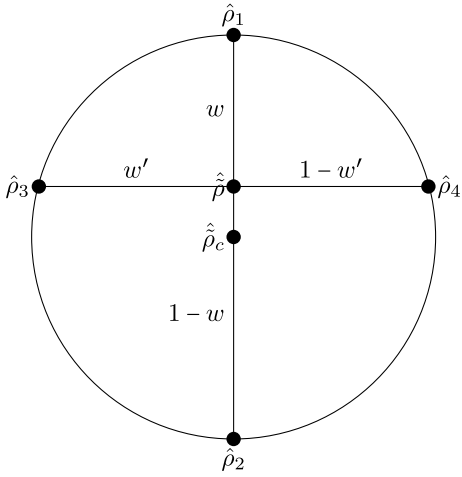


Fig. 1.28 Representation on a disk of the space of 2D pure and mixed states. Due to convexity, pure states are represented on the border while mixture in the interior of the disk. Note the mixtures $\hat{\hat{\rho}}$ and $\hat{\hat{\rho}}_c$. The latter is equal to $\frac{1}{2}\hat{I}$. Note also that the mixture $\hat{\hat{\rho}}$ can be both expanded using the pure states $\hat{\rho}_1, \hat{\rho}_2$ as well as $\hat{\rho}_3, \hat{\rho}_4$, where the geometric distance from $\hat{\rho}_1, \hat{\rho}_2$ are $w, 1-w$, respectively, as well as the geometric distance from $\hat{\rho}_3, \hat{\rho}_4$ are $w', 1-w'$, respectively. Note also that the first couple represents orthogonal states since they are opposite points of a diameter while the latter does not. Adapted from Auletta et al. (2009, p. 187)

for some arbitrary basis $\{b_j\}$ spanning the Hilbert space of the system.

It is evident that pure states are the special case of mixed states in which the system is found in one and only state with probability 1 (only one of the $\wp_j$ in Eq. (1.365) is equal to 1, all the others are 0). In fact, Eq. (1.362) or (1.365) allows us to say that any density matrix (whether pure or a mixture) can be written as a weighted sum of pure states (or projectors). Such a property is called *convexity*. In vectorial spaces, a set $\mathcal{C}$ in that space is said to be *convex* if e.g., $\forall x, y$ in $\mathcal{C}$ and all w in the interval $[0, 1]$, the point $wx + (1-w)y$ also belongs to $\mathcal{C}$. Thus, when dealing with density matrices, any arbitrary matrix $\hat{\rho}$ can be expanded e.g. as follows (Fig. 1.28)[144]:

$$\hat{\rho} = w\hat{\rho}_1 + (1-w)\hat{\rho}_2, \tag{1.367a}$$

$$\hat{\rho} = w'\hat{\rho}_3 + (1-w')\hat{\rho}_4, \tag{1.367b}$$

for some pure states $\hat{\rho}_j$'s ($j = 1, \ldots, 4$), and where one of the two weights $w, 1-w$ as well as one of the two weights $w', 1-w'$ are zero if $\hat{\rho}$ is a pure state. Otherwise, it is a mixture. Note that we have a *maximally mixed* state when all the weights are equal (it corresponds to a state of maximal entropy, on which issue we shall come back). In such a case, for any n dimension of the Hilbert space, due to the completeness relation (1.94), we have

[144] See also Landsman (2017, Sects. 2.2–2.3).

$$\hat{\tilde{\rho}}_c := \frac{1}{n} \sum_n \hat{P}_n = \frac{1}{n} \hat{I}, \tag{1.368}$$

where the subscript c denotes that the maximally mixed density matrix occupies geometrically always the centre of the (hyper-)sphere of the density operators as displayed in Fig. 1.28 (for the 2D space). Thus, a completely mixed state is one such that all components are equiprobable and contains every other state in its convex decomposition.

Trace

It is helpful to introduce the symbol Tr for denoting the *trace* of a (square) matrix, i.e. the sum of the diagonal elements of a matrix, i.e. in mathematical terms,

$$\mathrm{Tr}\, \hat{A} := \sum_j A_{jj}, \tag{1.369}$$

where $\hat{A}$ is a matrix. In the 2D case, if $\hat{\rho} = |\psi\rangle \langle\psi|$ and $|\psi\rangle$ is given by the superposition (1.358), we have

$$\mathrm{Tr}\, \hat{\rho} = \sum_{j=v,h} \langle j | \hat{\rho} | j \rangle, \tag{1.370}$$

which can be easily generalised.

The trace satisfies the following basic properties:

- $\mathrm{Tr}(\hat{A} + \hat{B}) = \mathrm{Tr}\, \hat{A} + \mathrm{Tr}\, \hat{B}$, (1.371a)
- $\mathrm{Tr}(C\hat{A}) = C\, \mathrm{Tr}\, \hat{A}$, (1.371b)
- $\mathrm{Tr}(\hat{A}\hat{B}) = \mathrm{Tr}(\hat{B}\hat{A})$, (1.371c)

where $C \in \mathbb{C}$ is a constant. Property (1.371c) is called the *cyclic property* of the trace, which can be generalised easily to n operators. From this property it follows that

$$\begin{aligned} \mathrm{Tr}(\hat{U}^\dagger \hat{A} \hat{U}) &= \mathrm{Tr}(\hat{U}\hat{U}^\dagger \hat{A}) \\ &= \mathrm{Tr}(\hat{A}\hat{U}\hat{U}^\dagger) \\ &= \mathrm{Tr}\, \hat{A}, \end{aligned} \tag{1.372}$$

where $\hat{U}$ is a unitary matrix. This result shows that the trace of a matrix (or an operator) is invariant under any unitary transformation.

Properties of the Density Operator and Purity

The density operator satisfies the following important properties:

- $\hat{\rho}^{\dagger} = \hat{\rho}$, (1.373a)
- $\mathrm{Tr}\,\hat{\rho} = 1$, (1.373b)
- $\langle \psi \left| \hat{\rho} \right| \psi \rangle \geq 0$ for all $|\psi\rangle$, (1.373c)
- $\hat{\rho}^2_{\mathrm{pure}} = \hat{\rho}_{\mathrm{pure}}$. (1.373d)

The first property means that the density matrix is Hermitian. This is evident as the probabilities $\wp_j$ are real (and non-negative) while the projectors $\hat{P}_j$ are themselves Hermitian (see e.g. Eq. (1.119)). About the latter property, since the density matrix for a pure state is a projector, the square of a projector is the projector itself (see Eqs. (1.89)). Since this is not true for mixtures, a good measure of the *purity* of a state is represented by

$$\mathrm{M} = \mathrm{Tr}\left(\hat{\rho}^2\right) \ . \tag{1.374}$$

It is clear that states inside the (hyper-)sphere of density operators will have $\mathrm{M} < 1$, while on the surface $\mathrm{M} = 1$. Let us consider again the centre $\hat{\tilde{\rho}}_{\mathrm{c}}$ of the sphere, as defined in Eq. (1.368) and represented in Fig. 1.28. As any other mixed state, this state has an infinite number of possible representations. If we take the vertical diameter, in the polarisation example with equal probability amplitudes, we have

$$\hat{\tilde{\rho}}_{\mathrm{c}} = \frac{1}{2}\left(|h\rangle\langle h| + |v\rangle\langle v|\right) \ . \tag{1.375}$$

Therefore, the matrix form of $\hat{\tilde{\rho}}_{\mathrm{c}}$ in the basis $\{|h\rangle\ , |v\rangle\}$ will be given by

$$\hat{\tilde{\rho}}_{\mathrm{c}} = \frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \frac{1}{2}\hat{I}\ , \tag{1.376}$$

and

$$\hat{\tilde{\rho}}^2_{\mathrm{c}} = \frac{1}{4}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \frac{1}{4}\hat{I}\ , \tag{1.377}$$

which yields $\mathrm{M} = 1/2$, and this corresponds to the minimum value of $\mathrm{Tr}\left(\hat{\rho}^2\right)$ for a bidimensional system. It is straightforward to generalise the previous result to a nD Hilbert space as follows:

$$\mathrm{M}^{(n)}_{c} = \frac{1}{n}\ , \tag{1.378}$$

which tells us that the purity of a completely mixed and equally weighted n-dimensional density matrix (of a centre) is simply the inverse of the number of possible outcomes (projections). It is then clear that for $n \to \infty$, $\mathrm{M}^{(n)}_{c} \to 0$.

Probability and Expectation Value

Note that both probabilities (1.101) and expectation values (1.102) can be expressed by making use of this formalism. The probability of obtaining the event described by the projector $\hat{P}_j = \left|o_j\right\rangle\left\langle o_j\right|$ given the initial state $\hat{\rho}$ of the system can be written as

$$\wp_j = \mathrm{Tr}(\hat{\rho}\hat{P}_j) = \langle\psi \mid o_j\rangle\langle o_j \mid \psi\rangle = \left|\langle o_j \mid \psi\rangle\right|^2, \tag{1.379}$$

while the expectation value of an observable $\hat{O}$ in the state $\hat{\rho}$ is given by

$$\left\langle\hat{O}\right\rangle_{\hat{\rho}} = \mathrm{Tr}(\hat{\rho}\hat{O}) = \left\langle\psi\left|\hat{O}\right|\psi\right\rangle. \tag{1.380}$$

1.4.2 Entangled and Product States

Composite Systems and Hilbert Spaces

As said, the formalism of the density matrix is very useful when dealing with composite systems (or compound systems). Another term used in the literature is *multipartite* system and, in the case of two subsystems, *bipartite* system. Let us suppose to have a system of two photons, both of which can be in horizontal or vertical polarisation (or in any superposition of the two polarisation states). Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be the Hilbert spaces for the polarisation states of photon 1 and photon 2, respectively, and

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 \tag{1.381}$$

the Hilbert space for the composite system of the two photons, where I recall that $\otimes$ is the tensor or direct product (already met in Eq. (1.319)), also known as the Kronecker product.[145] In fact, while $\mathcal{H}_1$ and $\mathcal{H}_2$ are identical 2D Hilbert spaces, they are nonetheless two independent Hilbert spaces. In other words, when we deal with composite systems with dimensions m, n, we need to consider a Hilbert space $\mathcal{H}$ for the whole system with dimension $m \times n$. When doubts may arise, I shall insert subscripts for specifying which is the Hilbert space of the system under consideration. In this case, we have two orthonormal bases

$$\{|v\rangle_1, |h\rangle_1\} \quad \text{and} \quad \{|v\rangle_2, |h\rangle_2\}, \tag{1.382}$$

where the first is a basis in $\mathcal{H}_1$ and the second a basis in $\mathcal{H}_2$. Thus, when we consider the two systems together, we need to deal with a basis in the Hilbert space $\mathcal{H}$ of the composite system that is given by the direct products of the basis vectors in $\mathcal{H}_1$ and $\mathcal{H}_2$:

[145] For the mathematical aspects see Byron and Fuller (1969–70, I, Sect. 3.11).

$$\{|h\rangle_1 \otimes |h\rangle_2\ ,\ |h\rangle_1 \otimes |v\rangle_2\ ,\ |v\rangle_1 \otimes |h\rangle_2\ ,\ |v\rangle_1 \otimes |v\rangle_2\}, \tag{1.383}$$

where each basis vector to the left of the symbol $\otimes$ pertains to the Hilbert space $\mathcal{H}_1$, while each basis vector to its right pertains to the Hilbert space $\mathcal{H}_2$. It is evident that the direct product basis (1.383) is an orthonormal basis. Therefore, the corresponding basis vectors for the 4D Hilbert space of the compound system, as a particular instance of (1.70), may be written in component form as

$$|h\rangle_1 \otimes |h\rangle_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix},\ |h\rangle_1 \otimes |v\rangle_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix},\ |v\rangle_1 \otimes |h\rangle_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix},\ |v\rangle_1 \otimes |v\rangle_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \tag{1.384}$$

Here and henceforth, when no equivocation is possible, I shall drop the symbol $\otimes$ between the kets, condense the kets into a single one, or both. For instance, the following expressions $|h\rangle_1 \otimes |h\rangle_2$, $|h\rangle_1 |h\rangle_2$, $|h_1 \otimes h_2\rangle$, $|h_1, h_2\rangle$, and $|h_1 h_2\rangle$ all represent the same direct product state. Moreover, here and henceforth, I shall use the state vectors with uppercase (often Greek) letters (like $|\Psi\rangle$, $|\Phi\rangle$, *etc.*) to denote the states of a composite system. Again for the sake of notational simplicity, here and henceforth, I shall suppress the indexes 1 and 2 that respectively denote systems 1 and 2 whenever no confusion may arise.

Separated Systems

We are interested in ascertaining which kind of relation, if any, can exist between these two systems. Let us first suppose that the two systems are *separated* (that is, there is no connection whatsoever between them) and each is in the state of linear polarisation at 45° (or $\frac{\pi}{4}$ in radians) relative to the horizontal polarisation, that is, we have

$$|\psi\rangle_1 = \frac{1}{\sqrt{2}}(|h\rangle_1 + |v\rangle_1), \quad |\psi\rangle_2 = \frac{1}{\sqrt{2}}(|h\rangle_2 + |v\rangle_2). \tag{1.385}$$

Since by *separated* is meant that no action whatsoever performed locally on one of the systems can have effects on the other one, then every possible combination of the allowed measurement outcomes of the two systems can occur, or, in other words, the probability distributions of the possible measurement outcomes of the two systems are statistically independent. If there are no additional constraints on the two systems, we shall expect that, in measurements of a sufficient large number of identically prepared systems, the composite system on average will have equal probability of $\frac{1}{4}$ to be found in the states described by the four product basis vectors (1.383). Therefore, we have

$$
\begin{aligned}
|\Psi\rangle_{12} &= |\psi\rangle_1 \otimes |\psi\rangle_2 \\
&= \frac{1}{\sqrt{2}}(|h\rangle_1 + |v\rangle_1) \otimes \frac{1}{\sqrt{2}}(|h\rangle_2 + |v\rangle_2) \\
&= \frac{1}{2}\left(|h\rangle_1 \otimes |h\rangle_2 + |h\rangle_1 \otimes |v\rangle_2 + |v\rangle_1 \otimes |h\rangle_2 + |v\rangle_1 \otimes |v\rangle_2\right). \quad (1.386)
\end{aligned}
$$

For instance, if we find that system 1 is, say, in the state of vertical polarisation, system 2 can be either vertically or horizontally polarised, both possibilities having equal probability. States of this kind are called *product states*, since they are mathematically represented by the simple product of the state vectors of the two (or more) involved systems, like on the first line of the previous equation. If we like to enlarge our considerations to other possible cases, we need to describe system 1 and 2 by the state vectors $|\psi\rangle_1$ and $|\psi\rangle_2$, respectively, having the form

$$
|\psi\rangle_1 = c_h |h\rangle_1 + c_v |v\rangle_1, \quad |\psi\rangle_2 = c'_h |h\rangle_2 + c'_v |v\rangle_2, \quad (1.387)
$$

with all coefficients $\in \mathbb{C}$, and describe the composite system by the state vector $|\Psi\rangle$

$$
\begin{aligned}
|\Psi\rangle_{12} &= |\psi\rangle_1 \otimes |\psi\rangle_2 \quad (1.388) \\
&= c_h c'_h |h\rangle_1 \otimes |h\rangle_2 + c_h c'_v |h\rangle_1 \otimes |v\rangle_2 + c_v c'_h |v\rangle_1 \otimes |h\rangle_2 + c_v c'_v |v\rangle_1 \otimes |v\rangle_2,
\end{aligned}
$$

where the normalisation condition needs to be satisfied. Here, the crucial point is that some of these double coefficients could be 0, which in this case could give rise to correlations between systems 1 and 2 and therefore to a different density matrix, as we shall see.

Before dealing with this problem, note that the density matrix of product states is a pure state. In general terms (and always for the 2D case), if we have the product state

$$
\begin{aligned}
\hat{\rho}_{jkmn} &= \hat{\rho}_{jk} \otimes \hat{\rho}_{mn} \\
&= \sum_{j,k,m,n} c_j c_k^\star c_m c_n^\star |j\rangle \langle k| \otimes |m\rangle \langle n|, \quad (1.389)
\end{aligned}
$$

where $\hat{\rho}_{jk} = |\psi\rangle \langle\psi|$, $\hat{\rho}_{mn} = |\psi'\rangle\langle\psi'|$ with

$$
|\psi\rangle = c_j |j\rangle + c_k |k\rangle \text{ and } |\psi'\rangle = c_m |m\rangle + c_n |n\rangle, \quad (1.390)
$$

and use the basis (1.384) with $|hh\rangle$ replaced by $|jm\rangle$, $|hv\rangle$ by $|jn\rangle$, $|vh\rangle$ by $|km\rangle$, $|vv\rangle$ by $|kn\rangle$, we can write the resulting compound density matrix as

$$
\hat{\rho}_{jkmn} = \begin{bmatrix}
\rho_{jjmm} & \rho_{jjmn} & \rho_{jkmm} & \rho_{jkmn} \\
\rho_{jjnm} & \rho_{jjnn} & \rho_{jknm} & \rho_{jknn} \\
\rho_{kjmm} & \rho_{kjmn} & \rho_{kkmm} & \rho_{kkmn} \\
\rho_{kjnm} & \rho_{kjnn} & \rho_{kknm} & \rho_{kknn}
\end{bmatrix}, \quad (1.391)
$$

where $\rho_{jkmn} = c_j c_k^\star c_m c_n^\star$. Clearly, we have the interference terms that are typical of a pure state.

Entangled States

We now consider the situation in which some coefficients in Eq. (1.388) are 0. Suppose, in particular, that the composite system at a certain time is described by the state vector

$$|\,\Phi\rangle_{12} = \frac{1}{\sqrt{2}}(|\,h\rangle_1 \otimes |\,h\rangle_2 + |\,v\rangle_1 \otimes |\,v\rangle_2). \tag{1.392}$$

This equation is surprising. According to it, when system 1 is in the state of horizontal polarisation so is system 2, and vice versa; similarly, when system 1 is in the state of vertical polarisation, the same occurs for system 2, and vice versa. In this case, the two systems are no longer separated. As a matter of fact, they can no longer be considered as independent since a certain measurement outcome on system 1 allows a prediction about system 2, and vice versa. Such a correlation is called *entanglement* and the relative state is called *entangled state*. Although surprisingly, entangled states have been found experimentally, as we shall discuss extensively in the following.

Many forms of entanglement are in fact possible and many observables can be involved. A very important form of entanglement that will occupy us in the following is represented by the *singlet state*, an entanglement of the spin observables $\hat{S}_z$ (1.321) of two particles:

$$|\,\Psi_0\rangle := \frac{1}{\sqrt{2}}\left(|\uparrow\rangle_1 \otimes |\downarrow\rangle_2 - |\downarrow\rangle_1 \otimes |\uparrow\rangle_2\right), \tag{1.393}$$

where I have used the basis (1.322). Note that, due to antisymmetry, the components of the two subsystems are here antiparallel instead of being parallel as in the previous example. In fact, this equation tells us that when the spin of particle 1 is found in spin up along the conventional direction represented by z, particle 2 is in a spin down state or vice versa: in other words, the spin orientations of the two particles are opposite (Sect. 1.3.1). Note that this is true also for arbitrary directions. In fact, entanglement, although bearing some commonalities with superposition is not basis-dependent like the latter but is rather a property of the state, irrespective of the observables that we could employ to describe or to measure the system. Note also that states (1.392) or (1.393) are maximally entangled, where states of this kind with nD Hilbert-space dimensions have all diagonal elements of the (pure) density matrix equal to $1/n$. Summarising, pure states are of two sorts: product states and entangled states.

1.4.3 Total and Marginal States

Marginal States

As mentioned, in the case of composite systems the formalism of the density matrix is especially helpful. Suppose that you need to exclude one of two systems from your description. This happens quite commonly when we are interested only in one of the systems and consider the other as e.g. its environment. Let us therefore start with the density matrix

$$\hat{\rho}_{12} = |\,\Phi_{12}\rangle\,\langle\Phi_{12}\,|\,, \tag{1.394}$$

where the ket is given by Eq. (1.392), and assume that we wish to omit the consideration of system 2. This can be done by performing a *partial trace* over system 2 (denoted by Tr_2), that is, by summing the diagonal elements with respect to system 2 only, or equivalently, by summing the expectation values of $\hat{\rho}_{12}$ on a complete set of orthonormal basis states of system 2. The resultant *reduced density matrix* of system 1 contains information only about system 1 and is given by

$$\begin{aligned}\hat{\varrho}_1 &= \mathrm{Tr}_2\,\hat{\rho}_{12}\\ &= \left\langle h\left|\hat{\rho}_{12}\right|h\right\rangle_2 + \left\langle v\left|\hat{\rho}_{12}\right|v\right\rangle_2\\ &= \frac{1}{2}(|\,h\rangle\,\langle h\,|_1 + |\,v\rangle\,\langle v\,|_1).\end{aligned} \tag{1.395}$$

In fact, the expressions $\left\langle j\left|\hat{\rho}_{12}\right|j\right\rangle$, $j = h, v$, according to Eq. (1.373c), can be considered as mean values of the density matrix $\hat{\rho}_{12}$ on states $|\,h\rangle$, $|\,v\rangle$. It is evident that the reduced density matrix $\hat{\rho}_1$ describes a mixture. In other words, the reduced states of a composite system prepared initially in an entangled state is a mixture. A reduced state is also called *marginal* state.

Partial trace allows us to write the marginal state for system 1 given a certain measurement outcome of system 2, what is called *conditioned* state,

$$\mathrm{Tr}_2\left[\left(\hat{I}_1 \otimes \hat{P}_k\right)\hat{\rho}_{12}\right] = \hat{\rho}_{1|k}, \tag{1.396}$$

where $\hat{I}_1$ is the identity transformation on system 1, $\hat{\rho}_{1|k}$ is the state of system 1 conditional on the outcome k for system 2, and $k = h, v$ (but generalisation to higher dimension is straightforward).

Schmidt Decomposition

Let us now introduce an important formalism and consider a composite system made of two finite subsystems with the same dimension N, in the state

$$|\Psi\rangle = \sum_{j,k=1}^{N,N} C_{jk} \left|a_j\right\rangle |b_k\rangle \ , \tag{1.397}$$

where $\left\langle a_k \mid a_j \right\rangle = \left\langle b_k \mid b_j \right\rangle = \delta_{jk}$ and $C_{jk} \in \mathbb{C}$. It is interesting to note that it is always possible[146] to convert the double sum into a single sum, i.e. rewriting the previous equation as

$$|\Psi\rangle = \sum_{n=1}^{N} c_n |v_n\rangle |w_n\rangle \tag{1.398}$$

by means of the unitary transformations

$$|v_n\rangle = \sum_j U_{nj} \left|a_j\right\rangle \ , \ |w_n\rangle = \sum_k U_{nk}^{'} |b_k\rangle \ , \tag{1.399}$$

where substitution of the latter equations into the (1.398) gives

$$C_{jk} = \sum_n U_{nj}^{\mathrm{T}} c_n U_{nk}^{'} \ , \tag{1.400}$$

and $U_{nj}^{\mathrm{T}} = U_{jn}$. This is called *Schmidt decomposition*,, after the name of the German–Baltic mathematician Erhard Schmidt (1876–1959), while the number of non-zero coefficients c_n's is called the *Schmidt number*. The $c_n^2 \in \mathbb{R}$ are the nonvanishing eigenvalues of the Hermitian matrices $\hat{C}\hat{C}^\dagger$ and $\hat{C}^\dagger\hat{C}$ (see Eq. (1.82f)), whose set of eigenvectors are $\{|v_n\rangle\}$ and $\{|w_n\rangle\}$, respectively. In fact, in the 2D Hilbert space, an example is the matrix

$$\hat{C} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \ , \tag{1.401}$$

and it is clear that $\hat{C}$ is itself Hermitian, i.e. $\hat{C} = \hat{C}^\dagger$; then, we get

$$\hat{C}^2 = \frac{1}{2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \ , \tag{1.402}$$

so that the identity operator is the multiple of $\hat{C}^2$. Its eigenvalues are degenerate and both $1/2$. Since these are c_n^2, we have that $c_n = \pm 1/\sqrt{2}$. Moreover, we are free to choose the eigenvectors of $\hat{C}^2$, since any vector is eigenvector of the identity. In particular, I choose

[146] See Schmidt (1907) and also Peres (1995). I follow here Auletta et al. (2009, Sect. 5.5.3).

$$|v_+\rangle = |w_+\rangle = |+\rangle := \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right) , \tag{1.403a}$$

$$|v_-\rangle = |w_-\rangle = |-\rangle := \frac{1}{\sqrt{2}}\left(|0\rangle - |1\rangle\right) , \tag{1.403b}$$

where $|0\rangle , |1\rangle$ represent a basis that will play an important role in this book representing the quantum units of information (by now, let us consider them as a further example of basis). Also the states $|+\rangle , |-\rangle$ will occupy an important role in the following, although, in such a context, I prefer to still use $|v_\pm\rangle , |w_\pm\rangle$ for convenience. Therefore, we can rewrite Eq. (1.398) as

$$\begin{aligned} |\Psi\rangle &= \frac{1}{\sqrt{2}}\left(|v_+\rangle\,|w_+\rangle - |v_-\rangle\,|w_-\rangle\right) \\ &= \frac{1}{\sqrt{2}}\left[\frac{1}{2}\left(|0\rangle\,|0\rangle + |1\rangle\,|1\rangle + |0\rangle\,|1\rangle + |1\rangle\,|0\rangle\right) - \frac{1}{2}\left(|0\rangle\,|0\rangle + |1\rangle\,|1\rangle - |0\rangle\,|1\rangle - |1\rangle\,|0\rangle\right)\right] \\ &= \frac{1}{\sqrt{2}}\left(|0\rangle\,|1\rangle + |1\rangle\,|0\rangle\right) . \end{aligned} \tag{1.404}$$

From our calculations, the two unitary matrices are easily derived (they are the matrices whose column vectors are the two vectors $|v_n\rangle$ and $|w_n\rangle$), that is

$$\hat{U} = \hat{U}' = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} . \tag{1.405}$$

It is finally straightforward to verify that, given these matrices, (1.400) holds.

If $\{|a_j\rangle\}$ and $\{|b_k\rangle\}$ are two orthonormal bases for the two distinct Hilbert spaces, then, given the unitarity of the transformations $\hat{U}$ and $\hat{U}'$, also $\{|v_n\rangle\}$ and $\{|w_n\rangle\}$ are two orthonormal bases for the respective spaces. It is interesting to observe[147] that a (pure) state $|\Psi\rangle$ of a composite system is a product state iff it has a Schmidt number 1. Moreover, it is a product state iff the corresponding reduced density matrices of the two subsystems are pure.

It is possible to extend the above proof to the case of two systems with different dimensions and formulate the following theorem: for any pure state $|\Psi\rangle_{12}$ of a composite system $\mathcal{S}_{12}$ it is always possible to find orthonormal sets of vectors $\{|j\rangle_1\}$ for system $\mathcal{S}_1$ and $\{|j\rangle_2\}$ for system $\mathcal{S}_2$ such that

$$|\Psi\rangle_{12} = \sum_j c_j\,|j\rangle_1\,|j\rangle_2 , \tag{1.406}$$

where c_j are non-negative real numbers with $\sum_j c_j^2 = 1$.

If, $\{|a_j\rangle\}$ and $\{|b_k\rangle\}$ are two orthonormal bases for two distinct Hilbert spaces with different dimensions, then $\{|j\rangle_1\}$ and $\{|j\rangle_2\}$ are are two, possibly incomplete, sets of orthonormal vectors for these two spaces. It should be stressed that Schmidt

[147] See Nielsen and Chuang (2000, 110), D'Ariano et al. (2017, Sects. 6.3 and 7.3).

decomposition must *not* be understood as a way of *reducing* the total Hilbert space. As a matter of fact, the involved transformations are not universal but rather depend on the coefficients considered and, therefore, are state-dependent.

Pure States are Extensions of Mixtures

Previously, I have said that a partial trace of a pure density matrix yields a mixture. Here, I shall show how, starting with a mixed state, we can write down a compound pure state of which the 'initial' mixed state is a reduced state. Suppose, for instance, that we have a system $\mathcal{S}$ in a mixed state

$$\hat{\tilde{\rho}}_{\mathcal{S}} = \sum_n w_n \, |\, a_n \rangle \, \langle a_n \, | \ , \tag{1.407}$$

where $w_n \geq 0$ (and $w_n \neq 0$ for $k \geq 2$ values of n), and $\{|\, a_n \rangle\}$ is a basis for the Hilbert space $\mathcal{H}_{\mathcal{S}}$ of $\mathcal{S}$. Now it is always possible to find a larger system $\mathcal{S} + \mathcal{S}'$ (provided that $\mathcal{S}'$ has at least k independent vectors, with $k \geq 2$) such that the state vector of the total system $\mathcal{S} + \mathcal{S}'$ is of the form (1.398), i.e.

$$|\, \Psi \rangle_{SS'} = \sum_n \sqrt{w_n} \, |\, a_n \rangle \, |\, b_n \rangle \ , \tag{1.408}$$

where $\{|\, b_n \rangle\}$ is a basis for the Hilbert space $\mathcal{H}_{\mathcal{S}'}$ of $\mathcal{S}'$. It can be easily shown that, by tracing out $\mathcal{S}'$,

$$\begin{aligned} \mathrm{Tr}_{\mathcal{S}'} \left(|\, \Psi \rangle_{SS'} \, \langle \Psi \, | \right) &= \sum_{mn} \sqrt{w_m w_n} \, |\, a_m \rangle \, \langle a_n \, | \, \mathrm{Tr} \left(|\, b_m \rangle \, \langle b_n \, | \right) \\ &= \sum_{mn} \sqrt{w_m w_n} \, |\, a_m \rangle \, \langle a_n \, | \, \delta_{mn} \\ &= \hat{\tilde{\rho}}_{\mathcal{S}} \ , \end{aligned} \tag{1.409}$$

we exactly recover the state (1.407). In other words, we have first considered the Hilbert space $\mathcal{H}_{\mathcal{S}}$ of a system $\mathcal{S}$ in a mixed state. Then, we have considered a second system $\mathcal{S}'$ with a Hilbert space $\mathcal{H}_{\mathcal{S}'}$ and built a compound system of these two in a larger Hilbert space $\mathcal{H}_{\mathcal{S}+\mathcal{S}'} = \mathcal{H}_{\mathcal{S}} \otimes \mathcal{H}_{\mathcal{S}'}$, which is described as pure state. Finally, we have traced out the system $\mathcal{S}'$. This shows that a mixed state of a certain system can always be obtained as a partial trace of a pure state of a larger system, a procedure that is called purification.

1.5 Summary

In this chapter, we have explored some basic concepts. I recall here in particular the following:

- The notion of self-interference: quantum systems make interference with themselves, what hinders to represent them as kinds of classical waves or classical fields.
- The notion of discontinuity, with its most relevant consequences: non-commutativity of some observables and uncertainty relations. Observables that do not commute are not jointly measurable.
- As a consequence, quantum systems do not satisfy the classical *omnimoda determinatio*. This shows that quantum systems, even when localised, cannot be understood as kinds of classical particles.
- Another crucial aspect is that all quantum systems obey the Schrödinger equation that deterministically rules the evolution in time of probability amplitudes. The probability distributions of observables can be computed on such a basis.
- All basic transformations are unitary or anti-unitary, and therefore they give rise to symmetries and groups.
- Superposition is a relative concept and therefore there are unitary transformations allowing to express the same state in different bases. In the case in which we deal with continuous observables, we need to use Fourier transforms.
- There are two different kinds of statistics, one for bosons (symmetric under permutation of coordinates) and one for fermions (antisymmetric under permutation of coordinates).
- When the Planck constant becomes growingly irrelevant with the increasing size of the systems, we attain the classical limit that marks the transition to the classical equations of motion.
- Another consequence is complementarity. This was understood by Bohr as a classical complementarity between waves and particles. However, as a consequence of self-interference and quantisation we cannot follow such a classical interpretation.
- Quantum systems can be represented by density operators and they are either pure states or mixtures, and the latter are weighted sums of the former.
- We have seen that quantum systems can be entangled.
- The Schmidt decomposition has been presented, showing that partial traces of pure states yield mixed reduced states while extension of mixtures yields pure global states.

References

Anscombe, G., and M. Elizabeth. 1971. *Causality and Determination*. Cambridge: Cambridge University Press; rep. in [Sosa and Tooley 1993, 88–104].

Apostol, Tom M. 1969. *Calculus*. New York: Wiley, 2nd ed.
Aristotle. 1831. *Opera* (Ed. Immanuel Bekker). Berlin: G. Reimerum.
Aristotle. 1950. *Physica*. Oxford: Clarendon, 1950, 1988; also in [Aristotle Op., 184–267].
Arnold, Vladimir I. 1978. *Mathematical Methods of Classical Mechanics* (tr. from russ.). New York: Springer, 1978; 2nd ed.: 1989.
Auletta, Gennaro. 2000. *Foundations and Interpretation of Quantum Mechanics: In the Light of a Critical-Historical Analysis of the Problems and of a Synthesis of the Results*. Singapore: World Scientific, 2000; rev. ed. 2001.
Auletta, Gennaro. 2004. Critical Examination of the Conceptual Foundations of Classical Mechanics in the Light of Quantum Physics. *Epistemologia* 27: 55–82.
Auletta, G., Fortunato, M., and Parisi, G. 2009. Quantum Mechanics, Cambridge, University Press, rev. paperback ed. 2013.
Auletta, Gennaro. 2011. *Cognitive Biology: Dealing with Information from Bacteria to Minds*. Oxford: University Press.
Auletta, G., and S.-Y. Wang. 2014. *Quantum Mechanics for Thinkers*. Singapore: Pan Stanford Publisher. https://doi.org/10.4032/9789814411721.
Baumgarten, Alexander Gottlieb. 1739. *Metaphysica* 1739; 7th ed., Halle 1779; rep. Hildesheim–New York, Georg Olms, 1982.
Bohr, Niels. 1913. On the Constitution of Atoms and Molecules. *Philosophical Magazine* 26: 1–25, 476–502, 857–875; rep. in [Bohr CW, II, 161–233].
Bohr, Niels. 1928. The Quantum Postulate and the Recent Development of Atomic Theory. *Nature* 121: 580–590; rep. in [Bohr CW, VI, 148–58].
Bohr, Niels. 1949. Discussion With Einstein on Epistemological Problems in Atomic Physics. in [Schilpp 1949, 201–241].
Bohr, Niels. 1996. *Collected Works*. Amsterdam: North Holland.
Bokulich, Alisa. 2008. *Reexamining the Quantum-Classical Relation: Beyond Reductionism and Pluralism*. Cambridge: University Press.
Born, Max. 1926. Quantenmechanik der Stoßvorgänge. *Zeitschrift für Physik* 38: 803–827.
Born, Max. 1949. *Natural Philosophy of Chance and Cause*. Oxford: Clarendon, 1949; New York, 1964.
Born, M., and P. Jordan. 1925. *Zur Quantenmechanik. Zeitschrift für Physik* 34: 858–888.
Born, M., W. Heisenberg, and P. Jordan. 1926. *Zur Quantenmechanik II. Zeitschrift für Physik* 35: 557–615.
Boscovich, Rogerius J. 1754. *De continuitatis lege and ejus consectariis, pertinentibus ad prima materiae elementa eorumque vires*. Roma, Collegio Romano.
Brown, Laurie M. (ed.). 2005. *Feynman's Thesis: A New Approach to Quantum Theory*. Singapore: World Scientific.
Byron Jr., F.W., and R.W. Fuller. 1969–70. *Mathematics of Classical and Quantum Physics*. New York: Dover.
Cohen-Tannoudji, Gilles. 1991. *Les constantes universelles*. Paris: Hachette.
Compton, Arthur H. 1923. A Quantum Theory of the Scattering of X-Rays by Light Elements. *Physical Review* 21: 483–502.
D'Ariano, M., G. Chiribella, and P. Perinotti. 2017. *Quantum Theory from First Principles: An Informational Approach*. Cambridge: University Press.
Davisson, C.J., and L.H. Germer. 1927. Diffraction of Electrons By a Crystal of Nickel. *Physical Review* 30: 705–740.
de Broglie, Louis. 1924. Sur la définition générale de la correspondance entre onde et mouvement. *Comptes Rendus à l'Academie des Sciences* 179: 39–40.
Descartes, René. 1641. *Meditationes de prima philosophia*. Paris; in: [Descartes Œ, VII].
Descartes, René, Œuvres (Ed. C. Adam/ P. Tannery), Paris, Vrin, 1907–1913; 1964–74
D'espagnat, Bernard., (ed.). 1995. *Veiled Reality*. Reading, MA: Addison/Wesley.
Dieks, Dennis. 1994. Modal Interpretation of Quantum Mechanics, Measurements, and Macroscopic Behavior. *Physical Review* A49: 2290–2299.

Dirac, Paul A.M. 1930. *The Principles of Quantum Mechanics*, 4th ed. Oxford: Clarendon, 1958, 1993.
Dirac, Paul A.M. 1933. The Lagrangian in Quantum Mechanics. *Physikalische Zeitschrift der Sowjetunion* 3.1: 64–72; rep. in [Brown 2005, 113–21].
Dirac, Paul A.M. 1939. A New Notation for Quantum Mechanics. *Mathematical Proceedings of the Cambridge Philosophical Society* 35: 416–418. https://doi.org/10.1017/S0305004100021162.
Eddington, Arthur Stanley. 1929. *Nature of the Physical World*. Cambridge: University Press.
Einstein, Albert. 1905. Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt. *Annalen der Physik* 17: 132–148.
Einstein, Albert. 1909. Über die Entwicklung unserer Anschauungen über das Wesen und die Konstitution der Strahlung. *Physikalische Zeitschrift* 10: 817–825.
Einstein, Albert. 1916. Strahlungs-emission und -absorption nach der Quanten-theorie. *Deutsche Physikalische Gesellschaft, Verhandlungen* 18: 318–323.
Einstein, Albert. 1993. *Mein Weltbild*. Ullstein: Frankfurt.
Englert, B.-G., M.O. Scully, and H. Walther. 1994. The Duality in Matter and Light. *Scientific American* 271 (6): 56–61.
Franson, J.D., and K.A. Potocki. 1988. Single-Photon Interference Over Large Distances. *Physical Review* A37: 2511–2515.
Fuchs, J., and C. Schweigert. 1997. *Symmetries, Lie Algebras and Representations*. Cambridge: University Press.
Gerlach, W., and O. Stern. 1922a. Der experimentelle Nachweis des magnetischen Moments des Silberatoms. *Zeitschrift für Physik* 8: 110–111.
Gerlach, W., and O. Stern. 1922b. Der experimentelle Nachweis der Richtungsquantelung im Magnetfeld. *Zeitschrift für Physik* 9: 349–352.
Gerlach, W., and O. Stern. 1922c. Das magnetische Moment des Silberatoms. *Zeitschrift für Physik* 9: 353–355.
Grangier, P., G. Roger, and A. Aspect. 1986. Experimental Evidence for a Photon Anticorrelation Effect on a Beam Splitter: A New Light on Single-Photon Interferences. *Europhysics Letters* 1: 173–179.
Heisenberg, Werner. 1925. Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen. *Zeitschrift für Physik* 33: 879–893; rep. in [Heisenberg GW, A.I, 382–96].
Heisenberg, Werner. 1927. Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. *Zeitschrift für Physik* 43: 172–198; rep. in [Heisenberg GW, A.I, 478–504].
Heisenberg, Werner. 1969. *Der Teil und Das Ganze*. München: Piper.
Heisenberg, Werner. 1984. *Gesammelte Werke*. Piper/Springer: München-Berlin.
Home, D., and M.A.B. Whitaker. 2007. *Einstein's Struggles with Quantum Theory: A Reappraisal*. Berlin: Springer.
Jammer, Max. 1966. *The Conceptual Development of Quantum Mechanics*. New York, McGraw-Hill, 1966; 2nd ed., Thomas Publisher, 1989.
Kant, Immanuel. 1763. *Der einzige mögliche Beweisgrund zu einer Demonstration des Daseins Gottes*; in [Kant A, I, 1–182].
Kant, Immanuel, and Kants Werke. 1968. *Akademie Textausgabe*. Berlin: W. de Gruyter.
Kastner, Ruth E. 2013. *The Transactional Interpretation of Quantum Mechanics: The Reality of Possibility*. Cambridge: Cambridge University Press.
Kuhn, Thomas S. 1978. *Black-Body Theory and the Quantum Discontinuity. 1894–1912*. Oxford, University Press, 1978; Chicago, The University of Chicago Press, 1987.
Lagrange, Joseph-Louis. 1788–89. *Mécanique analytique*. Paris.
Landau, Lev D. 1927. Das Dämpfungsproblem in der Wellenmechanik. *Zeitschrift für Physik* 45: 430–441.
Landau, Lev D., and E.M. Liftshitz. 1976. *Mechanics* (eng. tr.), I vol. of *The Course of Theoretical Physics*. Oxford: Pergamon.
Landsman, Klaas. 2017. *Foundations of Quantum Theory: From Classical Concepts to Operator Algebras*. Berlin: Springer.

Leibniz, Gottfried. 1686. Brevis Demonstratio Erroris Mirabilis Cartesii et aliorum .... *Acta Eruditorum Lipsiensium*, in [Leibniz MS, VI, 117–23].
Leibniz, Gottfried W. 1695. Specimen Dynamicum pro admirandis NaturæLegibus circa corporum vires et mutuas actiones detegendis et ad suas causas revocandis. Pars I. *Acta Eruditorum Lipsiensium*; rep. in [Leibniz MS, VI, 234–46].
Leibniz, Gottfried W. 1712–14. Monadologie in [Leibniz PS, VI, 607–623].
Leibniz, Gottfried W. 1715–16. Streitschriften zwischen Leibniz und Clarke [Leibniz PS, VII, 345–440]. Published before as *A Collection of Papers, which passed between the late Learned Mr. Leibniz and Dr. Clarke in the Years 1715 and 1716, relating to the Principles of Natural Philosophy and Religion*, London 1717.
Leibniz, Gottfried W. 1971. *Mathematische Scriften* (Ed. Gerhardt), Halle, 1860; rep. Hildesheim, Olms, 1971. (NEA)
Leibniz, Gottfried W. 1978. Philosophische Schriften (Ed. Gerhardt), Halle, 1875. *rep*. Olms: Hildesheim.
Margenau, Henry. 1950. *The Nature of Physical Reality: A Philosophy of Modern Physics*. New York: McGraw-Hill, 1950; Woodbridge, Conn.: Ox Bow Press.
Mehra, J., and H. Rechenberg. 1982–2001. *The Historical Development of Quantum Theory*. New York: Springer.
Misner, C.W., K.S. Thorne, and J.A. Wheeler. 1970. *Gravitation 1970, 1973*. San Francisco: Freeman and Co.
Newton, Isaac. 1704. *Opticks or A Treatise of the Reflections, Refractions, Inflections and Colours of Light*, 1704, 2nd ed. lat.; 1706, 3rd ed. 1721, London, 4th ed. 1730; New York, Dover, 1952, 1979.
Nielsen, M.A., and I.L. Chuang. 2000. *Quantum Computation and Quantum Information*. Cambridge: University Press 2002, 2011.
Pauli, Wolfgang. 1925. Über den Zusammenhang des Abschlusses der Elektronengruppen im Atom mit der Komplexstruktur der Spektren. *Zeitschrift für Physik* 31: 765–783.
Pauli, Wolfgang. 1927. Zur Quantenmechanik des magnetischen Elektrons. *Zeitschrift für Physik* 43: 601–623.
Peirce, Charles S. 1877. On a New Class of Observations, Suggested by the Principles of Logic, in [Peirce W, III, 235–37].
Peirce, Charles S. 1892. The Doctrine of Necessity Examined *Monist*2321–37; in [Peirce W, VIII, 111-125].
Peirce, Charles S. 1903. The Seven Systems of Metaphysics in [Peirce EP, II, 179–95].
Peirce, Charles S. 1982. *Writings, vols*. I-: Bloomington, Indiana University Press.
Peirce, Charles S. 1998. *The Essential Peirce, vols*. I-II: Bloomington, Indiana University Press.
Penrose, Roger. 2004. *The Road to Reality: A Complete Guide to the Laws of the Universe* J. Cape, 2004; London: Vintage.
Peres, Asher. 1995. Higher Order Schmidt Decompositions. *Physics Letters A* 202: 16–17.
Pfleegor, R.L., and L. Mandel. 1967a. Interference Effects at the Single Photon Level. *Physics Letters* 24A: 766–767.
Pfleegor, R.L., and L. Mandel. 1967b. Interference of Independent Photon Beams. *Physical Review* 159: 1084–1088.
Planck, Max. 1900a. Über die Verbesserung der Wien'schen Spektralgleichung. *Verhandlungen der Deutschen Physikalischen Gesellschaft* 2: 202–204.
Planck, Max. 1900b. Zur Theorie des Gesetzes der Energieverteilung im Normalspektrum. *Verhandlungen der Deutschen Physikalischen Gesellschaft* 2: 237–245.
Planck, Max. 1906. *Vorlesungen über die Theory der Wärmestrahlung*. Leipzig: Barth, 1906; 2nd ed., 913.
Planck, Max. 1922. *The Origin and Development of the Quantum Theory*. Oxford: Clarendon Press.
Planck, Max. 1931. *The Universe in the Light of Modern Physics*. London: George Allen.
Poincaré, Henri. 1902. *La science et l'hypothèse*, Paris, 1902; Flammarion, 1968.
Poincaré, Henri. 1905. *La Valeur de la science*. Paris: Flammarion.

Prugovečki, Eduard. 1971. *Quantum Mechanics in Hilbert Space*, 1971; 2nd ed.: New York: Academic Press, 1981.
Renninger, Mauritius K. 1960. Messungen ohne Störung des Meßobjekts. *Zeitschrift für Physik* 158: 417–421.
Rindler, Wolfgang. 2001. *Relativity: Special, General, and Cosmological*. Oxford: University Press.
Robertson, H.P. 1929. *The Uncertainity Principle. Physical Review* 34: 163–164.
Schilpp, Arthur (Ed.) 1949. *Albert Einstein. Philosopher-Scientist*. La Salle (Illinois), Open Court, 1949, 3rd ed. 1988.
Schlosshauer, Maximilian. 2007. *Decoherence and the Quantum-to-Classical Transition*. Berlin: Springer.
Schmidt, Erhard. 1907. Zur Theorie des linearen und nichtlinearen Integralgleichungen. I Teil: Entwicklung willkürlicher Funktionen nach Systemen vorgeschriebener. *Mathematische Annalen* 63: 433–476.
Schrödinger, Erwin. 1926a. Quantisierung als Eigenwertproblem I. *Annalen der Physik* 79: 361–376; rep. in [Schrödinger GA, III, 82–97].
Schrödinger, Erwin. 1926b. Quantisierung als Eigenwertproblem II. *Annalen der Physik* 79: 489–527; rep. in [Schrödinger GA, III, 98–136].
Schrödinger, Erwin. 1926c. Quantisierung als Eigenwertproblem III. *Annalen der Physik* 80: 437–90; rep. in : [Schrödinger GA, III, 166–219].
Schrödinger, Erwin. 1926d. Quantisierung als Eigenwertproblem IV. *Annalen der Physik* 81: 109–139; rep. in [Schrödinger GA, III, 220–250].
Schrödinger, Erwin. 1926e. Über das Verhältnis der Heisenberg-Born-Jordanschen Quantenmechanik zu der meinen. *Annalen der Physik* 79: 734–756; rep. in [Schrödinger GA, III, 143-65]
Schrödinger, Erwin. 1984. *Gesammelte Abhandlungen*. Wien: Verlag der österreichischen Akademie der Wissenschaften.
Schwartz, Matthew D. 2014. *Quantum Field Theory and the Standard Model*. Cambridge: University Press.
Sosa, E. and Tooley, M. (Eds.) 1993. *Causation*. Oxford University Press.
'T Hooft, Gerard., 2016. *The Cellular Automaton Interpretation of Quantum Mechanics*. Berlin: Springer.
Tinkham, Michael. 1964. *Group Theory and Quantum Mechanics*. New York: Dover; 2nd ed. 1992.
van Fraassen, Bas C. 1991. *Quantum Mechanics. An Empiricist View*, 1995. Oxford: Clarendon.
Wheeler, John A. 1988. World as System Self-Synthesized by Quantum Networking. *IBM Journal of Research and Development* 32: 4–15. https://doi.org/10.1147/rd.321.0004.
Wigner, Eugene P. 1959. *Group Theory*. New York: Academic.
Wilczek, Frank. 2006. *Fantastic Realities: 49 Mind Journeys and a Trip to Stockholm*. Singapore: World Scientific.

Chapter 2
The Main Problems



Randomness is where reason stops, it's a statement that things are accidental, meaningless, unpredictable, and happen for no reason.

Gregory J. Chaitin, *The Unknowable*

QM could be considered a cluster of riddles and the world that it describes as a kind of impenetrable sphinx. As a matter of fact, the best minds of several generations have tried to cope with these mysteries, and consistent advancements have been made step by step although to date no congruent general interpretation that would capture the agreement of the physicists' community has been provided. Here, I schematically summarise some of the main problems partitioning the whole in four broad subjects: significance of the quantum formalism, measurement, non-locality and causality. The possible solutions to these problems will be treated in the next chapters.

2.1 Quantum-Mechanical Formalism

2.1.1 Two Pictures of Quantum Dynamics

When dealing with QM, the major issue is likely the relation between theory and reality, a problem by which Einstein was very much concerned in general terms. The problem is big and can be considered to be the very object of this work. Since it presents many aspects and subtleties, it needs to be aggressed from different points of view and step by step. Here, I like to examine one of the most basic aspects, namely, the status of the mathematical formalism.

© Springer Nature Switzerland AG 2019

G. Auletta, *The Quantum Mechanics Conundrum*,
https://doi.org/10.1007/978-3-030-16649-6_2

Objectivism

As for every other physical theory, the formalism of QM (i.e. its laws or fundamental equations) is (are) addressed to describe determinate classes of physical systems and especially their evolution in time when dealing with specific classes of contexts (like different kinds of interaction, free motion, etc.). According to the influential German physicist and philosopher of science G. Ludwig (1918–2007), a physical theory has: a mathematical description (physical laws or laws of nature), an ontological domain and application prescriptions (connected with experimental procedures), where the latter two depend on the former.[1] Focussing on the ontological domain, it is natural to assume that the quantum formalism does describe physical objects, since it appears difficult to make any reality ascription at a physical level without assuming that we deal with some kinds of objects. On the other hand, if no reality ascription is possible, the relation between mathematics and reality breaks down: there was no reality or our formalism was only a construction of the human mind. Now, there is a minimal requirement of objectivity that is true even for mathematics since it is a basic character of scientific activity. When the scientist in whatever field has scrupulously followed the methods and procedures of his/her field and faithfully taken into account data and facts, he/she will always ascribe an error to those procedures. In other words, the cornerstone of scientific methodology is to assume that facts and data (or results) are stronger than whatever explanation we might give of them. And this cannot work without, at least implicitly, assuming that what we try to describe or explain has an objectivity that, at least partly, is independent of our theories.

Ontology

What about physics? Here, we could assume a reduced form of objectivism that bypasses the problem of ontological ascription, being satisfied with the capability of the formalism to make certain specific predictions.[2] However, as recalled by the Israeli–British physicist David Deutsch, no scientific theory can circumvent the necessity to provide satisfactory *explanations* of what it aims to describe[3]: the models of science are intended to be "maps" of the real world, in the sense that they *describe* the world, although not in all its details (otherwise the model would be totally unhelpful),[4] making it intelligible to us. Therefore, we expect from a physical theory, as being different from pure mathematics as well as from practical activities like, e.g. weather forecast, the capability to make ontological ascriptions: namely,

[1] Ludwig (1978, pp. 8–9). Physical reality also presents a fundamental domain (*Grundbereich*) that is relatively independent of the application prescriptions as well as of the mathematics.

[2] This has been clearly the standpoint of the young Heisenberg: see Heisenberg (1927).

[3] It may be discussed in which sense causal aspects need to be involved here (Margenau 1950, Sect. 8.1).

[4] Deutsch (1997, 2011). The same point is supported in Wallace (2012, Sect. 1.1). For the notion of scientific models as maps see Auletta (2011a, Sect. 2.1.5) and literature quoted there.

we expect that each physical theory defines and precisely delimitates the ontological status of the entities (like photons, electrons, fields) that it introduces or that is assumed to be the referent of the theory,[5] where with *ontological status* I understand in a first approximation the specification of the kind of objects our theory deals with.

Now, when we consider QM, several problems emerge. Namely, if quantum systems can display non-local aspects as a consequence of superposition and entanglement (Sects. 1.2.1 and 1.4.2), how can we say that they are real objects at all? In fact, we cannot conceive of a system that does not take a certain "location" somewhere in Nature that does not ideally occupy a well-limited region of space-time, according to Einstein's understanding of relativity. And this is to a certain extent also true of fields (at least of their sources). Moreover, as we have seen, many quantum systems are indiscernible (Sect. 1.3.1): in which sense can we speak of different systems if they need to be considered identical? Finally, it seems that measurement alters the properties of a quantum system. In fact, we expect that any system behaves classically when measured (that it gives a classical output), also when we have an initial superposition state, which by definition is not classical. We shall come back to this in the next section, but, in a first approximation, the problem is whether the state of the object system is a superposition or an eigenstate seems to depend on the observable that we are selecting (Sects. 1.2.2–1.2.3). Now, if the properties that we attribute to the system depend on our description or operations, how can we ascribe any objective reality to it? Here, in a first approximation, I understand with *property* the value that we can attribute to an observable, i.e. the eigenvalue of the corresponding operator, when the system is in a certain state.

Extracting Information

Here, we meet a further problem: CM treated variables (the counterpart of quantum-mechanical observables) and states (being a state defined through its variables and therefore being a kind of variable itself) on the same foot. Quantum-mechanically this is not the case. Now, it is true that also in QM the state can be considered an observable, being mathematically represented by a Hermitian operator: it either is a pure state (and is described by a single projector) or a mixture (in such a case is described by a weighted sum of projectors), as discussed in Sect. 1.4.1.[6] If we hypothesise that it is represented by a single projector, by measuring it we are not able to determine the state, since such a measurement corresponds to a test through which we ask the system whether or not it is in the state described by the projector (and the two alternatives define the test). In fact, projectors, independently of the dimensions of the system, are binary operators (Sect. 1.2.2). In other words, here we are not asking what is the state of the system, which is precisely the kind of question that, if answered, would allow us to extract the whole information from the system's state. At the opposite, we get either (i) yes, the system is in the state

[5] This has been pointed out already in Auletta (2000, Sect. 6.2). See also 'T Hooft (2016, p. 34).

[6] I summarise here the arguments presented, e.g. in Auletta and Wang (2014, Sects. 3.5–3.7 and 7.8).

you are asking about or (ii) no, the system is not in that state. To understand this difference, a classical example could be helpful. It is much easier to ascertain whether or not a certain object is an apple (the answer is indeed also here either yes or not) than to give a full description of the object (here the answer takes the form of a list of properties). On the other hand, if the state is a mixture, to measure a single system would mean again to get a result described by one of the projectors summed (each with its weight) in the mixture. The outcome of this measurement would again provide partial information. We have therefore evidence that the state of a quantum system cannot be determined through a single measurement; as a consequence, we cannot distinguish between non-orthogonal states with a single measurement as well as we cannot perfectly discriminate between two non-orthogonal states in general. This is an issue closely related with the no-cloning theorem,[7] which will be dealt with later.

In general, when measuring a certain observable, we only obtain one of its eigenvalues but we cannot ascertain the probability distribution of its possible properties in the state being measured if not known beforehand (it would be known through a procedure called preparation). Thus, this probability distribution remains unknown to us through a single outcome. However, as we shall see with more details in the following, it is such a probability distribution that would represent the complete quantum description of the state. In fact, in the discrete case (to keep things as simple as possible), for any arbitrary observable described by Eq. (1.96) and any arbitrary state $|\psi\rangle$, whose Hilbert-space dimension is the same, thanks to the resolution of identity (1.94), we can always write

$$|\psi\rangle = \hat{I}\,|\psi\rangle = \sum_j \hat{P}_j\,|\psi\rangle = \sum_j |o_j\rangle\langle o_j \mid \psi\rangle = \sum_j \langle o_j \mid \psi\rangle\,|o_j\rangle\,, \tag{2.1}$$

where the $\langle o_j \mid \psi\rangle$'s are the (in general, complex) probability amplitudes that would allow us a full reconstruction of the probability distribution of the measurement outcomes for that observable. We are certainly able to attribute a specific property given some measurement, which is an item of the list of properties that would constitute the complete knowledge of the state, but, as I shall argue, the problem is that we cannot consistently progress and cumulate the knowledge through different measurements, so that this limitation cannot be circumvented. Thus, we have a fundamental difference relative to CM.

Passive and Active Transformations

For this reason, QM distinguishes between the change of *state* and the change of *observable*, a distinction that makes no strict sense in CM. Somebody could see

[7] For the original papers on the no-cloning theorem see Wootters and Zurek (1982), Dieks (1982). For further extensions see D'ariano and Yuen (1996), Lindblad (1999). On the issue of state discrimination see D'ariano et al. (2017, Sect. 2.8.6).

in this an imperfection of the theory. Quite the opposite: it is a manifestation of its strength insofar it adds a new conceptual (and mathematical) dimension to our description of physical systems and therefore allows for distinctions that did not arise in a classical context.[8] These two different kinds of time evolution are called the *Schrödinger picture* (already discussed in Sect. 1.1.3) and the *Heisenberg picture*, respectively. Actually, as mentioned there, they are equivalent formulations, so that the fundamental tenet of CM is also satisfied, although in a different way. This means that we can pass from one to another thanks to a unitary transformation, which we know to express the notion of reversibility in QM (Sect. 1.2.5).[9] Before penetrating into the ontological issue it is crucial to consider this double dynamics from a formal point of view.

All of these time evolutions can be represented as rotations of some vectors in the Hilbert space, where I recall that states are represented by state vectors. In particular,

- An evolution in time of a certain state (according to the Schrödinger picture), called *active* transformation, can be considered as a rotation of the corresponding state vector in the Hilbert space, whilst
- The corresponding time evolution of one of its observables (according to the Heisenberg picture), called *passive* transformation, is the opposite rotation (in the opposite sense) of the vectors representing the basis (i.e. the possible outcomes that we would obtain if we performed a measurement) of that particular observable; obviously, different bases are possible according to what observable we are considering or measuring. Therefore, the Heisenberg picture represents a change of basis (Sect. 1.2.2).

An Example

Consider Fig. 2.1. Let us call γ the angle between $|\psi\rangle$ and $|x_2\rangle$ and β the angle between $|\psi'\rangle$ and $|x_1'\rangle$. Then, we can write the initial state as

$$|\psi\rangle = \cos(2\alpha + \beta)\,|x_1\rangle + \sin\gamma\,|x_2\rangle\ . \tag{2.2}$$

The state vectors $|x_1\rangle$ and $|x_2\rangle$ could be, e.g. eigenstates of the 1D 'position' observable

$$\hat{x} = x_1\,|x_1\rangle\,\langle x_1| + x_2\,|x_2\rangle\,\langle x_2|\,, \tag{2.3}$$

where, for the sake of simplicity, only two positions are considered. Then, Eq. (2.2) could be considered a discrete counterpart (in 2D Hilbert space) of Eq. (1.120). Now, according to the Schrödinger picture (see Eq. (1.15)), at a later moment t, the state of the system will be evolved to a state $|\psi(t)\rangle = |\psi'\rangle$:

[8] As recalled in Auletta and Wang (2014, Sect. 7.4).

[9] The reader can see a more complete exposition in Auletta et al. (2009a, Chaps. 2–3). For a shorter summary see Auletta and Wang (2014, Sect. 4.6).

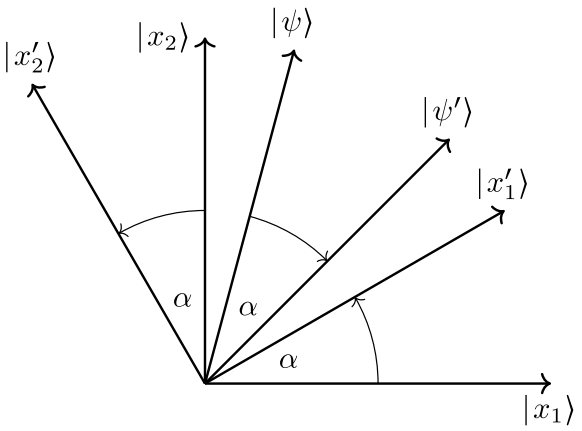


Fig. 2.1 Active and passive transformations. A transformation may be considered from two equivalent viewpoints: from the *active* point of view, the state vector $|\psi\rangle$ is transformed (here represented by a clockwise rotation by an angle α) into the state vector $|\psi'\rangle$ while the basis vectors $|x_1\rangle$ and $|x_2\rangle$ are kept fixed; from the *passive* point of view, the basis vectors $|x_1\rangle$ and $|x_2\rangle$ are transformed in a reverse manner (here represented by a counterclockwise rotation by the same angle α) to the basis vectors $|x_1'\rangle$ and $|x_2'\rangle$, respectively. Note that the original state $|\psi\rangle$ in the new basis $\{|x_1'\rangle, |x_2'\rangle\}$ is equivalent to the transformed state $|\psi'\rangle$ in the old basis $\{|x_1\rangle, |x_2\rangle\}$

$$\hat{U}_t\,|\psi\rangle = |\psi'\rangle = \cos(\alpha+\beta)\,|x_1\rangle + \sin(\alpha+\gamma)\,|x_2\rangle\ , \tag{2.4}$$

where the coefficients have changed (due to the rotation of the state vector in the Hilbert space). Now, in the Heisenberg picture, we can write the initial state (which is therefore kept fixed) in terms of the rotated basis $\{|x_1'\rangle, |x_2'\rangle\}$ (for instance, these two state vectors could be eigenstates of momentum), so that we obtain the same transformation in the latter picture:

$$|\psi\rangle = \cos(\alpha+\beta)\,|x_1'\rangle + \sin(\alpha+\gamma)\,|x_2'\rangle\ , \tag{2.5}$$

which shows that the state vector $|\psi\rangle$ in the basis $\{|x_1'\rangle, |x_2'\rangle\}$ is equivalent to the state vector $|\psi'\rangle$ in the basis $\{|x_1\rangle, |x_2\rangle\}$. Note indeed that it is a single unitary operator and its inverse that gives rise to passive and active transformations.

The Two Pictures

Let us now go behind the example and deal more generally with the two different pictures. I recall that, when the state is known, QM does not predict the results of individual measurements, but only their statistical mean, i.e. the weighted average of all possible measurement outcomes. Let $\hat{O}$ be an observable with eigenvalues o_j's described by Eq. (1.96). Its expectation value for a quantum system in the state $|\psi\rangle$ is given by the statistical mean (1.102), so that we can write the expectation value of

$\hat{O}$ in the state $|\psi(t)\rangle$ at time t in the Schrödinger picture as[10]

$$\langle \hat{O} \rangle_{\psi(t)} = \left\langle \psi(t) \left| \hat{O} \right| \psi(t) \right\rangle . \tag{2.6}$$

By making use of Eq. (1.304) with $t_0 = 0$, we have

$$\left\langle \psi(t) \left| \hat{O} \right| \psi(t) \right\rangle = \left\langle \psi(0) \left| \hat{U}_t^\dagger(t)\, \hat{O}\, \hat{U}_t(t) \right| \psi(0) \right\rangle , \tag{2.7}$$

which allows us to displace the action of the time evolution operator from the state vector to the observable, and therefore to go over to the Heisenberg picture. Starting from the latter equation, define the observable $\hat{O}(t)$ and the state vector $|\psi\rangle$ in the Heisenberg picture as

$$\hat{O}(t) := \hat{U}_t^\dagger(t)\, \hat{O}\, \hat{U}_t(t), \tag{2.8a}$$

$$|\psi\rangle := |\psi(0)\rangle = \hat{U}_t^\dagger(t)\, |\psi(t)\rangle , \tag{2.8b}$$

respectively, where for the latter equation I have used Eqs. (1.306)–(1.307). The first equation represents in fact the evolution in time of the observable $\hat{O}$, while in the LHS of the last one I have dropped the reference to time $t_0 = 0$ because we are considering the state as static in such a picture. In other words, the time evolution of an observable in the Heisenberg picture is precisely expressed as a time-dependent change of basis. Note that since an observable is represented by an operator, its time evolution is determined by both an unitary operator and its adjoint. By taking into account Eq. (2.8a), we can write

$$\left\langle \psi(0) \left| \hat{U}_t^\dagger(t)\, \hat{O}\, \hat{U}_t(t) \right| \psi(0) \right\rangle = \left\langle \psi \left| \hat{O}(t) \right| \psi \right\rangle . \tag{2.9}$$

Comparing Eqs. (2.7) and (2.9), we derive the relation

$$\left\langle \psi(t) \left| \hat{O} \right| \psi(t) \right\rangle = \left\langle \psi \left| \hat{O}(t) \right| \psi \right\rangle , \tag{2.10}$$

which establishes the equivalence between the Schrödinger (LHS) and Heisenberg (RHS) pictures. Moreover, this generalisation allows us to write the equivalent of the Schrödinger equation in the Heisenberg picture by taking the time derivative of $\hat{O}(t)$ and by making use of the following reformulation of the Schrödinger equation (1.15)

$$\frac{d\hat{U}_t(t)}{dt} = \frac{d}{dt} e^{-\frac{i}{\hbar}\hat{H}t} = -\frac{i}{\hbar}\hat{H}\hat{U}_t(t) = -\frac{i}{\hbar}\hat{U}_t(t)\hat{H}, \tag{2.11}$$

where I have displaced the time derivative from the ket to the unitary operator in order to be able to deal with the time evolution of the observable (see also Table 1.1,

[10] See Auletta et al. (2009a, Sect. 3.6). Auletta and Wang (2014, Sect. 7.5).

Sect. 1.1.3). Here, I have made use of the fact that $\hat{H}$ and $\hat{U}_t(t)$ commute. From Eqs. (2.8a) and (2.11), the time derivative of $\hat{O}(t)$ is found to be given by

$$\begin{aligned}\frac{d\hat{O}(t)}{dt} &= \frac{d}{dt}\left[\hat{U}_t^\dagger(t)\,\hat{O}\,\hat{U}_t(t)\right] \\ &= \frac{d\hat{U}_t^\dagger(t)}{dt}\,\hat{O}\,\hat{U}_t(t) + \hat{U}_t^\dagger(t)\,\hat{O}\,\frac{d\hat{U}_t(t)}{dt} \\ &= \frac{i}{\hbar}\hat{H}\hat{U}_t^\dagger(t)\,\hat{O}\,\hat{U}_t(t) - \frac{i}{\hbar}\hat{U}_t^\dagger(t)\,\hat{O}\,\hat{U}_t(t)\hat{H} \\ &= \frac{i}{\hbar}\left[\hat{H}\hat{O}(t) - \hat{O}(t)\hat{H}\right], \end{aligned} \tag{2.12}$$

where $\hat{O}$ is treated as a constant (since up to the last step it has no dependency on time) and in the second line use has been made of the product rule (1.206). By multiplying the above equation by $i\hbar$ and by expressing the last line in terms of the commutator, we finally have

$$i\hbar\frac{d}{dt}\hat{O}(t) = \left[\hat{O}(t), \hat{H}\right], \tag{2.13}$$

where use has been made of the antisymmetric property of the commutator given by Eq. (1.212a). This is known as *Heisenberg equation*. An instance of this equation is represented by Eq. (1.320) when the commutator does not evolve with time. Note that the Heisenberg equation has the same form of the classical Eq. (1.216). These considerations are obviously true for any quantum observable in the Heisenberg picture, provided that the counterpart of the latter in the Schrödinger picture does *not* depend explicitly on time.

But …

Although the two pictures are formally equivalent, Schrödinger's and Heisenberg's standpoints remain quite different and even opposite when more concrete physical considerations are taken into account. In fact, to deal with observables is very different than dealing with the evolution of the state of a system. In particular, as we shall see, they are operationally different, as far as to prepare a state and to select an observable are two different experimental procedures. To a certain extent, this duality is somehow connected with the complementarity principle (Sect. 1.3.3). Moreover, Heisenberg's formulation presents important analogies with Hamiltonian CM (see Eqs. (2.13), (1.215a), and (1.216)), as pointed out by Dirac (Sect. 1.3.2), while the Schrödinger equation is a counterpart of the classical wave equation

$$\frac{\partial^2 u}{\partial t^2} = c^2\nabla^2 u, \tag{2.14}$$

where the u is a scalar function of the position expressing the mechanical displacement of the wave, and this makes a difference at least from the point of view of heuristics.[11] This means that the equivalence of the two pictures does not efface some fundamental differences between these physical concepts. Thus, a formal equivalence does not solve the interpretational problem here at stake.

von Neumann Equation

It is worth mentioning that we can write also an unitary time transformation for density matrices (Sect. 1.4). In fact, I have said that density matrices are kinds of observables, although of particular kind. In the Schrödinger picture, the state vectors $\left|\psi_j(t)\right\rangle$ carry time dependency, and so does the density matrix. By explicitly writing out the time dependence, we write

$$\hat{\rho}(t) = \sum_j \wp_j \left|\psi_j(t)\right\rangle\left\langle\psi_j(t)\right|, \tag{2.15}$$

which is a pure state when we have a single probability $\wp_j = 1$. So, here we make no assumption on whether this density matrix represents a pure state or a mixture. It is noted that the density matrix can be considered as a peculiar observable precisely because usual observables in the Schrödinger picture do not carry time dependence. We will discuss about this subtle point below. Using the Schrödinger equation (1.15), we find, in analogy with Eq. (2.12), that the time derivative of the density matrix is given by

$$\begin{aligned}
\frac{d}{dt}\hat{\rho}(t) &= \sum_j \wp_j \frac{d}{dt}\left(\left|\psi_j(t)\right\rangle\left\langle\psi_j(t)\right|\right) \\
&= \sum_j \wp_j \frac{d}{dt}\left[\hat{U}_t(t)\left|\psi_j\right\rangle\left\langle\psi_j\right|\hat{U}_t^\dagger(t)\right] \\
&= \sum_j \wp_j \left[\frac{d}{dt}\hat{U}_t(t)\left|\psi_j\right\rangle\left\langle\psi_j\right|\hat{U}_t^\dagger(t) + \hat{U}_t(t)\left|\psi_j\right\rangle\left\langle\psi_j\right|\frac{d}{dt}\hat{U}_t^\dagger(t)\right] \\
&= -\frac{i}{\hbar}\hat{H}\sum_j \wp_j \left|\psi_j(t)\right\rangle\left\langle\psi_j(t)\right| + \frac{i}{\hbar}\sum_j \wp_j \left|\psi_j(t)\right\rangle\left\langle\psi_j(t)\right|\hat{H} \\
&= \frac{i}{\hbar}\left[\hat{\rho}(t)\hat{H} - \hat{H}\hat{\rho}(t)\right],
\end{aligned} \tag{2.16}$$

[11] On this see Bokulich (2008, p. 55).

where I have made use again of Eq. (1.304). This result can be rewritten as

$$i\hbar \frac{d}{dt}\hat{\rho}(t) = \left[\hat{H}, \hat{\rho}(t)\right], \tag{2.17}$$

where I have used again the antisymmetric property (1.212a) of the commutator. Note that we have obtained an analogous of the Heisenberg equation (2.13) but by exchanging the positions of the Hamiltonian and the evolving observable, in accordance with the noted peculiarity of the density matrix considered as observable. This equation is known as the *von Neumann equation*, after the name of the Hungarian mathematician, physicist and computer scientist John von Neumann (1903–1957). Just as the Schrödinger equation describes how a state vector evolves in time, the von Neumann equation describes how the corresponding density matrix evolves in time. For the case of a time-independent Hamiltonian that we have implicitly assumed here, the solution to von Neumann equation (2.17) is given by

$$\hat{\rho}(t) = \hat{U}_t(t)\hat{\rho}(0)\hat{U}_t^{\dagger}(t), \tag{2.18}$$

where $\hat{\rho}(0)$ is the initial density matrix of the system at time $t_0 = 0$. In the case of pure states, the above result could also be obtained substituting Eq. (1.15) into Eq. (2.15).

Liouville Equation

Note that the von Neumann equation is a kind of quantum counterpart of the classical *Liouville equation* (or continuity equation), after the name of the French mathematician Joseph Liouville (1809–1882), which rules the dynamics of a statistical ensemble of classical systems.[12] Let us denote with $\rho(q, p; t)$ the density of representative points that at time t are contained in the infinitesimal volume element $d^n p\, d^n q$ in the classical phase space Γ around q and p, where I recall that with *phase space* we denote the Cartesian representation whose coordinates are represented by position and momentum (Fig. 1.25, Sect. 1.2.4). Then, it is possible to show that we have

$$\frac{d\rho}{dt} = \{\rho, H\} + \frac{\partial \rho}{\partial t} = 0 \tag{2.19}$$

or, by applying Eq. (1.215a) we get an instance of Eq. (1.216), i.e.

$$\frac{\partial \rho}{\partial t} = \{H, \rho\}\ . \tag{2.20}$$

[12] Landau and Lifshitz (1976, Sect. 46).

Then, the Liouville theorem states that the density of representative points in the phase space Γ is constant.

Dirac Picture

I mention that a third picture is the so-called *Dirac picture*, after the name of Paul Dirac, and treats the energy observable (the Hamiltonian) of the system as composed of two terms, a free term (describing the unperturbed evolution of the system) and an interaction term (describing the effects of some perturbation). It is therefore especially helpful when dealing with small perturbations, a problem that goes beyond the scope of this work. However, in many cases, it is also convenient to use this picture for splitting some Hamiltonian of a compound system in a part that pertains to each subsystem and an interaction part. For instance, in the case of two arbitrary subsystems 1 and 2, we could write

$$\hat{H}_{1+2} = \hat{H}_1 + \hat{H}_2 + \hat{H}_{12}, \tag{2.21}$$

where $\hat{H}_{1+2}$ is the total Hamiltonian, $\hat{H}_j$, $j = 1, 2$ are the subsystems 1's and 2's Hamiltonians and $\hat{H}_{12}$ is the interaction Hamiltonian.

2.1.2 Probability and Interference

Evolution of Probabilities

If we ask, on the basis of such examination, what quantum dynamical evolution (for both states and observables) means, there is a clear answer: it is an evolution of probability amplitudes, from which probabilities can be computed, allowing us to extract (partial) information from the system. This is certainly true and represents one of the major differences with CM. In fact, due to the basic postulates of CM, we can always describe the evolution of any state with all its single properties considered as real objects possessing independent reality.

The probability amplitude is the analogous of a wave amplitude (Fig. 1.1, Sect. 1.1.1): we have indeed seen that quantum-mechanical systems can be understood in a first (and very rough) approximation as kinds of waves (see also Sect. 1.3.3). However, there is an important difference too: any classical wave (apart from electromagnetic waves, which are in fact quantum-mechanical systems) can be considered as a characteristic (elastic) deformation or distortion of a physical medium (like water or air), but there is no medium in which quantum "waves" move or of whose physical stuff they are made (this was discovered for light already between the end of the nineteenth and the beginning of the twentieth centuries when the theory of aether was confuted). So, they are just evolving probability amplitudes.

Classical Probability Theory

However, there is also a connection. It is well known that classical waves can make interference (Fig. 1.3, Sect. 1.1.3). The same is true for the quantum probability amplitudes, as recalled in Sect. 1.4. This has a momentous consequence, since quantum probability does no longer satisfy some of the basic characters of classical probability.

First, let us summarise (classical) probability theory. The set of all possible outcomes of an experiment is called the *sample space*, usually denoted by Ω. A subset of the sample space to which a probability can be assigned is called an event. In other words, an event is a set of outcomes of the experiment (note that different outcomes may be grouped in a single event). In particular, the sample space Ω itself is an event; by definition it always occurs (because it comprehends all possible outcomes). At the other extreme, the empty set $\emptyset$ is also an event; by definition it never occurs.

Let A and B be two events, if A is a proper subset of B (denoted in the set theory by $A \subset B$), when A occurs, B must also occur. The union of A and B (denoted by $A \cup B$) is the event obtained by combining (summing) the elements of A and B. The intersection of A and B (denoted by $A \cap B$) is the event whose elements are common to both A and B. If the intersection of events A and B is empty, (denoted by $A \cap B = \emptyset$), then A and B are said to be mutually exclusive (or disjoint) events.

It is an experimental fact that if an ordinary coin is flipped N times, coming up heads N_h times, the ratio of N_h to N is nearly $\frac{1}{2}$, and the large we make N the closer the ratio approaches $\frac{1}{2}$. We express the results by the following statement: the probability of a coin coming up heads is $\frac{1}{2}$. This illustrates the intuitive frequency concept of probability. Let us generalise this as follows. If a random experiment has N possible outcomes, all mutually exclusive and equally likely, and the number N_A of these outcomes lead to the event A, then the probability of A is defined by

$$\wp(A) = \frac{N_A}{N}, \tag{2.22}$$

which is the classical definition of probability. Classical probability satisfies the following basic rules for alternative events A, B (the first three are called Kolmogorov's axioms, after the name of the great Russian mathematician Andrey N. Kolmogorov (1903–1987))[13]:

- $0 \leq \wp(A) \leq 1$ for all $A \subset \Omega$, (2.23a)
- $\wp(\Omega) = 1$, (2.23b)
- $\wp(A \cup B) = \wp(A) + \wp(B) - \wp(A, B)$, (2.23c)

where $\wp(A, B)$ denotes the *joint probability* of events A and B, i.e. $A \cap B$, and

[13] Kolmogorov (1933), Gnedenko (1969, p. 48). The latter is a classical textbook on probability theory. See also Landsman (2017, Sect. 1.1).

$$\bullet \ \wp(A \cup B) = \wp(A) + \wp(B) \text{ if } A \cap B = \emptyset. \tag{2.23d}$$

It may be noted that Eqs. (2.23c) and (2.23d) are referred to as the addition rule for probability and that the latter is a special case of the former. Equation (2.23d) can be generalised as follows:

$$\wp\left(\bigcup_j X_j\right) = \sum_j \wp(X_j) \tag{2.24}$$

for n disjoint events X_j. An event always defines a complementary event, namely, the complementary set (the event not occurring), and together these define a so-called *Bernoulli trial*, after the name of the Swiss mathematician Jacob Bernoulli (1654–1705): did the event occur or not?

A Classical Example

To understand the difference between classical and quantum probability, note, in first instance, that, classically, a probability only expresses the subjective ignorance or incertitude about certain events. For instance, before throwing a dice we do not know which face we shall get. However, we consider that this is possible in principle. Moreover, the probability to obtain, say, a six remains well distinct and relatively independent from the probability to obtain a two. The only connection between these two probabilities is represented by the way in which the event space is sampled. In the case in which we throw one dice, we have a probability of 1/6 to get any of the six numbers. When throwing two dice, the problem is different, since we need to consider the possible combinations between the faces of the two dice. In this case, it is easy to understand that we can get a two only if the dice show both a one. However, we can obtain a seven in each of these cases: (i) 1 and 6 or 6 and 1, (ii) 2 and 5 or 5 and 2, (iii) 3 and 4 or 4 and 3. Now, it is clear that, with two dice, to get a seven is six times more probable than to get a two.

A Quantum-Mechanical Example

Suppose now a slightly different arrangement. Suppose that we ask about the probability that somebody (or a group of persons) takes one of two ways out of a certain closed space like a room. The situation is classically essentially the same as that of the dice throwing. Quantum-mechanically, however, due to superposition or entanglement, the probability to take one path is not independent from the probability to take the other one, as it is evident by considering again Fig. 1.3, Sect. 1.1.3. This kind of interdependence has nothing to do with the way in which we sample the space of the events. It is rather like an invisible and non-local thread that connects the two possible events. This is an aspect of the quantum-mechanical formalism that is puzzling and of which we should ask if it has any ontological substrate at all: if, in a given experimental context, events have not yet happened (and in fact,

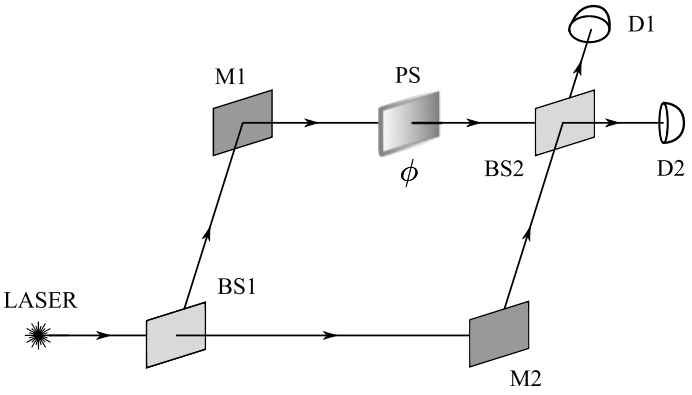


Fig. 2.2 Scheme of the Mach–Zehnder interferometer. The apparatus is shown in three dimensions with a top-lateral view. See Auletta et al. (2009a, Sect. 1.2) for further details

also quantum-mechanically, only one of the two will subsequently happen in our universe), and thus we cannot have localisation of the system, how can we speak of an interdependency at all? Between what?

To elucidate the problem, let us make use of an interferometer, i.e. a device for producing interferences (here, among photons), as shown in Fig. 2.2. Note that here and in the following essentially Mach–Zehnder interferometry is used (after the names of the Czech inventor Ludwig Mach (1868–1951) and the Swiss physicist Ludwig L. A. Zehnder (1854–1949)). Suppose a laser that produces a wave train that is short compared with the total distance covered. The source works at such a low intensity that, at most, only a quantum of energy (photon) is carried in each pulsation. The wave train is split into two beams by means of the BS1 device (beam splitter), and recombined using the BS2 device (that acts both as a beam merger and as a beam splitter as a consequence of the interference between the two paths); moreover, the two split beams are reflected fully by the two mirrors, M1 and M2 and the phase of the upper component is shifted by the phase shifter PS, a device producing phase differences between the two components. Beam splitters are half-silvered mirrors that allow the light to be partly reflected and partly transmitted.[14] The proportion of these two components can be adjusted but, for the sake of simplicity, here and in the following if nothing is added I always assume that the reflection and transmission coefficients are equal to $1/\sqrt{2}$ (BS1 and BS2 send the two components along the two paths with equal probability 1/2). Alongside these instruments, two photon detectors (D1 and D2) are also inserted at the final end. Note that the state of the photon, initially in $|d\rangle$, just after the first beam splitter can be written as the following superposition:

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|d\rangle + |u\rangle\right), \tag{2.25}$$

[14] A beam merger is like a beam splitter but works the other way around.

where $|d\rangle$ denotes the component that has taken the lower path ('down' in short) and $|u\rangle$ denotes the component that has taken the upper path ('up' in short). If we represent the basis that we are using as usually, that is, we define

$$|d\rangle := \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad |u\rangle := \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{2.26}$$

we can express the action of any of the two beam splitters by means of the unitary operator (also called the Hadamard operator, after the name of the French mathematician Jacques Hadamard (1865–1963)) of the form

$$\hat{U}_{\text{BS}} := \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}. \tag{2.27}$$

Two subsequent applications of the same unitary operator brings the system back to the initial state, so that the Hadamard transformation is the inverse of itself ($\hat{U}_{\text{BS}} = \hat{U}_{\text{BS}}^{\dagger}$), but, being a unitary transformation, this does not imply that it is an observable. This also clarifies the notion of beam merger.

When the upper beam goes through the phase shifter, we have

$$\frac{1}{\sqrt{2}} \left(|d\rangle + |u\rangle \right) \overset{\text{PS}}{\longmapsto} \frac{1}{\sqrt{2}} \left(|d\rangle + e^{i\phi} |u\rangle \right), \tag{2.28}$$

where $\phi \in [0, 2\pi]$ in radians (or $0^\circ \leq \phi \leq 360^\circ$). Of course, the phase difference can be experimentally arranged. The above equation expresses in terms of map or transformation what in traditional equation-like terms would be

$$\frac{1}{\sqrt{2}} \left(|d\rangle + e^{i\phi} |u\rangle \right) = \hat{U}_{\phi} \left[\frac{1}{\sqrt{2}} \left(|d\rangle + |u\rangle \right) \right], \tag{2.29}$$

where $\hat{U}_{\phi}$ is a unitary phase-shift operator whose form it is not important now. It is interesting to see what happens when the two beams or components pass the second beam splitter. In this case, we have the transformations

$$|d\rangle \overset{\text{BS2}}{\longmapsto} \frac{1}{\sqrt{2}} \left(|1\rangle + |2\rangle \right), \tag{2.30a}$$

where with $|1\rangle$ and $|2\rangle$ I indicate here the two paths leading to the detectors D1 and D2, respectively. The transformation on $|u\rangle$ has, instead, a different form thanks to the Hadamard operator:

$$|u\rangle \overset{\text{BS2}}{\longmapsto} \frac{1}{\sqrt{2}} \left(|1\rangle - |2\rangle \right). \tag{2.30b}$$

Note that all transformations performed by the BSs, BMs, PS and Ms are unitary. So, up to detection, every operation is reversible. Inserting the latter two transformations

into Eq. (2.28), we obtain the final state $|f\rangle$ of the photon *after* it leaves the beam splitter BS2, but *before* it is set to be detected at D1 or D2

$$|f\rangle = \frac{1}{2}\left[\left(1+e^{i\phi}\right)|1\rangle + \left(1-e^{i\phi}\right)|2\rangle\right]. \tag{2.31}$$

It is important to note that the state $|f\rangle$ is precisely of the same form as that in, e.g. Eq. (1.53), with the coefficients of components $|1\rangle$ and $|2\rangle$ given by

$$c_1 = \frac{1}{2}\left(1+e^{i\phi}\right) \quad \text{and} \quad c_2 = \frac{1}{2}\left(1-e^{i\phi}\right), \tag{2.32}$$

respectively. The final probabilities to detect the photon are given by

$$\wp_1 = |c_1|^2 = \left|\frac{1}{2}\left(1+e^{i\phi}\right)\right|^2 = \frac{1}{4}\left(1+e^{i\phi}\right)\left(1+e^{-i\phi}\right) = \frac{1}{2}\left(1+\cos\phi\right) \tag{2.33a}$$

and

$$\wp_2 = |c_2|^2 = \left|\frac{1}{2}\left(1-e^{i\phi}\right)\right|^2 = \frac{1}{2}\left(1-\cos\phi\right). \tag{2.33b}$$

For deriving the above probabilities I have taken Euler formulas (1.181) into account, from which, by summing and subtracting, we get

$$\cos\phi = \frac{1}{2}\left(e^{i\phi}+e^{-i\phi}\right), \quad \sin\phi = \frac{1}{2i}\left(e^{i\phi}-e^{-i\phi}\right). \tag{2.34}$$

Normalisation of probability is satisfied as it is evident. I also recall that trigonometric functions satisfy the Pythagorean trigonometric identity

$$\cos^2\phi + \sin^2\phi = 1. \tag{2.35}$$

Interferometry Interaction-Free Experiment

Let us now elucidate the issue at stake by means of the already mentioned interaction-free experiment (Sect. 1.2.4).[15] For the intelligence of the experiment note that interference is realised by letting the two main paths cross at the second beam splitter (BS2). Therefore, if we block one of the possible paths that the system can take, there is no longer interference. Now, interference determines specific effects on the detection results (collected at detectors D1 and D2). Thus, blocking the interference will alter the probabilities of these outputs, what allows us to infer the presence of an obstacle on one of the two paths. In order to investigate the quantum-mechanical

[15]The first idea was proposed in Renninger (1960). For some additional material see Auletta (2000, pp. 353–358), Auletta et al. (2009a, Sect. 9.6).

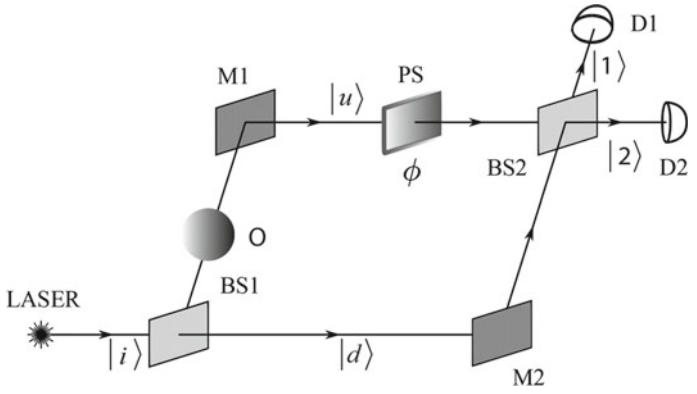


Fig. 2.3 Scheme of the interaction-free measurement with a Mach–Zehnder interferometer

roots of this result, let the evolution of the system *in absence of the object* (and of the phase shifter) be described as follows:

$$\begin{aligned} |i\rangle &\overset{\text{BS1}}{\longmapsto} \frac{1}{\sqrt{2}}\left(|d\rangle + |u\rangle\right) \overset{\text{M1, M2}}{\longmapsto} \frac{1}{\sqrt{2}}\left(|d\rangle + |u\rangle\right) \\ &\overset{\text{BS2}}{\longmapsto} \frac{1}{2}\left(|1\rangle + |2\rangle\right) + \frac{1}{2}\left(|1\rangle - |2\rangle\right) = |1\rangle \ , \end{aligned} \tag{2.36}$$

where $|\,i\rangle$ is the initial state (an incoming photon entering the interferometer from the left) and BS2 makes the transformations (2.30a) and (2.30b). Note that, after BS2, if the phase difference ϕ is zero (or the PS is inactive), detector D1 detects all photons (due to constructive interference) while at D2 we have the dark-output (no photon, due to destructive interference). Then, the photon leaves the interferometer moving up, and it is detected by D1.

Suppose now that *the object O is present* in the upper path, as shown in Fig. 2.3.[16] Then, the evolution of the system is described by

$$\begin{aligned} |i\rangle &\overset{\text{BS1}}{\longmapsto} \frac{1}{\sqrt{2}}\left(|d\rangle + |u\rangle\right) \\ &\overset{\text{O}}{\longmapsto} \frac{1}{\sqrt{2}}\left(|d\rangle + |a\rangle\right) \overset{\text{M1, M2}}{\longmapsto} \frac{1}{\sqrt{2}}\left(|d\rangle + |a\rangle\right) \\ &\overset{\text{BS2}}{\longmapsto} \frac{1}{2}\left(|1\rangle + |2\rangle\right) + \frac{1}{\sqrt{2}}|a\rangle \ , \end{aligned} \tag{2.37}$$

where $|a\rangle$ represents the state of the photon when it is absorbed by the object. From the last line of Eq. (2.37), we have three possible outcomes:

[16] The MZ interferometer version of the interaction-free measurement was proposed for the fist time in Elitzur and Vaidman (1993) .

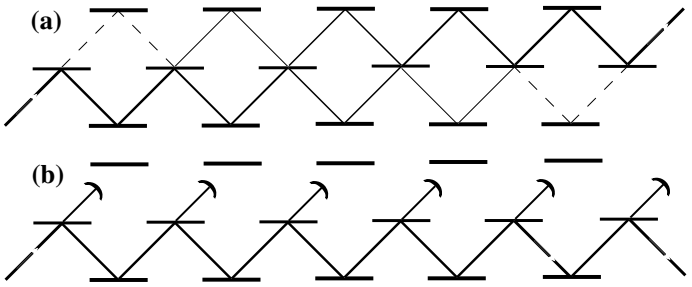


Fig. 2.4 **a** The principle of coherently repeated interrogation. The probability that the photon is reflected at each of the N beam splitters is $\cos^2(\pi/2N)$ (while the relative phase is 0). A single photon incident from the lower left gradually transfers to the upper right half of the system. After N stages, where N depends on the BS reflectivities, the photon will certainly exit via the upper port of the last BS. **b** Introduction of detectors prevents the interference. At each stage the state is projected back into the bottom half of the system if the respective detector does not fire. After all stages there is a good chance that the photon now exits via the lower port of the last BS, indicating the presence of the detectors. Adapted from Auletta et al. (2009a, p. 317)

- Detection by D1,
- Detection by D2,
- Absorption by the object,

with probabilities of 1/4, 1/4 and 1/2, respectively. Thus, with probability 1/4, we can detect the presence of an object without interacting with it, precisely in those cases in which the detector D2 does click. In the absence of the object, in fact, it could not click. In other words, we have altered the detection probabilities by the *sole presence* of an object, and not through some kind of action on the photon. Note that the probability to detect the obstacle without interaction can be considerably increased if we arrange a sequence of beam splitting. In ordinary conditions, after all N stages, the photon will certainly leave by the upper exit, but, if we insert a series of detectors (which together represent here the 'obstacle') in the upper half of the apparatus which prevent the interference (Fig. 2.4b), then there is a small chance that the photon takes the upper path and triggers a detector, and a large probability $\cos^2(\pi/2N) \simeq 1 - \pi^2/4N^2$ that it continues to travel on the lower path. The non-firing of each detector projects the state onto the lower half.

Significance of the Experiment

Coming back to the previous experiment, what is amazing here is that the photons that arrive at the detector D2 are precisely those that *have not met* the obstacle (otherwise they would have been absorbed by the latter). In other words, the state (2.25) of the photon, due to the presence of the obstacle, is projected into one of the two components, either $|d\rangle$ or $|u\rangle$, precisely where the obstacle is not. This means that we are able to detect the presence of the obstacle although the physical

systems that are measured did not interact with it (from here the name *interaction-free* measurement). This is interesting, but so far we have not explained why we can detect the presence of an obstacle on one path of an interferometer without interacting with it. This can now be clarified as follows: it is precisely this non-local thread that, making the two paths interdependent, makes the 'wave' component that has not met the obstacle 'sense' that there is such a one on the other path. It is like when the fisher knows that he/she has taken a fish by feeling that something happened at the other end of the fishing net. In fact, since, without obstacle, we have interference with a single photon and this cannot be split, we cannot say that the interference is produced through the interference of two parts of the photon following each one a classical trajectory of which we could know at each time the 'location' through which it passes. Then, the only conclusion is that the photon is in a *global* state such that there is no localisation at all but a connection between probability amplitudes or components of the global state.

Now, the crucial point is the following: if the presence of this thread allows us to make predictions about the presence of an obstacle that otherwise could not be formulated, then it should be something real. However, we immediately run in a serious difficulty here. The word *thread* seems to have misled us so that we likely forgot that we are dealing here with probability amplitudes and not with something real in any ordinary sense, and the so-called "thread" is finally only a (metaphoric) term for expressing some kind of relation among these probability amplitudes, as, e.g. displayed in Eq. (1.359): the last two terms of the last row are the interference terms, and they mathematically express such a quantum relation among probability amplitudes. Is this only a pure formal aspect? However, if the presence or absence of the latter makes an *observable* difference, we cannot avoid the conclusion that it should be somehow real.

Difference Between Classical and Quantum Probabilities

The issue at the stake here concerns not only such possible interconnections but also the character of probabilities as such: if probabilities in QM can be altered in the way described so far, then it is likely that in many situations classical laws of probability are no longer satisfied, as I have anticipated. In fact, if the presence of an obstacle in one of the two paths of an interferometer eliminates the interference, it is clear that in some cases we increase the probability of detection at one of the two detectors, precisely in those cases in which the phase difference between the two paths is such that there is a destructive interference at that detector. Then, the presence of an obstacle increases in those cases the probability to obtain a result (see also Fig. 1.17, Sect. 1.2.4, for further evidence), what is certainly impossible from a classical point of view. In fact, classically speaking, the presence of an additional obstacle cannot increase the probability of detection in any case, because a blocked stream contributes to nothing.

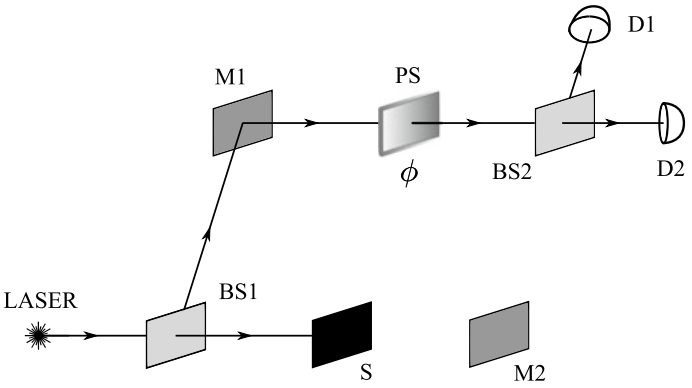


Fig. 2.5 Mach–Zehnder interferometer with a movable screen S. Here, it is inserted in the lower path. Adapted from Auletta and Wang (2014, p. 89)

To display these differences in the most effective way, let us go back to the Mach–Zehnder experiment and propose a variant of the interaction-free measurement.[17] Let us rewrite the probabilities (2.33) for a single photon to be detected at detectors D1 and D2 as

$$\wp_{u+d}(\mathrm{D1}) = \frac{(1+\cos\phi)}{2} \text{ and } \wp_{u+d}(\mathrm{D2}) = \frac{(1-\cos\phi)}{2}\,, \tag{2.38}$$

respectively. As we have seen, this result is due to interference between two "alternatives": photon taking the upper path and photon taking the lower path. This justifies the notation above where the subscript $u+d$ means that the photon can take both paths. If we block the lower path by inserting a movable screen S between BS1 and M2 (Fig. 2.5), then it is clear that the photon will be absorbed by the screen with probability $1/2$: in other words, we have an obstacle on one of the paths as in the interaction-free measurement, although we are not interested now in the changes of the final state as such but in the change induced on the *probabilities* of certain happenings. With probability 1/2, the photon will take the upper path and will be reflected by M1. At the symmetric beam splitter BS2 it will have 50% chance of going to detector D1 and 50% chance of going to detector D2. The overall probabilities for a photon impinging on BS1 to be detected at D1 or D2 will be therefore $\wp_u(\mathrm{D1}) = 1/4$ and $\wp_u(\mathrm{D2}) = 1/4$, respectively. A similar analysis can be performed in the case we insert the screen S in the upper path between BS1 and M1: again the probabilities for detection at D1 or D2 will be $\wp_d(\mathrm{D1}) = \wp_d(\mathrm{D2}) = 1/4$. We immediately see that the probability for detection at, say, detector D1 when both paths are open is in general *not* equal to the sum of the probabilities of being detected at D1 after taking the two paths separately, i.e.

[17] I have followed here Auletta et al. (2009a, Sect. 1.4).

$$\wp_{u+d}(\mathrm{D1}) \neq \wp_u(\mathrm{D1}) + \wp_d(\mathrm{D1}) \,, \tag{2.39}$$

except in the cases $\phi = \pi/2$ and $\phi = 3\pi/2$. This is due to the fact that, when both paths are open, the state of the photon after BS1 is not a mere addition of the two alternatives but rather a quantum superposition of them. This contradicts the basic structure of classical probability. In classical probability theory, given two events A and B, as an immediate consequence of the classical law (2.23c), we have that

$$\wp(A \cup B) \leq \wp(A) + \wp(B) \,. \tag{2.40}$$

In the example of Fig. 2.5 this inequality is violated for all values of $\phi \neq \{\pi/2, 3\pi/2\}$ either for D1 or for D2.

Let us compare this result with that obtained in the case the experimental setup shown in Fig. 2.5 is replaced by its classical analogue. In this classical device, photons are replaced by bullets and the beam splitters by random mechanisms that send each bullet in one of the two paths, with equal probability over many runs. Then, if both paths are open, the probability of detection at both D1 and D2 is equal to 0.5, i.e. $\wp_{u+d}(D1) = \wp_{u+d}(D2) = 1/2$. On the other hand, if one of the two paths is blocked, we have $\wp_u(D1) = \wp_u(D2) = \wp_d(D1) = \wp_d(D2) = 1/4$. It clearly results that, in this classical example, Eq. (2.39) becomes an equality and therefore the requirement (2.40) is obviously satisfied.

The result (2.39) is strictly related to the fact that, due to the superposition principle, quantum probabilities are calculated as square moduli of the corresponding amplitudes and that, therefore, in QM *amplitudes* and not the corresponding probabilities sum linearly. In particular, when more than one 'alternative' (or 'path') lead to the same measurement outcome, one has first to sum the amplitudes corresponding to the different 'alternatives' (as in Eq. (2.31)) and then to calculate the square modulus in order to obtain the probability of that measurement outcome.

Then, the inference that we can make from these ideal experiments is that probability amplitudes and therefore probabilities express something real that defy our (classical) expectations. In fact, we have agreed that we need to consider objective what is able to induce a change of our hypotheses, and, in a physical context, this should be some kind of physical reality. This makes a huge difference relative to CM, where, as noticed, probabilities only express our own incertitude about events that are perfectly well determined in themselves. Since QM rejects the principle of the *omnimoda determinatio* (i.e. the assumption that there is a reality perfectly determined in itself), as recalled in Sects. 1.2.1 and 1.2.4, there must be here a connection between the indetermination of the quantum systems and this objectivity of probabilities, a connection that we need still to explore.

2.1.3 What Does the Quantum-Mechanical Formalism Tell Us?

I have explained (in Sect. 2.1.1) that there are two main kinds of time evolution in QM: the Schrödinger picture dealing with the evolution of states and the Heisenberg picture dealing with the evolution of observables. There is also a third category that was often introduced: that of property, where with property I recall that, in such a context, it is meant the value that we attribute to an observable in some state (one of its eigenstate). Many prefer to speak of *effect* instead of property. This terminology is fully appropriate. Nevertheless, as we shall see, there are problems with the way in which both notions are mostly used. Thus, I prefer to maintain for the time being the notion of property and to deal with the analysis of these notions later on.

So far, the notions of state, observable and property are actually the only three elements of this examination that not only do not have a pure formal (mathematical) character but seem to be the referent of that formalism. In fact, they have been involved in the whole examination of the present section. Therefore, we could assume that this triplet of categories represents somehow the ontological substrate of the theory by determining what a quantum system is.[18] However, we have also seen how elusive is the reality described by QM. This demands care. Then, we need certain further refinements of our notions in order to understand in which sense this triplet can be invested of ontological value if any. In conclusion, the relation between formalism and ontology is problematic and needs to be still clarified.

2.2 The Measurement Problem

2.2.1 It is Not a Problem for Classical Physics

In the previous section, we have seen that CM treats states and observables (or variables) on the same foot: since it is assumed that we can extract the whole information contained in a state and since a state is fully defined through its pertinent variables, it is the same operation (or at least a combination of operations) that answers both questions:

(i) What is the value of an observable?
(ii) Which is the initial state of the system?

In fact, since all classical observables commute, they are also jointly measurable (Sect. 1.2.4), so that we can always think of a composite measurement that will provide us a complete information of the state of a classical system.

[18] What is to a certain extent also supported in Margenau (1950, Sects. 8.5 and 15.4).

It is clear, also from a pure classical point of view, that some experimental setups are better suited for measuring certain observables and not others. For instance, for measuring the position of a system we need a rigid grid (that allows us to precisely localise the system), but for measuring the momentum we need rather a spring or some other elastic device (allowing to absorb the impact of an object moving with a certain speed and having a certain mass).[19] Nevertheless, the principle of continuity acknowledged by CM, as recalled in Sect. 1.1.2, ensures us that measurements of position and momentum performed in subsequent times (with whatever order, due to the commutativity of classical observables) allow us to reconstruct the whole state of the system. Therefore, classically, state and observable conflate into a single physical entity (what justifies why two different pictures were not traditionally considered in CM: Sect. 2.1.1). Moreover, since the values that a certain variable possesses are assumed to be properties of the system being in that state and therefore contribute to the very definition of a state, also properties conflate in the same entity together with states and observables. Ultimately, there is only the state characterised by its parameters or degrees of freedom and relative properties. Due to both the assumed completeness and the assumed correctness of the theory (a point to be dealt with in the next section), we have here a full isomorphism between theory and reality, and so the categories state, variable (observable) and property are assumed to resume the reality of a physical system at a certain time.

Thus, classically, no measurement problem can arise, as far as to measure only means to become aware of (or verify) the state with its properties that the system already possesses in reality before our taking notice of it.[20] Of the latter statement is possible to deny one of its two parts: either the irrelevancy of being aware or the fact that properties describe immediately ontological facts. As I shall argue, I am inclined to reject the second part and to reconsider in critical terms the first part. I recall that one of the physicists who more contributed to our current understanding of measurement procedures was Günther Ludwig. He strongly stressed that the triviality of the measurement problem in CM is in fact a mistake that hinders a clear understanding of the process.[21]

Now, the question raised in Sect. 2.1.3 becomes more interesting: since states (as well as observables and properties), in a classical framework, are considered to be real and as the ontological substrate to which we refer when describing the system (at least at a certain moment), it becomes mandatory to ask how QM considers these notions.

[19] For this argument see Bohr (1949a, pp. 219–220).

[20] In D'ariano et al. (2017, Sect. 5.1) this is expressed by saying that measurement outcomes become irrelevant in a classical framework.

[21] Ludwig (1983, I, pp. 55–56).

2.2.2 *It is a Double Problem for Quantum Mechanics*

Measurement and Detection

Like for CM, all time evolutions in QM (the time evolution of both states and observables) are reversible. We express this by means of a unitary operator. As mentioned, most of the transformations in QM are unitary. However, there is a particular kind of operation that is not unitary: *detection*. Traditionally, many authors speak of "measurement" in the same sense in which I am using now the term "detection". However, there is an ambiguity here: measurement can be understood both as the final act of detection or as the whole process that comprehends in itself also stages that precedes detection. The interaction-free measurement (Sect. 2.1.2) allows us to understand this: indeed, it is called interaction-free *measurement* and not interaction-free *detection*. As a matter of fact, also in such a case there is detection (without a detection event nothing would have happened or could be inferred). However, what we have discovered is that the detected photons are those that have not interacted with the obstacle, and this is precisely what allows to speak of measurement (but not of a detection) without interaction. Thus, to avoid any ambiguity, in the following I prefer to use the word *detection* to denote the final step of the process through which information about the object system is acquired and employ the term *measurement* for describing the whole process and sometimes the final step only when no possible misunderstanding can arise. Following this convention, I shall call *apparatus* the coupling device representing the experimental setup (or physical context) and *detector* the detection device. Sometimes, when speaking of the specific part of the apparatus that shows the graduate scale with the possible values of the measured observable, I shall use the term *pointer* or also *metre*.[22]

A Non-unitary Dynamics

That detection is not unitary (reversible) is quite clear by considering that, when we measure a system, in many situations its initial state is a superposition (Sect. 2.1.1). Let us recall that all possible eigenvalues of an observables (which allow us to ascribe properties to the system under consideration) are associated with one of its eigenvectors, that is, with one of the vectors in the Hilbert space that are elements of an orthonormal basis (its eigenbasis) that we use for expanding the state of the system that we are measuring (Eq. (1.113)). Both in the classical case and in most controlled situations for the quantum case, this state is known in advance: in QM through a procedure called preparation, which represents the first stage of the measurement process. Suppose, for the sake of the argument that we knew such a state. Since the state vector of the system represents by assumption a known superposition of those eigenvectors, it is clear that we could bring this state vector to coincide with a particular one among these eigenvectors by simply rotating it and letting coincide

[22] Ludwig used the term registration apparatus for denoting the detector (Ludwig 1983, I, p. 8).

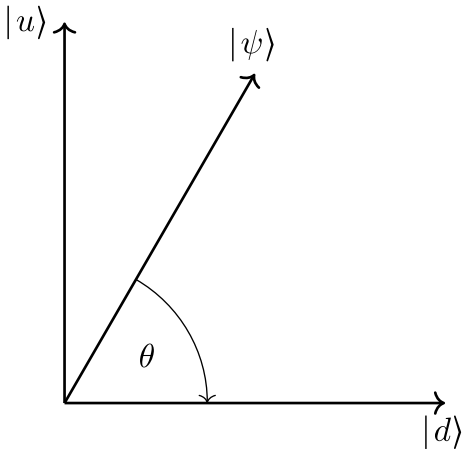


Fig. 2.6 Unitary transformation of a state vector $|\psi\rangle$ making and angle θ with the state $|d\rangle$. By performing a CW rotation of this angle on the plane it will coincide with $|d\rangle$

with the latter one (Fig. 2.6). This operation has nothing mysterious and is perfectly allowed by QM (it is unitary, indeed). Here, the situation appears not so different from the classical counterpart.

However, as mentioned, the point is that often we do not know the input state of the system in QM.[23] This is quite common for spontaneous processes of photon absorption or emission. But also in experimental context such a problem may arise. For instance, coming back to the interaction-free measurement, suppose that we do not know where is the obstacle. If it is in the upper path, before BS2 the state is represented by the second row of Eq. (2.37). However, if it is in the lower path, the system, before BS2, is rather in the state

$$\frac{1}{\sqrt{2}}\left(|u\rangle + |a\rangle\right). \tag{2.41}$$

Clearly, the two vectorial representations are not identical, and it cannot be the same transformation to bring the system, say, to state $|2\rangle$. In such cases the situation is deeply different from the previous example: the initial state vector makes now an unknown angle with the eigenvector that is associated with the subsequent result of the detection. Since this angle can be arbitrary (of any magnitude), there does not exist a *single* unitary operation that would allow us to transform any initial superposition into a given possible detection outcome. Strictly speaking, with *detection outcome* we understand one of the eigenvalues of the measured observable. However, such eigenvalues are associated with eigenvectors of the same observable and thus we can broadly speak of an output state. Thus, this transformation cannot be unitary and therefore it must be irreversible (Sect. 1.2.5). This is quite surprising: we have a formalism (the Schrödinger equation or the Heisenberg one) that should describe the quantum dynamical processes, but, when we perform an operation (a detection)

[23] These arguments are also summarised in Auletta and Wang (2014, Sect. 9.1).

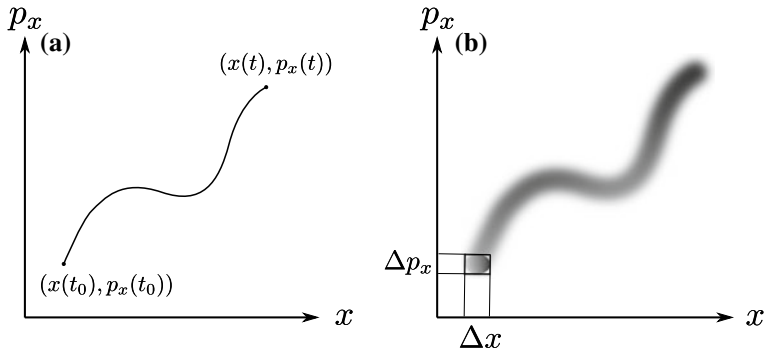


Fig. 2.7 **a** In CM, given a system's initial state $(x(t_0),\ p_x(t_0))$ at time t_0, the evolved state $(x(t),\ p_x(t))$ of the system at later time t is uniquely determined (the same is true also for antecedent time). Thus, the state of a classical system at any given time is represented by a point in the phase space, and its time evolution is represented by a single trajectory (a well-defined curve). As we shall see, this constitutes *classical* determinism. **b** According to the uncertainty relation $\Delta x\ \Delta p_x \geq \hbar/2$, the quantum-mechanical phase space can be thought of as being divided into cells, each having a minimal area $\Delta x\ \Delta p_x \approx \hbar/2$, and the state of a quantum system is roughly represented by an elliptical spot that is bounded by the cell (Fig. 1.25, Sect.1.2.4). As a consequence, it is not possible to ascribe a definite phase space trajectory to a quantum system, but we have rather a multipath dynamics: the most probable states of a quantum system can at best be thought of as a fuzzy contour representing a bundle of trajectories. As we shall see this constitutes *quantum* determinism. Adapted from Auletta and Wang (2014, p. 154)

on a quantum system, then that formalism does no longer describe such a dynamics! This is precisely the opposite of what we expect from a true physical theory.

No Cumulation of Information

There is second problem that is somehow related with the previous one. As I have mentioned, classical systems are assumed to be both perfectly determined and knowable (Sects. 1.1–1.2). This means that a classical system occupies always a determined location in its phase space. And, as a consequence, the state of a classical system is represented by a single point as well as its time trajectory by a single curve (Panel (a), Fig. 2.7). It is further assumed that this curve (all of its points) is (are) perfectly knowable to us. Is this assumption justified? Can we support a full determination of the state together with its properties? In Sect. 1.1.2, I have cast some doubts on this assumption.

In whatever way we consider this assumption, the situation is deeply different for QM.[24] Here, the commutation relations among many observables hinder the joint measurability of any observables (Sect. 1.2.4). Moreover, due to the uncertainty relation the quantum-mechanical phase space can at best be thought of as being divided

[24]The following arguments have been presented in Auletta et al. (2009a, Sect. 2.3) and a little expanded in Auletta and Wang (2014, Sect. 6.9).

into cells, each of an approximate area representing the minimal value allowed by those relations, and the state of a quantum system is roughly represented by an elliptical spot that is bounded by the cell. As a consequence, the time evolution in phase space is not represented by a single curve but as a pipe that represents a bundle of trajectories (Panel (b), Fig. 2.7). Thus, any attempt at reducing the uncertainty of position will immediately induce a corresponding increase of the uncertainty in momentum, and vice versa, always in accordance with the uncertainty relation: if there is no increase in the area, only changes in its shape, becoming more or less elliptical according to the operation at the hand, are allowed (analogously to what happens classically with the incompressibility of the density of representative points in phase space determined by Eq. (2.20)). In other words, although we are free in the choice of the observable to measure (by hypothesis),[25] the consequences of detection are *not controllable* by us. As anticipated in Sect. 2.1.1, if we decide to measure the position observable, we shall obtain a certain kind of information that should be contained in the state of the system but we also alter the state of the system in some way. On the other hand, if we decide to measure the momentum observable, we shall obtain *another* kind of information that should still be contained in the initial state; however, we alter also in this case the state of the system but *in another way*. These two pieces of information represented indeed two different and incompatible properties due to the non-commutativity of these quantum observables and the uncontrolled change of the state. The crucial problem is that we cannot put together these two pieces of information, what shows that we cannot choose a measurement with the purpose to get information about both the position and momentum observables as in the classical case (although, as we shall see, kinds of 'intermediate' situations are still possible). Summarising, we are not completely free even in the *single* measurement act as regarding its consequences, which we cannot control, although we are still free to choose alternative experimental setups at any moment (i.e. to choose to measure different observables). How is it possible?

As mentioned in the previous subsection, both classically and quantum–mechanically, the experimental contexts for measuring position and momentum are not identical. Hence, we cannot measure these two quantities jointly, instead we need a succession of at least two measurements. However, in CM, the already mentioned assumption of continuity together with the commutativity of classical observables, the specific deterministic character of classical laws and the possibility to reduce to zero the measurement error ensure us that we can combine the information obtained through these two consecutive measurements in order to fully reconstruct the state of the system. However, this is precisely what cannot happen in QM, where, due to quantisation and uncertainty relations as well as to a multipath dynamics ruling evolution of probability amplitudes, the order in which we perform two or more measurements (detections) does matter (see e.g. Fig. 1.17, Sect. 1.2.4): each measurement will discontinuously change the state of the system, and, after such a jump, the system (if survived) will follow a dynamics that renders future outcomes not pre-

[25] As pointed out in Conway and Kochen (2006, 2009). See also Bohr (1949a, p. 230), Landsman (2017, Sects. 6.2–6.3).

dictable, making impossible to cumulate the information extracted from subsequent measurements. There is no way to circumvent this problem. Thus, we can understand why dealing with observables raises questions that cannot be easily answered from Schrödinger's point of view (Sect. 2.1.1).

The Observer

This means that the observer of a quantum system cannot be considered as a detached spectator who can look at physical systems without determining certain consequences on those systems, as it is still assumed to be true for classical physics. Knowledge here is not a contemplative activity but is deeply and dynamically involved in what we are used to call reality. This statement, which can sound quite surprising and whose full meaning can be understood only later on, should be correctly understood. It does not imply that also the uncertainty relations themselves depend on certain measurements that we perform on quantum systems (and therefore on the knowledge that we acquire or do not acquire). This inference seems quite natural and in fact such an epistemology was, in some measure, supported by the young Heisenberg in his historical paper on uncertainty relations.[26] However, in Sects. 1.2.4 and 2.1.2 I have recalled that already in 1960 it was shown that this is not necessarily the case. The interpretation that I would like to support tells that uncertainty relations are a consequence of quantum interference terms and non-local correlations as manifested in superposition and entanglement, but not of the operations that we may perform on a quantum system. On the contrary, the results that we can obtain through these operations are deeply conditioned and framed by the constraints imposed by those terms through uncertainty relations, so that the results and their possible influences on a subsequent measurement cannot be controlled by the observer. Thus, when corrected, the point of view of Heisenberg, i.e. that detection is an uncontrollable operation, remains fully true and is even invested of a deeper meaning.

2.2.3 von Neumann's Problem

Projection Postulate

Thus, the problem of measurement seems to imply a conflict with quantum-mechanical (reversible) laws. The first scholar to have deeply considered this problem was J. von Neumann, and his treatment influenced generations of physicists and philosophers (for this reason, we might call this von Neumann's problem in short).[27]

---

[26] Heisenberg (1927).

[27] I address the reader to von Neumann (1932, Chap. 6). This textbook was for a longer time the sole formal and conceptual treatment of quantum measurement. Apart from QM, the contributions

Let us consider this problem in detail.[28] Suppose that the state before detection is (or has been prepared in) a superposition of eigenstates $\left|o_j\right\rangle$'s of the observable $\hat{O}$ that we wish to measure, i.e. (see Eq. (1.113))

$$|\psi\rangle = \sum_j c_j \left|o_j\right\rangle \ , \tag{2.42}$$

At a certain moment we need to assume that the apparatus has a basis $\left|a_j\right\rangle$ that can be put in correlation with the above terms. As an instance, I shall make use of the basis (2.26), $\{|d\rangle\ , |u\rangle\}$, which denote some up and down states (not necessarily paths in an interferometer). For this 2D case such a process could be mathematically described as follows:

$$|\psi\rangle\ |A_0\rangle \longmapsto c_d\,|d\rangle\ |a_d\rangle + c_u\,|u\rangle\ |a_u\rangle\ , \tag{2.43}$$

where $|a_d\rangle$ and $|a_u\rangle$ are the apparatus states corresponding to the system states $|d\rangle$ and $|u\rangle$, respectively, and $c_d, c_u \in \mathbb{C}$ satisfy the normalisation condition. Note that the transformed state is of the kind of the parallel entangled state (1.392). Here, we have assumed that the apparatus is in an initial ready state $|A_0\rangle$ uncoupled with the system's state (the state for the two systems of the left hand is a product state), while the object system is already in the superposition of the states up and down. For the sake of simplicity, I have suppressed the direct product symbol, but I stress that the apparatus kets $|a_d\rangle$ and $|a_u\rangle$ and the system kets $|d\rangle$ and $|u\rangle$ pertain to two different Hilbert spaces (or subspaces): $\mathcal{H}_\mathcal{S}$ for the object system and $\mathcal{H}_\mathcal{A}$ for the apparatus. Generalisation to the n-dimensional case using Eq. (2.42) is easily provided:

$$|\psi\rangle\ |A_0\rangle \longmapsto \sum_{j=1}^{n} c_j \left|o_j\right\rangle \left|a_j\right\rangle . \tag{2.44}$$

Successively, let us assume that, as a result of detection, we obtain the component $|o_k\rangle$ of the initial state that corresponds to $|a_k\rangle$ for the apparatus. As we have seen, there is no way to obtain this result from a superposition like (2.42) by means of a unitary evolution. Then, how can we reconcile the ordinary (unitary) quantum-mechanical time evolution with the experimental evidence of measurement? A conflict of this type (between a general theoretical framework and some experimental evidences) is very common in science—see, for example, Planck's problem in Sect. 1.1.1—and it is the very source of scientific research because it pushes to seek for other hypotheses beyond the existing theories.[29]

---

of von Neumann are fundamental in many fields, ranging from cellar automata going through game theory up to the foundations of modern computation.

[28] I follow here the exposition in Auletta et al. (2009a, Sect. 9.1).

[29] This is the essence of abduction and induction, a problem on which I shall come back later.

So far, it is clear that, in order to get a specific outcome (one of the eigenvectors of the measured observable), the quantum state needs to be changed. Let us see how and to this purpose make use of the density matrix formalism. The density matrix corresponding to the initial pure state (2.42) of the system can be written, in analogy with Eq. (1.366), as

$$\hat{\rho} = |\,\psi\rangle\,\langle\psi\,| = \sum_j |c_j|^2 |o_j\rangle\langle o_j| + \sum_{j\neq k} c_j c_k^* |o_j\rangle\langle o_k| \,. \tag{2.45}$$

The two sums on the RHS of Eq. (2.45) represent the diagonal and off-diagonal (coherent) parts, respectively. It is clear that, in order to obtain a determined outcome, the latter sum has to vanish during measurement: in this case, we must obtain a classical statistical mixture. But not all forms of mixtures are adequate in order to describe a measurement. The necessary condition for obtaining a determined result when measuring is that the apparatus $\mathcal{A}$ that performs the measurement on the object system $\mathcal{S}$ can extract information from $\mathcal{S}$ (can tell us what is the value of the observable we are measuring). This is in turn possible only if the state of the object system is a mixture of *eigenstates* of the measured observable. Thus, during the measurement process, we must obtain for the system a mixture of the type

$$\hat{\tilde{\rho}}_S = \sum_j |c_j|^2 |o_j\rangle\langle o_j| \,. \tag{2.46}$$

Of course, there is no unitary evolution that can account for the transformation from the state before the measurement, represented by Eq. (2.45), to the state after the measurement, represented by Eq. (2.46). In fact, we can write the change from Eq. (2.45) to Eq. (2.46) as

$$\hat{\rho} \longmapsto \sum_j \hat{P}_j \hat{\rho} \hat{P}_j \tag{2.47}$$

for the different eigenvalues o_j of the measured observable $\hat{O}$, and where $\hat{P}_j = \left|\,o_j\right\rangle\left\langle o_j\,\right|$ are the projectors on the different eigenstates of $\hat{O}$. The map (2.47) is a mathematical formulation of what is commonly known as *projection postulate*. Note that we need to distinguish between the notions of projector and projections. *Projectors* can represent any kind of state (and their combination also mixed states). *Projections* are the application of projectors to given initial states.

An Example

Let us consider the previous 2D state space example (see also Fig. 1.10, Sect. 1.2.2), which is the exact counterpart of Eqs. (1.359) and (1.362):

$$\begin{aligned}\sum_{j=d,u} \hat{P}_j \hat{\rho} \hat{P}_j &= |o_d\rangle \langle o_d| (c_d |o_d\rangle + c_u |o_u\rangle) \left(\langle o_d| c_d^* + \langle o_u| c_u^*\right) |o_d\rangle \langle o_d| \\ &+ |o_u\rangle \langle o_u| (c_d |o_d\rangle + c_u |o_u\rangle) \left(\langle o_d| c_d^* + \langle o_u| c_u^*\right) |o_u\rangle \langle o_u| \\ &= |c_d|^2 |o_d\rangle \langle o_d| + |c_u|^2 |o_u\rangle \langle o_u| \\ &= |c_d|^2 \hat{P}_d + |c_u|^2 \hat{P}_u \\ &= \hat{\hat{\rho}}_S, \end{aligned} \tag{2.48}$$

where I have made of the usual scalar products $\langle o_d \mid o_d\rangle = \langle o_u \mid o_u\rangle = 1$, $\langle o_d \mid o_u\rangle = \langle o_u \mid o_d\rangle = 0$.

The Mind of the Observer

One could be led to the conclusion that either the quantum-mechanical laws fail to describe measurement, and in particular its classical result, or these laws hold always true and detection becomes simply illusionary. Alternatives of this kind seem to exhaust all possibilities. In fact, the world is more complex and surprising as we often expect, and in most cases there are other solutions to a problem. In fact, von Neumann did not consider the classicality of the apparatus as a possible solution to the measurement problem, as Bohr was inclined to in his formulation of the complementarity principle (Sect. 1.3.3). The reason is that von Neumann would not introduce a duality (quantum/classical) *inside physics*. Then, it is quite natural to raise the question, as von Neumann himself did, of whether there is an additional (and not physical) kind of reality intervening in the detection process that cannot be led back to quantum-mechanical laws and could therefore account for this duality. He remarked that there was indeed a non-physical factor: the mind of the observer attending at the measurement process and being witness of the final detection event. In his own words, it is "an entity that brings out of the physical dimension". Therefore, it seemed quite natural to ask whether such a violation of the physical laws should be attributed to the observer's mind. In fact, von Neumann did not fully clarify which one of the following two interpretations he personally followed[30]:

- The mind could have indeed the power to change the physical state of the observed quantum system.
- The mind could perceive something that in fact does not correspond to what happens in the physical reality at the quantum level.

According to the first interpretation, the mind has a real power to modify the evolution of a quantum system. According to the second, it produces rather an illusionary representation of what quantum reality is. The first interpretation can be called a *subjectivist* one since it stresses the power of a subjective agency; the second one could be called an *objectivist* one since it assumes that quantum-mechanical laws faithfully represent reality, so that the latter reduces to such a description. As we

[30] As pointed out in Tarozzi (1996).

shall see, the community of physicists in fact divided itself in these two parties. I wish also to point out from the start, that none of these two interpretations is free of problems. In particular, the subjectivist one seems to run into the big difficulty of the violation of the causal closure of physical phenomena (on which I shall come back), and in fact many have subsequently departed from this idea. The objectivist interpretation appears to nullify the significance of the act of detection itself (and of our experience of the physical world).

Psychophysical Parallelism

This may be the reason why von Neumann himself seemed to support a kind of middle way, the so-called psychophysical parallelism, on the outline of the Occasionalists, a seventeenth-century Cartesian school of philosophy, according to which "the subjective apperception remains in fact external to the physical dimension" although everything happens "as it happened in the physical world". However, a difficulty here is represented by the fact that the psychophysical parallelism was historically introduced for explaining how there could be agreement between mind and physical reality when there are reasons for assuming that no causal connection between these two dimensions is possible (also this problem will be considered below). Here, at the opposite, these two parts show disagreement and not agreement (since the mind precisely perceives what *does not* happen in the physical world, at least without the concourse of the mind), what makes this interpretation rather obscure, or at least leaves unanswered the crucial question of why the observer's mind should be systematically misled when dealing with quanta without apparently showing a similar impairment when observing classical phenomena (this is also a problem for the objectivist interpretation). Moreover, psychophysical parallelism presupposes a kind of dualism between the mental and the physical dimensions, a hypothesis that was in fact introduced in philosophy by R. Descartes[31] but that sounds inappropriate to current scientific ears, since it accepts the ultimate existence of these two regions of the world without providing any possible answer why it should be so.[32] As a matter of fact, this third line of thought was not very successful in the following. In the next chapter, we shall consider the other two mentioned schools.

In whatever way we interpret the role of the observer, if we admit its relevance then one of the crucial consequences of this analysis is that there should be a dividing line between observed object and observer, what is called von Neumann's (but sometimes also Heisenberg's) *cut*. Now, the main problem is the significance of this cutting line

[31] Descartes (1641a).

[32] As we shall discover, as far as Descartes' standpoint would be limited to the statement that the operations of the mind respond to different processes than the operations of the body, I would agree. But this does not imply the existence of two separated substance (in the metaphysical language of that time).

and how to place it and on which grounds without introducing a certain arbitrariness in the foundations of the theory.[33]

2.2.4 What Are the Consequences?

Summarising, detection appears as a jump-like, abrupt, irreversible (non-unitary), nonlinear and even non-causal transformation that seems to represent a break with quantum theory and its basic laws.[34] It may even appear in contradiction with any physical explanation as such. This transformation has been called the *reduction of the wave packet* due to the fact that the induced disturbance, which affects the system when measured, reduces the wave function to one of the possible outcomes constituting its components. In other words, the superposition the system enjoyed before the disturbance seems cancelled and (through a projection) we get only one single component out of two or many ones: in the wave-packet reduction, the probability of the system being in one of the possible states that constitutes the superposition is reduced to zero except for the state in which the wave function collapses, being its value 1 (being detected, that state is by definition real and therefore certain). Moreover, since the wave function does not represent the distribution of its energy, as it has already been pointed out in Sect. 1.2.1, the wave-packet reduction is made without any transfer of energy. Naturally, after such a discontinuity the system (if it is not absorbed by the detector) resumes its continuous evolution on the lines of Schrödinger equation, so that its whole evolution could be understood as an alternation of collapses or scatterings and unitary evolution.[35] The notion of "wave-packet reduction" was first introduced by Heisenberg in 1927, as he studied the case of a highly energised electron in an atom, and subjected to a measurement of its position (Sect. 1.2.4).

Summarising the previous examination, we cannot extract the whole information that is contained in the state of a quantum system due to the basic discontinuity of quantum measurement, and it is such a discontinuity that opens a dangerous gap between theory and reality. Therefore, the main conclusion seems to be that the relation between theory and reality is at least disturbed, what puts upside down even the act of observing, suggesting a kind of idealistic epistemology. In fact, both the subjectivist and the objectivist interpretations share a sort of idealistic background: the former because it suggests that reality may depend on the knowledge that we have of it, and the latter also (although in a subtler way) since our mathematical formalism is finally what reality is.[36] It is not surprising that measurement has always

[33] This problem worried very much J. Bell: see Bell (1981).

[34] For a summary see Auletta (2000, Chap. 14). See also Conway and Kochen (2006).

[35] This wide subject is known as the path-integral approach: see Feynman and Hibbs (1965). For a synthesis see also Auletta et al. (2009a, Sect. 10.8).

[36] In fact, one distinguishes between the subjectivist idealism of Ancient philosophy whose reference scholar is the Greek philosopher Plato (V–IV century BC) and the objectivist idealism of modern

been considered one of the major problems of this theory and that the most brilliant minds have tried (most unsuccessfully) to solve it.[37] More recent developments of the theory (especially due to experimental progresses) and an extensive philosophical examination could put us in a better situation today.

2.2.5 *Schrödinger Cat*

The evolution through time of a quantum system as described by its wave function is determined by the Schrödinger equation. On this basis, quantum theory can forecast at all times the system's probabilistic evolution. However, this evolution is modified as soon as we want to verify it experimentally, since a disturbance is introduced in the system, and this disturbance corresponds to the second type of possible transformations of the system's state: detection. In a paper published in 1935, Schrödinger, weighing up all the problems arisen from QM and the interpretations provided up to the date, examined some possible consequences of the portray of reality suggested by quantum theory.[38] In this context, he proposed an ideal experiment, later known as the Schrödinger cat paradox, aiming at highlighting the problems that we are discussing here.

The experiment consisted of a cat caged in a box in which there is also a jar filled with volatile poison, and a hammer that can, any time, fall upon the jar and break it (Fig. 2.8). The hammer is connected to a trigger mechanism, which in turn

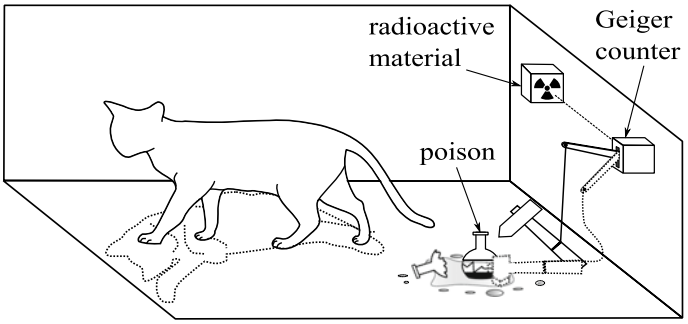


Fig. 2.8 Pictorial representation of the Schrödinger's cat thought experiment. Adapted from http://en.wikipedia.org/wiki/File:Schrodingers_cat.svg

times whose most important representatives are the German Philosopher Friedrich W. J. Schelling (1775–1854) and the German theologian and philosopher George W. F. Hegel (1770–1831).

[37] For a reprint of all main standpoints see Wheeler and Zurek (1983). See also Auletta (2000, Part IV).

[38] Schrödinger (1935). For historical context and reconstruction see Mehra and Rechenberg (1982–2001, VI, p. 738 ff.). For a synthesis of the problem with more recent approaches and experiments see Auletta et al. (2009a, Sect. 9.3) while for a short summary see also Auletta and Wang (2014, Sect. 9.8).

is attached to a Geiger counter activated by a radioactive material, say an atom. If the atom decays, the metre is activated and the hammer falls on the jar, thereby breaking it: the jar lets out the poison and the cat dies. If we happen to observe the box after a while, the cat will be either dead or alive according to whether the atom is decayed or not. But, given the quantum probabilistic nature of the radioactive decay (see also Sects. 1.2.1 and 2.1.2), the wave function that describes the system before observation finds itself in a superposition of "dead cat" and "alive cat", which appears to contradict our common experience. In fact, according to the quantum-mechanical description, we have an entanglement of two possibilities: undecayed atom—inactive hammer—living cat and decayed atom—active hammer—dead cat. What is most surprising is that, if we want to make sure what happens in the end and open the box, then, according to the projection postulate, the entanglement is cancelled and we are met with only one of the two possible states (and we find the cat is either alive or dead).

From a formal view point, let us denote with $|d\rangle_A$ and $|u\rangle_A$ the eigenstates corresponding to the decayed and undecayed atom, respectively. Similarly, $|a\rangle_H$, $|a'\rangle_H$ represent the hammer when active and not active, respectively, and $|\frown\rangle_C$, $|\smile\rangle_C$ dead and living cat, respectively. Then, after some time, the state of the system 'atom + hammer + cat' shall be described by the entangled state

$$|\Psi\rangle = c_d\,|d\rangle_A\,|a\rangle_H\,|\frown\rangle_C + c_u\,|u\rangle_A\,|a'\rangle_H\,|\smile\rangle_C\ , \tag{2.49}$$

were $c_d, c_u \in \mathbb{C}$ are coefficients subject to the usual normalisation condition. The state (2.49) is quite paradoxical, and, in particular, three facts have to be remarked: first, due to the correlation between the system, hammer and the cat states, the superposition character of the atomic state has been mapped onto the entangled state of the compound system (as it happens for ordinary measurement as displayed in Eq. (2.44)). As a consequence, a measurement of the state of either systems would immediately transfer the information onto the other (I take here the notion of information transfer in the most usual sense of the word but a careful examination will be provided later on). Second, this is an entangled state between a microscopic system (the atom) and a macroscopic one (the cat). Third, what appears weird is a cat being dead–alive, what sounds very bizarre from a biological point of view.

As I have recalled elsewhere, the Schrödinger's cat ideal experiment deals with two problems: the measurement problem (and, in particular, the difficulty that we cannot know a priori what is the actual value of an observable) and the issue whether it is quantum-mechanically possible to speak of an object as separated.[39] Actually, the main concern of Schrödinger was about the second point and seems to have been focussed on the indissoluble unity between observer and observed system, so that his preferred solution was rather to propose that the source of the paradox be in the *extrapolation* that we try to make by considering the cat (and its health) alone. We shall come back on this point of view that is certainly correct but also requiring some additional clarifications.

[39] Auletta (2000, p. 362).

This thought experiment reinforces what we have learnt so far: one could be tempted to conclude that either QM is in fact unable to explain our macroscopic physical world or the experience that we ordinarily have of this world is fully wrong, since we experience a gap between observation and the quantum state that we attribute to the system. As I have previously warned, such alternatives are generally not the sole solutions to scientific problems, and we should consider with a certain suspicion such apparently exclusive statements.

2.3 The Problem of Non-locality

2.3.1 The Paradox Proposed by Einstein–Podolsky–Rosen

Is QM Complete?

The three physicists Albert Einstein, Boris Podolsky (1896–1966) and Nathan Rosen (1909–1995), whom I shall collectively denote by the acronym EPR, in their famous 1935 paper,[40] set forth a paradoxical argument in order to assess whether QM could constitute a complete description of the physical world. The paper's intent is clearly shown in its starting words: bearing well in mind the distinction between an autonomous reality with respect to any theory, and the conceptual apparatus of the latter, it claims that the scientific theories should represent reality, which means that a theory is a successful model if it is correct and the description it gives is complete. With their own words[41]

> Any serious consideration of a physical theory must take into account the distinction between the objective reality, which is independent of any theory, and the physical concepts with which the theory operates. These concepts are intended to correspond with the objective reality, and by means of these concepts we picture this reality to ourselves. In attempting to judge the success of a physical theory, we may ask ourselves two questions: (1) "Is the theory correct?" and (2) "Is the description given by the theory complete?" It is only in the case in which positive answers may be given to both of these questions, that the concepts of the theory may be said to be satisfactory. The correctness of the theory is judged by the degree of agreement between the conclusions of the theory and human experience. This experience, which alone enables us to make inferences about reality, in physics takes the form of experiment and measurement. It is the second question that we wish to consider here, as applied to QM.

A theory is wholly correct if each of its elements has its counterpart in reality. And, likewise, a theory is complete if each element of reality has its counterpart in the theory. Correctness and completeness together would ensure the fully scientific status

[40] Likely the most influential one in the history of physics.

[41] Einstein et al. (1935, p. 138). For this section I invite the reader also to consider Auletta (2000, Part IX), Auletta et al. (2009a, Sect. 16.1), Auletta and Wang (2014, Chap. 10).

to our representation of reality. On this basis, the *criterion of reality* established by EPR tells us that if we are able to predict with certainty a property of an observable pertaining to a system without disturbing the latter, then this property must be real independently of any operation that we could perform on the system, and it is called an *element of reality*. In this argument, EPR assume a Separability principle or condition according to which, if there is no interaction between two systems (and *a fortiori* no disturbance), the reality of certain properties of one of the two systems cannot depend on whatever operation we perform on the other system.

EPR aimed to prove that there are elements of reality that do not have a counterpart in the quantum theory (are not described by it); more specifically, if we can predict the simultaneous value of two mutually non-commuting observables, we deal with a reality without a counterpart in the theory, since QM, according to the uncertainty relations (Sects. 1.2.4 and 2.2.2), tell us that we cannot have simultaneously determinate values for both observables.[42] Then, the key argument is to show evidence that, if the aforesaid criterion is assumed as a sufficient condition of the physical reality, then QM does not satisfy the condition of completeness. In such a case, QM would be rather a kind of statistical theory that describes reality only approximately.

It should also be noted that the argument is a negative one (is a counterargument). It is in fact very difficult and likely impossible, as shown by the Austrian–British epistemologist Karl Popper (1902–1994), to prove positively that a hypothesis or even whether a theory is complete or even correct, which would demand a kind of infinite or total experience of our world.[43] Einstein was well aware of this problem.[44] Then, the aim of EPR is to show a counter evidence that would put into discussion the completeness of quantum theory. It is also important to consider that the kind of incompleteness we are discussing here (inability to account for certain experimental results) is not to mix with the incompleteness of axiomatised formal systems. It seems in fact that any theory, as far as it can be axiomatised, shows indecidibility.[45] However, EPR were likely thinking that a theory should also display formal completeness.

The proposed criterion of reality implicitly assumes the validity of *omnimoda determinatio* (Sects. 1.2.1 and 1.2.4); it can have many applications since it only deals with the existence of elements of reality, without specification of their nature or reference to particular physical quantities. In other words, the criterion of reality has a philosophical character, since it is not restricted to a specific physical theory and neither to physical problems as such. Then, what is interesting in the EPR argument is that a philosophical assumption about reality (possessing even a metaphysical character) together with (i) physical laws (those of QM) and (ii) general scientific principles (that of separability) can lead to empirical consequences that are testable

[42] Auletta (2000, p. 534).

[43] Popper (1934).

[44] See Einstein (1930, 1934).

[45] Chaitin et al. (2011, pp. 75–79).

and in fact have been subsequently tested.[46] This shows that philosophical ideas are not irrelevant to the scientific enterprise, as mentioned in the introduction.[47]

EPR's Ideal Experiment

The argument considers two particles, which come into interaction in a time interval and after this time stop to interact. The quantum-mechanical theory holds that a certain wave function describes at any moment the system composed of those two particles, showing certain interdependencies (in fact, it is an entangled system). From this it follows that, as we measure the momentum or position of one of the particles, we can automatically know the momentum or position of the other one without having disturbed the system, and in such a way we would satisfy also the requirement of separability. The point is that if we decide to measure, say, the position we could have measured the momentum as well, or vice versa. This means that we could know both position and momentum of the other particle without having disturbed it, satisfying in this way the criterion of reality.

Let us suppose that the composite system with a single Cartesian component is described by the wave function in the position representation

$$\Psi(x_1, x_2) = \int_{-\infty}^{\infty} \psi_x(x_2)\varphi_x(x_1)dx, \tag{2.50}$$

where x_1, x_2 are the positions of particles 1 and 2, respectively, and the eigenfunctions of particles 1–2 have the following form (Eq. (1.129)):

$$\varphi_x(x_1) = \delta(x_1 - x) \text{ and } \psi_x(x_2) = \delta(x - x_2 + x_0), \tag{2.51}$$

where x_0 represents some fixed position and the second delta corresponds to the eigenvalue $x + x_0$ of $\hat{x}_2$. Being $\Psi(x_1, x_2)$ an entangled state, it is evident that it allows us to predict the position of particle 2 if we know that of particle 1 or vice versa (Sect. 1.4.2).

We know that we can describe the same system by expanding the wave function in another basis: this is performed by Fourier transforms and their inverses (Sect. 1.2.3). Then, we can write the wave function $\Psi(x_1, x_2)$ in the position representation explicitly in terms of the momentum eigenfunctions of particles 1 and 2 as

[46] This insight can be found in Tarozzi (1988), more recently reformulated in Auletta et al. (2009b): see also further references therein. It can be further considered Auletta (2011a, Sect. 2.2.6). It might be also interesting to have a look at Eddington (1939, pp. 46–48).

[47] Margenau insists that metaphysical conceptual elements are even necessary (Margenau 1950, pp. 12–13).

$$\Psi(x_1, x_2) = \frac{1}{2\pi\hbar}\int_{-\infty}^{\infty} e^{\frac{i}{\hbar}p_x(x_1-x_2+x_0)}dp_x$$
$$= \int_{-\infty}^{\infty} \psi_{-p}(x_2)\varphi_p(x_1)dp_x, \tag{2.52}$$

where the minus sign for the momentum value of the second particle is due to the fact that the two particles go into opposite directions, and the eigenfunctions of momentum for the two particles are (Eq. (1.148))

$$\varphi_p(x_1) = \frac{1}{\sqrt{2\pi\hbar}}e^{\frac{i}{\hbar}p_1x_1} \text{ and } \psi_{-p}(x_2) = \frac{1}{\sqrt{2\pi\hbar}}e^{-\frac{i}{\hbar}p_2(x_2-x_0)}, \tag{2.53}$$

with eigenvalues $p_x, -p_x$, respectively, where I have dropped the subscript x in the momentum variable for the sake of notation. Similarly, we can predict the momentum of particle 2 if we know that of particle 1 or vice versa. On the grounds of the aforementioned criterion of the physical reality, EPR concluded that both the momentum and the position of the particles were real *already before* performing any measurements, contrary to what the quantum uncertainty relations ensure us.

Description and Prediction

Note that EPR acknowledge both the descriptive and the predictive characters of the quantum theory: the *descriptive* because the two particles are described by a common wave function, and the *predictive* since it is thanks to the form of this description that we are able to predict the value of position and momentum of the second particle when we know the position and the momentum, respectively, of the particle that we have detected. However, the aim of EPR is to show that such description and predictions are nonetheless insufficient to catch some elements of reality. This would in turn imply severe restrictions in the scope of both quantum-mechanical description and prediction: as already mentioned, description would have only a statistical value and prediction (assuming that no violation of separability is thought to be possible) had a classical explanation in terms of an element of reality that exists (although possibly unknown) in full independence of both the quantum–mechanical description and the physical operations that we can perform on the involved systems.

EPRB Model

In order to set up a real experiment to check the EPR argument, we should deal with "continuous variables" such as position and momentum (that is, with observables that in some situations, especially when non-interacting, can possess a continuous spectrum), and that is rather difficult to realise in practice. This is why in 1951 the Anglo-American physicist David Bohm (1917–1992), proposed to abandon this kind of variables and instead use other observables which could take discrete values only,

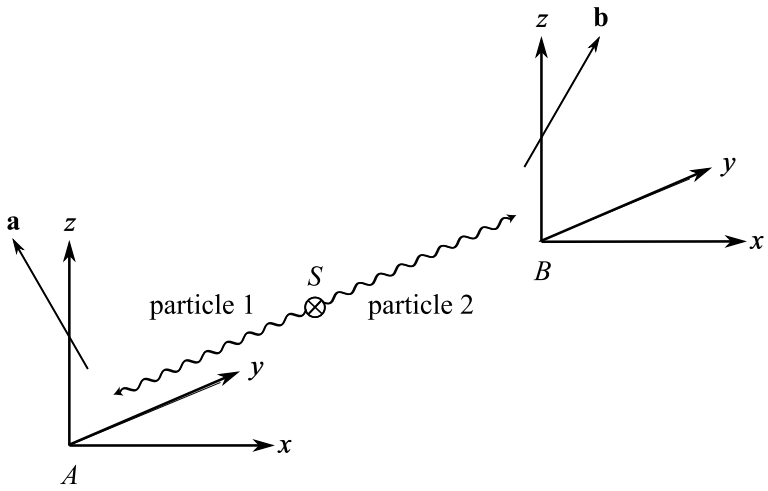


Fig. 2.9 Schematic overview of the EPR–Bohm experiment. Two spin $\frac{1}{2}$ particles are produced in a spin singlet state from a common source S (e.g. by decay of a spin 0 particle). After the time the two particles no longer interact, the spin of particle 1 in a direction **a** and the spin of particle 2 in another direction **b** are measured with two apparatus A and B, respectively. The Euclidean vectors **a** and **b** are taken to be unit vectors as they represent here spatial directions. Adapted from Auletta and Wang (2014, p. 304)

such as the spin of electrons (Sect. 1.3.1)[48]: in this way, the argument was greatly simplified and came a bit closer to its experimental verification. For this reason, one often speaks of the EPR–Bohm (or also EPRB) model.

Such a model consisted of two spin particles (fermions) that are in the singlet state (1.392), that is, the total spin of it is 0 (Fig. 2.9). Making us the z-direction basis (1.322), I rewrite here this state for commodity:

$$|\Psi_0\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow_z\rangle_1 \otimes |\downarrow_z\rangle_2 + |\downarrow_z\rangle_1 \otimes |\uparrow_z\rangle_2\right). \tag{2.54}$$

At a certain point the particles are let to break up so that later in time the distance is sufficiently great to make sure that they cannot interact (and in this way, we should satisfy EPR's requirement of separability). Taking into account that there is no interaction, the total state of the system keeps on being a singlet, while preserving the angular momentum, and this could be the classical mechanism allowing us predictions. Indeed, this entails the fact that, if a measurement of the spin component in the z direction of particle 1 leads to the result $+1$, the spin component of particle 2 in the same direction must give the value -1 and vice versa.[49] In fact, it is straightforward to show that

[48] The model was first proposed in Bohm (1951, pp. 614–623).

[49] I follow here Auletta et al. (2009a, Sect. 16.2) and Auletta and Wang (2014, Sect. 10.3).

$$
\begin{aligned}
(\hat{\sigma}_{1z} \otimes \hat{\sigma}_{2z}) \,|\, \Psi_0 \rangle \; &= \frac{1}{\sqrt{2}} (\hat{\sigma}_{1z} \,|\uparrow_z\rangle_1 \otimes \hat{\sigma}_{2z} \,|\downarrow_z\rangle_2 - \hat{\sigma}_{1z} \,|\downarrow_z\rangle_1 \otimes \hat{\sigma}_{2z} \,|\uparrow_z\rangle_2) \\
&= -\frac{1}{\sqrt{2}} (|\uparrow_z\rangle_1 \otimes |\downarrow_z\rangle_2 - |\downarrow_z\rangle_1 \otimes |\uparrow_z\rangle_2) \\
&= -\,|\, \Psi_0 \rangle \;, \qquad (2.55)
\end{aligned}
$$

where $\hat{\sigma}_{1z}$ and $\hat{\sigma}_{2z}$ are the z components of the spin observable of particles 1 and 2 (see Eqs. (1.324)), respectively, and use has been made of the eigenvalue equations

$$
\hat{\sigma}_z \,|\uparrow_z\rangle = |\downarrow_z\rangle \;, \;\; \hat{\sigma}_z \,|\downarrow_z\rangle = -\,|\uparrow_z\rangle \;, \qquad (2.56)
$$

according to the last line of Table 1.2, Sect. 1.3

The observable $\hat{\sigma}_{1z} \otimes \hat{\sigma}_{2z}$ represents the measurement of the spin components of both particles in the z direction. The above expression means that $|\, \Psi_0 \rangle$ is an eigenstate of $\hat{\sigma}_{1z} \otimes \hat{\sigma}_{2z}$ with the eigenvalue given by -1. However, it can be shown that this result holds true for the spin components of both particles in an arbitrary direction, namely, we have

$$
(\hat{\boldsymbol{\sigma}}_1 \cdot \mathbf{a}) \otimes (\hat{\boldsymbol{\sigma}}_2 \cdot \mathbf{a}) \,|\, \Psi_0 \rangle = -\,|\, \Psi_0 \rangle \;, \qquad (2.57)
$$

where $\mathbf{a}$ is an arbitrary unit vector and $\hat{\boldsymbol{\sigma}}_1, \hat{\boldsymbol{\sigma}}_2$ the Cartesian 3D spin matrices of the two particles. In other words, the state $|\, \Psi_0 \rangle$ is rotationally invariant. This is because it is an eigenstate of the magnitude squared of the total spin operator $(\hat{\boldsymbol{\sigma}}_1 + \hat{\boldsymbol{\sigma}}_2)^2$ with the eigenvalue 0. In this way, we have also confirmed that entanglement is a property of the state that is independent of the basis used (Sect. 1.4.2). In order to see this rotational invariance, it suffices to show that $|\, \Psi_0 \rangle$ turns out to be also an eigenstate of $\hat{\sigma}_{1x} \otimes \hat{\sigma}_{2x}$ and $\hat{\sigma}_{1y} \otimes \hat{\sigma}_{2y}$. This can be achieved either by explicit calculations or by expressing $|\, \Psi_0 \rangle$ in terms of the respective eigenstates of $\hat{\sigma}_x$ and $\hat{\sigma}_y$. I shall show both approaches for the x direction.

First, recall the action of the x Pauli spin matrix on the eigenstates of $\hat{\sigma}_z$:

$$
\hat{\sigma}_x \,|\uparrow_z\rangle = |\downarrow_z\rangle \;, \;\; \hat{\sigma}_x \,|\downarrow_z\rangle = |\uparrow_z\rangle \;. \qquad (2.58)
$$

Then, by applying the observable $\hat{\sigma}_{1x} \otimes \hat{\sigma}_{2x}$ to the singlet state, we obtain

$$
\begin{aligned}
(\hat{\sigma}_{1x} \otimes \hat{\sigma}_{2x}) \,|\, \Psi_0 \rangle \; &= \frac{1}{\sqrt{2}} (\hat{\sigma}_{1x} \,|\uparrow_z\rangle_1 \otimes \hat{\sigma}_{2x} \,|\downarrow_z\rangle_2 - \hat{\sigma}_{1x} \,|\downarrow_z\rangle_1 \otimes \hat{\sigma}_{2x} \,|\uparrow_z\rangle_2) \\
&= \frac{1}{\sqrt{2}} (|\downarrow_z\rangle_1 \otimes |\uparrow_z\rangle_2 - |\uparrow_z\rangle_1 \otimes |\downarrow_z\rangle_2) \\
&= -\,|\, \Psi_0 \rangle \;. \qquad (2.59)
\end{aligned}
$$

Let us now consider the second procedure and, taking advantage of the formulae (1.328a), make use of the following superpositions:

$$|\uparrow_z\rangle = \frac{1}{\sqrt{2}}(|\uparrow_x\rangle + |\downarrow_x\rangle), \quad |\downarrow_z\rangle = \frac{1}{\sqrt{2}}(|\uparrow_x\rangle - |\downarrow_x\rangle). \tag{2.60}$$

Then, we can rewrite $|\Psi_0\rangle$ in terms of $|\uparrow_x\rangle$ and $|\downarrow_x\rangle$ as

$$\begin{aligned} |\Psi_0\rangle &= \frac{1}{\sqrt{2}}(|\uparrow z\rangle_1 \otimes |\downarrow z\rangle_2 - |\downarrow z\rangle_1 \otimes |\uparrow z\rangle_2) \\ &= \frac{1}{2\sqrt{2}}[(|\uparrow_x\rangle_1 + |\downarrow_x\rangle_1) \otimes (|\uparrow_x\rangle_2 - |\downarrow_x\rangle_2) - (|\uparrow_x\rangle_1 - |\downarrow_x\rangle_1) \otimes (|\uparrow_x\rangle_2 + |\downarrow_x\rangle_2)] \\ &= -\frac{1}{\sqrt{2}}(|\uparrow_x\rangle_1 \otimes |\downarrow_x\rangle_2 - |\downarrow_x\rangle_1 \otimes |\uparrow_x\rangle_2), \end{aligned} \tag{2.61}$$

which, apart from the global phase factor -1, is a singlet state but with the projection of the spin along the x direction. Therefore, if the spin component in the x direction of particle 1 is measured first, followed by a subsequent measurement of the spin component of particle 2 in the same direction, then the two results must have opposite values.

Nature of the Problem

This shows that rotationally invariance is not as obvious as it seemed to spring from the mere (classical) conservation of the total angular moment. In fact, in QM the spin is quantised and the states display self-interference. Thus, if we measure a directional component of one of the particles' spin, the rotational invariance makes the measurement result of the other particle's spin opposite, whatever direction we have chosen and whatever result we have obtained on the first particle. But what is amazing is that we have a similar result also for other discrete quantum observables. In fact, if this is true for a pair of observables pertaining each to one of the two subsystems, this must be true for *any* pair of such observables.[50] As Schrödinger said, it is as the system would know the answer to *any question*.[51] In this way, if we do not draw the conclusion that QM is incomplete, it seems that we would be forced to assume that, e.g. the spin measurement on a particle could always determine the result of the measurement on the other particle taken separately, thus infringing the separability condition.[52]

Here lies, for EPR, the essence of the paradox:

- It strikes as absurd to maintain that the measurement of an observable (position, momentum or spin) in a particle should also, and at the same time, determine that of the other one, also when they are far apart from one another and no longer interact. There are two aspects to consider here. If the measurement of a particle's observable determines that of the other one, then the observables seem to do not

[50] Schrödinger (1936).

[51] Schrödinger (1935).

[52] See also Auletta (2000, pp. 536–537).

possess any *physical reality*, according to the criterion set forth above (a problem that we already met when dealing with measurement in the previous section). In this way, one could be led to the conclusion that there is no possible reality at all according to QM, since no basic criterion for discriminating between what is real and what is not could be applied (see also Sect. 2.1.1). This was a constant worry of Einstein.[53] The second aspect of the problem is how to understand causality if *distant* system can influence each other instantaneously violating the separability condition.

- If, at the opposite, we hold that the criterion of the physical reality and separability be true, we are led to the conclusion that the quantum theory cannot be considered to be a complete theory, and, as already mentioned, we need to postulate the existence of some supplementary parameters or hidden variables that quantum physics did not take into account, and which are necessary for determining, independently of measurement, the position or momentum (as with the case put forward by EPR). This can be called a *classical* realistic stance. Such hidden variables should explain that reciprocal influence of the two particles in terms of propagation of effects, so that that influence would be no longer instantaneous and both causality and separability would be satisfied. In this way, as anticipated, QM would at most be a statistical *description* of reality, and, in order to have a complete description of reality (which is classically deterministic by definition: see e.g. Fig. 2.7, Sect. 2.2.2) satisfying those *predictions*, we are forced to introduce a theory that would take into account the existence of those hidden variables.

2.3.2 Classical Physics and Locality

Limits on the Speed of Signals

The paper of EPR joined two different and very relevant issues: the problem of the *reality* of quantum-mechanical observables or of properties (and through them of physical systems as such) and the problem of *locality* (included separability). This connection served a specific purpose: the aim of EPR was to demonstrate that QM was not a complete theory since it failed to account for the reality of certain observables or properties, and any attempt at denying such a reality, according to them, would in turn generate a major conflict with the requirement of locality.[54] Why should this issue play a so relevant role in the controversy on QM? Because classical physics, after Einstein's relativity theory, set very specific requirements about the propagation of causal effects of our universe (and it is not by chance that the two points listed at the end of the previous subsection deal both with causality). It was not always so. As a matter of fact, classical physics up to the end of the nineteenth century did not acknowledge specific limits of the speed with which signals can be

[53] See e.g. Einstein (1948).

[54] For a historical reconstruction of the problems around the EPR paper see Jammer (1974, Chap. 6).

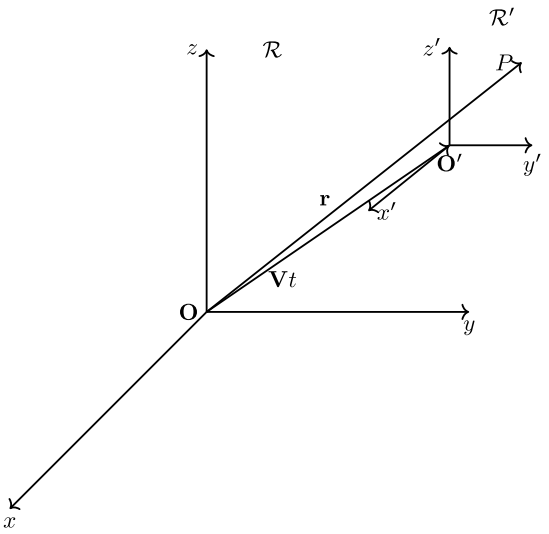


Fig. 2.10 Relationship between two different inertial reference frames $\mathcal{R}$ and $\mathcal{R}'$ under Galilei transformations. $\mathcal{R}'$ is in motion with respect to $\mathcal{R}$ with constant velocity $\mathbf{V}$. Adapted from Auletta et al. (2009a, p. 112)

exchanged and so events be also casually connected: Einstein discovered that, in order to causally connect two events, it is in fact necessary that at least a light signal be exchanged. I define a *signal* as any modification or perturbation of a physical or chemical 'medium'. Although this does not imply that speed could be also infinite,[55] studies of the American philosopher of physics John Earman proved that an infinite speed was not a contradictory concept in that context.[56] It was the great contribution of A. Einstein to have shown that nothing can overcome the speed of light (299, 792, 458 m/s) and that for material particles this is even an unreachable limit.[57] This is evident by considering that electrons (who have a small mass relatively to charge) can in fact reach a velocity that is near to that of light (about 85%), but their kinetic energy is no longer described by its classical-mechanical formulation as in the Hamiltonian given by Eq. (1.12), but satisfies the relation (1.3), that is, electrons at very high speed behave like kinds of photons, and c becomes an asymptotic limit.[58]

Galilean Relativity

The classical-mechanical view of relativity was expressed by the Galilei transformations (Fig. 2.10), after the name of the Italian astronomer, physicist and

[55] As pointed out in Maudlin (1994, p. 23).

[56] Earman (1986).

[57] Einstein (1905, 1909). For a summary of special and general relativity I invite the reader to read Einstein (1917, 1922, 2006). For a good exposition of relativity see French (1968). while for a clear presentation with philosophical examination of some consequences see Friedman (1983).

[58] French (1968, Chap. 1). Note that French already remarked that CM did not set specific limitations on the speed of bodies.

philosopher Galileo Galiei (1564–1642).[59] Note that Galilei transformations still rule non-relativistic QM. Consider the two reference frames $\mathcal{R}$ and $\mathcal{R}'$ such that $\mathcal{R}'$ moves with respect to $\mathcal{R}$ with constant velocity $\mathbf{V}$: note in fact that both CM and special relativity (SR) deal with inertial (not accelerated) motion and reference frames, while general relativity takes into account accelerated frames. We assume that at time $t = 0$ the origins of the two frames coincide, i.e. $O = O'$. The relation between the space-time coordinates in $\mathcal{R}$ and $\mathcal{R}'$ may then be written as

$$\mathbf{r} = \mathbf{r}' + \mathbf{V}t \quad \text{and } t = t' \,, \tag{2.62a}$$

where the vectors $\mathbf{r}$ and $\mathbf{r}'$ represent the position of a point particle P in $\mathcal{R}$ and $\mathcal{R}'$, respectively. Note that time is unchanged, that is, in such a theory is a fixed universal parameter, the so-called *absolute time*. This is a fundamental requirement for classical physics, e.g. the principle of action–reaction requires simultaneity and therefore absolute time.[60] Obviously, the inverse transformations are given by

$$\mathbf{r}' = \mathbf{r} - \mathbf{V}t \quad \text{and } t' = t \,. \tag{2.62b}$$

Galilean relativity tells us that all kinds of experiments performed in cabins moving with inertial motion relative to each other do not display observable differences or also give rise to the same results.

Michelson–Morley Experiment

The departure point of SR was understanding that the speed of light must be constant in vacuum, making c the basic relativistic constant. This was the result of Einstein's interpretation of a famous experiment performed by the American physicist Albert A. Michelson (1852–1931) and optician–astronomer Edward W. Morley (1838–1923) (Fig. 2.11). The Michelson–Morley experiment (a kind of interferometry) was performed in 1887 at what is now Case Western Reserve University, and is considered to be the first strong evidence against the theory of a luminiferous aether permeating the whole universe (this was understood as a privileged reference frame, a heritage of Newton's assumption of an absolute space), which was generally postulated at that time. The experiment compared the speed of light in perpendicular directions, in an attempt to detect the relative motion of matter through the stationary background luminiferous aether: since the revolution of the Earth around the Sun would happen on this background, the difference in relative motion (the so-called "aether wind"), according to the direction of the Earth, could be detected. Since no difference was remarked, the theory of the aether was discredited.

[59] Galilei (1632, 2nd day).

[60] As pointed out in Rindler (2001, p. 6).

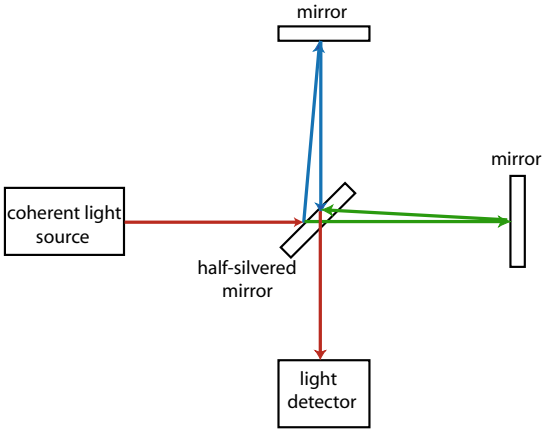


Fig. 2.11 The Michelson–Morley interferometer. A single source of monochromatic light is sent through a half-silvered mirror that is used to split it into two beams travelling at right angles to one another (top view). After leaving the splitter, the beams travel out to the ends of long arms where they are reflected back into the middle by mirrors. They are then recombined on the far side of the splitter in an eyepiece, producing a pattern of constructive and destructive interference based on the length of the arms. Any slight change in the amount of time the beams spent in transit would then be observed as a shift in the positions of the interference fringes. If the aether were stationary relative to the sun, then the Earth's motion would produce a shift of about 0.04 fringes. Adapted from Auletta et al. (2009a, p. 16)

Derivation of Lorentz Transformations

The genius of Einstein was to understand that such an experiment was showing that the speed of light is also invariant relative to the motion of the Earth, and therefore should be invariant everywhere in vacuum (having dropped the aether). Moreover, as anticipated, Einstein postulated that it represents the maximal speed attainable in our universe, and therefore c comes to represent an upper bound to the speed of information-conveying signals. In particular, this speed limit must apply to particles, since they can convey messages. As a consequence, the speed limit c actually does guarantee the invariance of causality. Let us call this *Einstein's locality* in short. Another consequence of the relativistic speed limit is that 'rigid bodies' and 'incompressible fluids' has become impossible objects, even as idealisations or limits. For, by definition, they would transmit signals instantaneously.[61]

On the basis of Einstein's locality, for light propagation, we need to set

$$c = \frac{dx}{dt} = \frac{dx'}{dt'} \tag{2.63}$$

[61] Rindler (2001, p. 56). However, it may be noted that in the original EPR experiment the two entangled systems are considered as connected by a rigid bar.

for two different reference frames $\mathcal{R}$, $\mathcal{R}'$, where for the sake of simplicity I am considering a 1D configuration space. Moreover, we need to assume that the motion of objects, if relatively slow, should be approximately described by the Galilei transformations (2.62), since this is what we actually observe at our scale. Let us further assume that velocities are equal and with opposite directions and with magnitude V. This means that the above Galilei transformations for space (in the 1D form)

$$x = (x' + Vt'), \quad x' = (x - Vt) \tag{2.64}$$

need to be corrected by a factor γ that goes to 1 when the speed goes to 0. In order to derive this factor, let us assume that x and x' are linear functions of x', t' and x, t, respectively, i.e.[62]

$$x = ax' + bt', \quad x' = ax - bt, \tag{2.65}$$

with $a, b \in \mathbb{R}$. Let us also assume that the motion of the origin of the unprimed reference frame as observed in the primed reference frame is such that $x = 0$ and analogously for the motion of the origin of the primed reference frame as observed in the unprimed reference frame, so that $x' = 0$. Then, from the above equations, we have

$$V = \frac{dx}{dt} = \frac{b}{a}. \tag{2.66}$$

By considering now a light signal moving at a constant speed c and expressing x and x' using Eq. (2.63), we obtain from Eq. (2.65)

$$cdt = (ac + b)dt' \text{ and } cdt' = (ac - b)dt. \tag{2.67}$$

By multiplying the LHSs and the RHSs of these two equations and making use of Eq. (2.66), we obtain

$$c^2 dtdt' = a^2 dtdt'(c^2 - V^2) \text{ or } (c^2 - V^2)a^2 - c^2 = 0. \tag{2.68}$$

This equation in a has a positive solution given by

$$\gamma := a = \left(\sqrt{1 - \beta^2}\right)^{-1}, \text{ with } \beta := \frac{V}{c}. \tag{2.69}$$

This, together the second of Eqs. (2.64)–(2.65) and the transformation for time that can be derived from the second of Eq. (2.67), allows us to write the new transformation rules that are known as *Lorentz transformations* (after the name of the Dutch physicist Hendrik A. Lorentz (1853–1928)) that can be written as[63]

[62] I follow here the derivation in French (1968, Chap. 3).

[63] Lorentz (1904).

$$x' = \gamma(x - Vt), \quad y' = y, \quad z' = z,$$
$$t' = \gamma\left(t - \frac{V}{c^2}x\right) = \gamma\left(t - \frac{\beta}{c}x\right), \tag{2.70}$$

where it is assumed that motion in the x direction will not affect the other coordinates. The inverse transformations are easily derived. These equations tell us that time is *relative* since it depends on the speed of the system: approaching more and more the speed of light it "slows down".[64] Moreover, it is also evident that length (or volume) is shrunk with $V \to c$. Note that the time difference between two readings of the same clock is less in the reference frame in which that clock is at rest than in any other frame. This allows us a definition of synchronisation. So, time dilation can be considered as a consequence of this definition of local simultaneity and of an experiment involving different clocks at different relative speeds. Similar considerations are true for length (or volume) contraction, so that in general we cannot perceive such differences in a single reference frame. This clarifies the notion of relativity.

Summarising, Einstein could derive the Lorentz transformations by stating the constancy of light speed and the SR principle tells us that all inertial frames are equivalent for the performance of all physical experiments.[65] Thus, each inertial frame is symmetric in the sense that a given physical experiment can be set up everywhere (space homogeneity), face in any direction (space isotropy), be repeated at any time (time homogeneity) and the outcome will be the same.[66] Note that the relativity principle plus the assumption of causality invariance (causal interconnections cannot be altered by changing reference frame) led us to the conclusion that there must also be an invariant velocity, but does not tell us which. If it would be infinite, Galilean relativity would follow, what shows that the assumption of the invariance of the light speed in vacuum is an additional assumption (ultimately derived from experimental evidence).[67]

Relativistic Physical Quantities

Let us now consider the expressions (1.3) and (1.4). If both need to be true, by simple substitution of $m = p_x/v_x$ in the former (for the 1D case) we get for the kinetic energy[68]

$$E = \frac{c^2 p_x}{v_x}. \tag{2.71}$$

Now, according to Newtonian mechanics, any difference in the kinetic energy of a body is due to the action of some potential, i.e. of some external force F, so that (always in the 1D case)

[64] French (1968, Chap. 4).
[65] Rindler (2001, p. 12).
[66] Rindler (2001, pp. 39–40).
[67] Rindler (2001, p. 15).
[68] I follow here the derivation in French (1968, Chap. 1).

$$dE = Fdx = \frac{dp_x}{dt}dx = v_x dp_x. \tag{2.72}$$

Multiplying the latter two equations, we obtain

$$EdE = c^2 p_x dp_x, \tag{2.73}$$

so that, by integrating both sides, we have

$$E^2 = c^2 p_x^2 + E_0^2, \tag{2.74}$$

where E_0 is some constant energy, which turn out to be the energy at rest. A similar argument is true for mass. Thanks to Einstein's relation (1.5) and Eq. (2.74), we can write the inertial mass that now becomes the rest mass as

$$m_0 = \frac{E_0}{c^2} = \frac{\sqrt{E^2 - c^2 p_x^2}}{c^2}, \tag{2.75}$$

where with *rest mass* it is understood the mass measured in the system's own rest frame (in which the momentum is zero) or, approximately, when the velocity of the system is very small relative to c (so that there are negligible relativistic effects on the mass). In general, for the mass of a moving body at speed v, we have

$$m(v_x) = \gamma m_0. \tag{2.76}$$

Thus, the energy of a moving body at speed v need to be modified according to

$$E(v) = \gamma E_0. \tag{2.77}$$

In conclusion, the new forms of momentum (1.4) and energy (1.5) in 3D become

$$\mathbf{p} = m(\mathbf{v})\mathbf{v} \text{ and } E = m(\mathbf{v})c^2 = \sqrt{(m_0 c^2)^2 + c^2 \mathbf{p}^2}. \tag{2.78}$$

Note that the mass of a body varies not only with speed. Due to Eq. (1.5), the gravitational binding energy of a body is a negative-mass contribution, so that the total mass of the body cannot be equal to the sum of the masses of its parts, at the opposite of what happens for the electric charge of classical electromagnetism that is strictly additive as a consequence of the linearity of Maxwell equations.[69] It may be further noted that that the vast amount of matter energy is constituted by thermal motion (heat), interatomic and intermolecular coherence forces (involved in chemical reactions), excited states of atoms (emitting radiation), nuclear bonds (the basis of atomic energy), but 99% is represented by the mass of the particle.[70]

[69] Rindler (2001, p. 8).

[70] Rindler (2001, p. 113).

Geometric Considerations

In this way, we have extended the notion of a 3D Euclidean space $\mathbb{E}^3$ (plus an independent time represented by a 1D Euclidean space $\mathbb{E}^1$) characterising Galilean relativity to a 4D space-time of SR. There are some elements to be pointed out. A Euclidean space is like a vector space but without origin. It is an *affine* space (an affine geometry is a Euclidean geometry but without using or "when forgetting" the metric notions of distance and angle). However, the Galilean space-time is not simply a Euclidean one. In fact, there is no absolute space here (but still absolute time), and thus it is not simply the product $\mathbb{E}^1 \times \mathbb{E}^3$, but is a so-called *fibre bundle*, with base space $\mathbb{E}^1$ and fibre $\mathbb{E}^3$.[71] A fibre bundle $\mathcal{B}$ is a manifold defined in terms of two other manifolds, the base space $\mathcal{M}$ and the fibre $\mathcal{V}$. In mathematics, a *manifold* is a topological space that resembles Euclidean space near each point in the sense that each point of an nD manifold has a neighbourhood that is homeomorphic to the Euclidean space of dimension n. In other words, a manifold can be curved but locally looks as a Euclidean space. The simplest example of fibre bundle is a product space in which a parallelogram is generated starting from a base space that is a line (Fig. 2.12a). Note that each fibre needs to be mapped to a single point in the base space.[72] Now, it turns out that Galilean particle histories (world lines) are cross sections of the bundle, i.e. continuous images of the base space $\mathcal{M}$ in $\mathcal{B}$ which meet each individual fibre in a single point (thus, for the annular in Fig. 2.12b, are any closed curves in $\mathcal{M} \times \mathcal{V}$).

Of course, in a relativistic context, there is no longer an absolute time either and thus its geometry cannot be represented by a fibre bundle. The geometry of the relativistic space-time is called *Minkowskian*, after the name of the Lithuania–German mathematician Hermann Minkowski (1864–1909)[73]: since both space and time are now functions of the speed, they become interdependent in a 4D space-time. This makes the square of the space-time distance

$$ds^2 = dc^2t^2 - dx^2 - dy^2 - dz^2 \tag{2.79}$$

between two events a relativistic invariant, where it is noted that time has an inverted sign relative to space dimensions, what means that it is not an ordinary spatial coordinate. Consequently, also the Dalembertian, after the name of the French mathematician, mechanician, physicist and philosopher Jean Baptiste Le Rond d'Alembert (1717–1783),

$$\Box = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2} \tag{2.80}$$

for 3D bodies becomes a relativistic invariant; note that it is the negative of the Laplacian (1.159) with the additional presence of the partial time derivative. In the

[71] Penrose (2004, Sects. 17.1–17.2), Geroch (1978, Chaps. 1–4).

[72] Penrose (2004, Chap. 15).

[73] Minkowski (1907–1908). See also Malament (2012, p. 110).

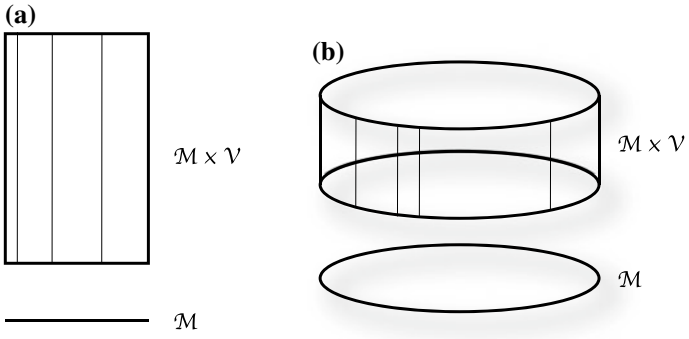


Fig. 2.12 **a** The simplest example of fibre bundle: the product space. Some fibres are shown (the thin vertical lines). **b** A more complex example: here the base space $\mathcal{M}$ is represented by a circle and the $\mathcal{B} = \mathcal{M} \times \mathcal{V}$ space as an annular. Note that the latter can also be twisted giving rise to Möbius strips

limit of an infinite dense collection of light rays and of world lines of infinitesimal point particles, we recover the concept of a continuous 4D manifold, i.e. a 4D 'space' with some smoothness properties.[74] Note that we have two idealisations here: (i) an infinitesimal point particle and (ii) a continuous manifold. We deal in particular with *differentiable manifolds*: a nD differentiable manifold is defined as a set of points tied continuously and differentiable, such that points in any sufficiently small region can be put in biunivocal correspondence with an open set of points of $\mathbb{R}^n$.[75]

Light Cones

According to SR, any physical system traces a trajectory in space-time and at each point of this space-time (that can be associated with an event E) a past and a future light cone are determined, whose (hyper-)surfaces represent the propagation of light which englobes both the past events that can have causal effect on that point and the future event on which E can have causal influence (Fig. 2.13). The surface of the light cones is the loci of events at zero interval from the "source" event (for which the square distance (2.79) is zero): this interval is called *light-like*. Since any exchanged signal must be inside such light cones (nothing reaching the light speed of massless photons),[76] subluminal causal effect between two events needs to satisfy $ds^2 < 0$ and the space-time interval between them is called *time-like*, while events that cannot be connected by light signals (and are outside the light cone), i.e. for which we have

[74] Misner et al. (1970, p. 10).

[75] Misner et al. (1970, p. 241).

[76] Although one could hypothesise the existence of superluminal particles (the so-called tachyons) without physical inconsistencies, it is also admitted that this move would be of no use for solving the EPR problem (Maudlin 1994, pp. 72–80).

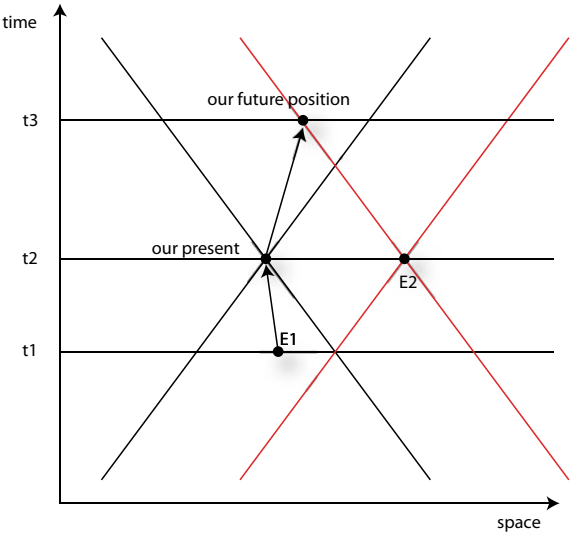


Fig. 2.13 Two light cones. For simplicity, space is 1D. Our present (our here-now) can be associated with whatever arbitrary event. Our trajectory in space-time is shown. The event E1 can influence ourselves at time t_2 but not the event E2 (whose light cones are in red). E2 will have instead an influence on our state at a later time t_3, when our trajectory will bring us to intersect the future light cone of E2

$ds^2 > 0$, are causally independent and are called *space-like*. Obviously, time (clocks) is (are) time-like while space is space-like. A further consequence is that events that can be viewed as simultaneous in certain reference frame are not necessarily so in another one if happening outside of its light cones. Thus, there is not only no universal time order but also no universal (but local) simultaneity.

Quadrivectors

Let us now introduce the space-time quadrivector

$$x^\mu := (x^0, x^1, x^2, x^3) = (ct, x, y, z) = (ct, \mathbf{x}), \tag{2.81}$$

where in a relativistic context we choose $\mathbf{x} = \mathbf{r}$. Greek indices are usual when dealing with space-time (when dealing with usual spatial coordinates we shall use Latin indices). Taking into account all what has been said, any general Lorentz transformation of the kind (2.70) and its inverse from a reference frame $\mathcal{R}$ to another reference frame $\mathcal{R}'$ take the compact forms

$$x'^\mu = \Lambda^\mu_{\ \nu} x^\nu \text{ and } x^\nu = \Lambda^\nu_{\ \mu} x'^\mu, \tag{2.82}$$

where I have adopted Einstein's summation convention (already introduced for the Levi-Civita tensor (1.231)): a repeated 'dummy' index is understood to be summed over, e.g.

$$\frac{\partial x^{\mu}}{\partial x^{\nu}}\xi^{\nu} := \sum_{\nu=0}^{3} \frac{\partial x^{\mu}}{\partial x^{\nu}}\xi^{\nu}, \tag{2.83}$$

where ξ^{ν} is some vector. The explicit matrix of the first Lorentz transformation (2.82) is

$$\begin{pmatrix} x'^0 \\ x'^1 \\ x'^2 \\ x'^3 \end{pmatrix} = \begin{bmatrix} \Lambda^0{}_0 & \Lambda^0{}_1 & \Lambda^0{}_2 & \Lambda^0{}_3 \\ \Lambda^1{}_0 & \Lambda^1{}_1 & \Lambda^1{}_2 & \Lambda^1{}_3 \\ \Lambda^2{}_0 & \Lambda^2{}_1 & \Lambda^2{}_2 & \Lambda^2{}_3 \\ \Lambda^3{}_0 & \Lambda^3{}_1 & \Lambda^3{}_2 & \Lambda^3{}_3 \end{bmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}. \tag{2.84}$$

A *contravariant* quadrivector is a vector x^{μ} transforming in this way. This is then true for the energy–momentum quadrivector

$$p^{\mu} := (p^0, p^1, p^2, p^3) = \left(\frac{E}{c}, p_x, p_y, p_z\right) = \left(\frac{E}{c}, \mathbf{p}\right) \tag{2.85}$$

of invariant length

$$p^{\mu}p_{\mu} = \frac{E^2}{c^2} - \mathbf{p}\cdot\mathbf{p} = \frac{p^2c^2 + m_0^2c^4}{c^2} - p^2 = m_0^2c^2, \tag{2.86}$$

where m_0 is the rest mass (2.75), and the relativistic expression of the total energy is given by the second Eq. (2.78). Note that the vector p_{μ} is a covariant vector, carrying a subscript instead of a superscript. To introduce covariant vectors is a crucial step due to the presence of Minkowski signature $(+---)$ as displayed in the square distance (2.79).[77] In all generality, I introduce the *covariant* quadrivector as

$$a_{\mu} = g_{\mu\nu}a^{\nu}, \tag{2.87}$$

where the $g_{\mu\nu}$ is the metric tensor

$$g_{\mu\nu} = g_{\nu\mu} := \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}, \tag{2.88}$$

which is non-zero only for indices $\mu = \nu$. Of course, we also have

$$a^{\mu} = g_{\mu\nu}a_{\nu}. \tag{2.89}$$

Note that its inverse $g^{\mu\nu}$ has exactly the same form so that $g_{\alpha\mu}g^{\mu\nu} = \delta^{\nu}_{\alpha}$. This is the *Minkowski metric* (and replaces Euclidean distance and metrics). Thus, contravariant and covariant quadrivectors are defined as

[77] But the convention $(-+++)$ is also used.

$$a^\mu := (a^0, \mathbf{a}) \text{ and } a_\mu := (a^0, -\mathbf{a}), \tag{2.90}$$

respectively. This allows us to define the relativistic scalar product

$$g_{\mu\nu}a^\mu b^\nu = a^\mu b_\mu = a_\mu b^\mu := a^0 b^0 - \mathbf{a}\cdot\mathbf{b} = a_0 b_0 - a_1 b_1 - a_2 b_2 - a_3 b_3, \tag{2.91}$$

which is obviously invariant under Lorentz transformation (determining Lorentz invariance), while the length $|\mathbf{a}|$ of a vector $\mathbf{a}$ is given by the square root of the scalar product

$$\mathbf{a}\cdot\mathbf{a} = |\mathbf{a}|^2 = g_{\mu\nu}a^\mu a^\nu, \tag{2.92}$$

according to the rules of the scalar product. This formalism allows us to express the space-time interval (2.79) in compact form

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu = g^{\mu\nu}dx_\mu dx_\nu = dx^\mu dx_\mu. \tag{2.93}$$

Note that it is required that

$$g_{\mu\nu}\Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta = g_{\alpha\beta}, \tag{2.94}$$

which is called the *Poincaré* transformation, after the name of the French mathematician, physicist, engineer and philosopher of science Henri Poincaré, and must hold for any tensor used in the theory.

Rotations and Boosts

Lorentz transformations need to take into account two different changes: changes in velocity or boosts (which connect two uniformly moving bodies), considered so far, and rotations. Rotations $\hat{\mathbf{R}}$ about the x, y, z axes, in agreement with Eq. (1.226), have components

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\theta_x & \sin\theta_x \\ 0 & 0 & -\sin\theta_x & \cos\theta_x \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta_y & 0 & -\sin\theta_y \\ 0 & 0 & 1 & 0 \\ 0 & \sin\theta_y & 0 & \cos\theta_y \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta_z & \sin\theta_z & 0 \\ 0 & -\sin\theta_z & \cos\theta_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{2.95}$$

respectively, with the angle $0 \le \theta \le 2\pi$. For boosts, I adopt the usual conventions (Eq. (2.69)):

$$\gamma = \cosh\phi\,,\ \beta\gamma = \sinh\phi\,,\ \beta = \tanh\phi\,, \tag{2.96}$$

where $0 \le \beta < 1$, and the parameter $0 \le \phi < \infty$, representing the hyperbolic angle that differentiates two frames of reference in relative motion, is called *rapidity*, being

a measure of relativistic velocity. The hyperbolic trigonometric functions are

$$\cosh x = \frac{e^x + e^{-x}}{2}, \quad \sinh x = \frac{e^x - e^{-x}}{2}, \quad \tanh x = \frac{\sinh x}{\cosh x}, \tag{2.97}$$

with $\cosh^2 x - \sinh^2 x = 1$. Thus, the Lorentz transformations (2.70) can be expressed as

$$\begin{aligned} x' &= x\cosh\phi - ct\sinh\phi, \quad y' = y, \quad z' = z, \\ ct' &= ct\cosh\phi - x\sinh\phi, \end{aligned} \tag{2.98}$$

while the boosts $\hat{\mathbf{B}}(\phi)$ in the x, y, z directions with rapidity ϕ have components

$$\begin{bmatrix} \gamma_x & \beta\gamma_x & 0 & 0 \\ \beta\gamma_x & \gamma_x & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} \gamma_y & 0 & \beta\gamma_y & 0 \\ 0 & 1 & 0 & 0 \\ \beta\gamma_y & 0 & \gamma_y & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} \gamma_z & 0 & 0 & \beta\gamma_z \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \beta\gamma_z & 0 & 0 & \gamma_z \end{bmatrix}, \tag{2.99}$$

respectively. They reader may recover from specialised textbooks the Lorentz inverse transformations in both cases. Note that parity and time reversal (Sect. 1.2.5) are Lorentz transformations. They represent special cases because they cannot be written as products of rotations and boosts. The Lorentz transformations make a group, called the Lorentz group, that is constituted by (i) the three *rotation generators* (which are angular momentum operators) $\hat{\mathbf{J}}$, with $\hat{\mathbf{R}}(\theta) = e^{\theta\cdot\hat{\mathbf{J}}}$, along the three Cartesian components defined by the partial derivatives of the rotation matrices with respect to three Cartesian angles of rotation as well as (ii) by the three *boosts generators* $\hat{\mathbf{K}}$, with $\hat{\mathbf{B}}(\phi) = e^{\phi\cdot\hat{\mathbf{K}}}$, along the three Cartesian directions constituted by the partial derivatives of the three boost matrices with respect to the rapidity parameter, i.e.

$$\left.\frac{\partial\hat{B}_x}{\partial\phi}\right|_{\phi=0} = \hat{K}_x. \tag{2.100}$$

Thus, the *generators* of the group in Cartesian components are

$$\begin{aligned} \hat{J}_x &= i\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad \hat{J}_y = i\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}, \quad \hat{J}_z = i\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \\ \hat{K}_x &= i\begin{bmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \hat{K}_y = i\begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \hat{K}_z = i\begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix} \end{aligned} \tag{2.101}$$

and any group element can be uniquely written up to a discrete transformation as

$$\Lambda = e^{i(\theta_n \hat{J}_n + \phi_n \hat{K}_n)}, \text{ with } n = x, y, z, \tag{2.102}$$

on the outline of the other generators introduced in Sect. 1.2.5. While rotations constitute a subgroup, boosts do not. Then, an infinitesimal generalised Lorentz transformation (on the outline of what we have done for momentum and angular momentum in Sects. 1.2.3 and 1.2.4) can be written as[78]

$$\delta x_\mu = i \left[\sum_{n=1}^{3} \theta_n (\hat{J}_n)_{\mu\nu} + \phi_n (\hat{K}_n)_{\mu\nu} \right] x_\nu, \tag{2.103}$$

where θ_n and ϕ_n are infinitesimal angles. This shows that the Lorentz group, that comprehends both rotations and boosts, is a Lie algebra. In fact, the generators (2.101) satisfy the Lie bracket relations in a similar way to, e.g. angular momentum (1.230):

$$\left[\hat{J}_l, \hat{J}_m\right] = i\epsilon_{lmn}\hat{J}_n, \left[\hat{J}_l, \hat{K}_m\right] = i\epsilon_{lmn}\hat{K}_n, \left[\hat{K}_l, \hat{K}_m\right] = -i\epsilon_{lmn}\hat{J}_n, \tag{2.104}$$

where ϵ is the Levi-Civit a tensor (see Fig. 1.20, Sect. 1.2.4). Note also that

$$\left[\hat{J}_m + i\hat{K}_m, \hat{J}_n - i\hat{K}_n\right] = 0. \tag{2.105}$$

As we shall see, when we add space and time translations, we obtain the Poincaré group, (after the name of H. Poincaré), which is a ten-generator non-Abelian (non-commutative) Lie group. Abelian (commutative) groups take their name from the Norwegian mathematician Niels Henrik Abel (1802–1829).

Invariance and Covariance

Physics needs to deal with both invariances (symmetries in general describe some kind of invariance) and covariances, where with *covariance* is understood a vector or a parameter that does change in different reference frames but is ruled by e.g. the Lorentz transformations. Indeed, physics could not exist if not able to deal with variable or relative quantities. What relativity tells us is that we have full relativity of two reference frames only when the physical systems move with constant velocity relative to each other and maintaining the same direction (are inertial). Thus, in many situations, there are identifiable differences in motion. In this context, the paradox of the twins, one (A) travelling in the space and the other (B) remained on the Earth, is particularly insightful. Suppose that A first goes away from the Earth, in a second moment reverse the direction moving towards our planet and finally arrives at his/her

[78] Schwartz (2014, Chap. 10).

start point and stops. Relativistic calculations show that time elapsed more slowly for A than for B (B is now older than A). Now, the question is: can we exchange the roles of A and B, perceiving B in relative motion relative to A? This is impossible, because there is no symmetry here: the crucial event is the inversion of the motion direction, which is in fact a change of the reference frame of A, while B remains in the same inertial frame throughout the whole journey of A. At the opposite, when only inertial motion is involved, no difference could be detectable since it is simply conservation of the state and could be considered as a null action. Of course, in a universe in which only inertial motion be allowed, no interaction could occur.

Dynamical and Static Relations

Coming back to our main examination, the relation between (special) relativity and QM is more tangled as it appears at first sight. The limit represented by the speed of light, i.e. *Einstein's locality*, as said sets limitations on the range of dynamical (and causal) exchanges in our universe. However, it does not tell us that those dynamical exchanges are the *only* way to connect physical systems. Nevertheless, this was an implicit assumption that almost nobody would put in question at the time of EPR. In other words, EPR as most physicists of their time assumed that *dynamical independence* is sufficient condition for separability. The reason is simple: to admit that physical systems can display interconnections that are not of dynamical (and *a fortiori*, not of causal) kind would imply that they can be related without displaying local physical interactions, and this seems to contradict the causal–local structure of SR.

From an epistemological point of view (and for the intelligence of what follows), we can distinguish between

- Dynamical relations (implying local interactions).
- Pure (static) correlations that do not imply exchange of signals (and thus neither local interactions).

I mention that the American philosopher and logician Charles S. Peirce (1839–1914), the father of pragmatism, had distinguished between dyadic and triadic relations.[79] To a certain extent, his division corresponds to the above one. In fact, he says[80] that "a dyad consists of two subjects brought into oneness" and among his favoured examples of triadic relations is "A gives B to C", which is clearly an action, what requires a kind of dynamics (a local interaction).[81]

[79] Peirce (1903a, pp. 272–273), Peirce (1903b, pp. 170–171), Peirce (1903c, 1.345–346), Peirce (1958, 5.472).

[80] Peirce (1958, 1.326).

[81] Nevertheless, starting from the fact that dyadic relations represent a kind of resistance to certain classes of action, he interpreted the latter as dynamical while understood triadic relations as displaying accordance with a law, what can be considered even as a reversal of the above distinction (Peirce 1958, 5.472). This was pointed out in Auletta (2011a, Sect. 3.2.3). See also Auletta (2016).

Thus, the question is if static relations exist at all. By now, let me stress that admitting the existence of static relations (pure correlations) in our universe would violate one of the basic assumptions of classical physics: that all natural systems can in principle be reduced to a sum of their parts. In fact, this is the ground of the reductionist methodology of CM. There is, however, evidence that such a possibility need to be seriously considered. In fact, during the twentieth century, also the theory of chaotic and complex systems have shown that such non-dynamical correlations play a role also in other domains of physics.[82] It may be recalled that one of the fathers of this theory was H. Poincaré.[83]

Locality and Separability

Thus, it would be suitable to distinguish here between two different principles[84]:

- *Einstein's locality* (or locality in short, when no ambiguity arises), concerning the impossibility to have superluminal signals and
- The *condition of separability*, namely, the assumption that when two systems are not dynamical connected, they can be considered as separated (they display no other form of connection).

What I am claiming is that what is in fact at stake with the EPR argument is separability and not Einstein's locality. The (now a classical) work of the physicists Sandu Popescu and David Rohrlich has in fact shown that these two principles set two very different quantitative bounds on an inequality pertaining to the family of the so-called Bell's inequalities.[85] This particular inequality is called CHSH inequality (after the name of the scholars who introduced it: J. Clauser, M. Horne, A. Shimony and R. Holt).[86] In fact, as we shall see, Einstein's locality sets a bound of four while separability sets a much stronger bound of two. On a formal plane, this shows that physical systems are thinkable that, although not violating locality, can nevertheless violate separability.

However, we should take into account that such a distinction between separability and Einstein's locality is a later development. As a matter of fact, to EPR locality and separability constituted two interconnected concepts: since locality is understood as a sufficient condition of separability, a violation of separability would denote a violation of locality. Thus, according to them there can be no physical influence of *any* kind (whether dynamical or static) on a distant system through local operations on another system that is isolated from the former. This is evident in Einstein's later reflection, when he reunified these two aspects under the principle of the local

[82] The reader might take into account the rather technical exposition in Ruelle (1990).

[83] Poincaré (1907).

[84] As first pointed out in Howard (1985).

[85] This distinction was first formulated in Popescu and Rohrlich (1994). See also Auletta (2011c), Ferrero and Sánchez–Gómez (2013). We shall come back on this problem.

[86] Clauser et al. (1969).

closeness of effects (*Nahewirkung*).[87] Note that this (as it is the case for the reality criterion) is a philosophical principle that expresses the fundamental idea of Einstein about causality. In the following examination, I shall first discuss the reasons of EPR, then deal essentially with separability, and finally come back to the local closeness of effects.

2.3.3 Reality and Correlations

Problems with the Notion of System

Let us try to explain the motivations of EPR for making (Einstein's) locality a sufficient condition of separability: if we allowed distant (static) correlations among parts of physical systems (renouncing to reductionist methodology), this would endanger the very notion of physical system (and therefore of reality). In fact, a well-defined physical system can always be isolated (at least in principle) from the rest of the world and relies on properties and processes that can be confined to a very specific (and controllable) portion of space (and time or space-time), i.e. that can be locally controllable, as, for instance, it happens in our laboratories. I remark here that in such a vision a physical system needs to be a kind of material object or body, although one could presume that also fields need to be somehow confined in some space-time region. Now, to admit that the system that we have in front of us in a lab can be interconnected with other systems out of any control and that therefore its properties could change in ways that are not locally predictable makes apparently crumble the notion of physical system as such (and also of experiment). In fact, I have already mentioned that this constitutes a difficulty for QM (Sects. 2.1.1 and 2.2.4).

Causal Closure

To deal with this problem we need to go back to pure philosophical issues. I ask the reader to be patient since this excursus will be important for the following, especially taking into account von Neumann's examination of measurement (Sect. 2.2.3) and its developments. G. W. Leibniz was the first scholar to assume that there are pure (static) correlations in Nature without causal interaction at all (an assumption that, in a time dominated by the reductionist methodology has contributed to his isolation from the rest of the scientific community). In fact, he says that "no created substance acts in true sense on others, in rigorous metaphysical terms, but everything comes out of its specific roots since every one represents separately the whole universe in its own way". (In such a context I recall that Leibniz posthumously inspired the "internalist" research program of CM between the end of the eighteenth century and beginning of the nineteenth century, based on functions like the Lagrangian and then

[87] Einstein (1948).

the Hamiltonian, in opposition to the Newtonian view based entirely on forces acting on a body from the 'exterior'.)

To understand Leibniz's standpoint, we need to come back to the mentioned problem of the mind–body relation as dealt by Descartes with. I cannot enter a problem on which there are perhaps millions of comments, rather I wish to recall a specific aspect of the issue. Descartes, although being a dualist, believed that the mind can act on the body and in particular it could cause a motion in the body like a deviation in the trajectory of a limb, e.g. an arm.[88] Leibniz, who had discovered that momentum was conserved in Nature (although there was some ambiguity in his definition of this concept), could not accept such an influence since it would imply a violation of the conservation of momentum and, as a consequence, also of the causal closure of physical phenomena,[89] according to which physical causal processes need to be self-consistent (any appeal to other factors than the physical ones is not allowed): the mind would indeed act as a kind of *Deus ex machina* from the exterior on the physical world by changing trajectories of physical objects. Let us call this fundamental assumption *Leibniz's principle of causal closure*. Of course, this has a direct relevance for the subjectivistic interpretation of measurement, as we shall see. Now, the followers of Descartes in the seventeenth century (the so-called Occasionalists) had no other choice than to renounce to this kind of action. Since they (following Descartes) conceived the mind and the body as two different kinds of substances, they coherently concluded that no natural causal interconnection at all could be conceivable between them. As a consequence, they also resorted to a *Deus ex machina*, but this time the Creator: it would be God who, being almighty, could connect a human act of will and a motion of the body.[90] They argued that, in the same way in which we are able to arbitrarily connect a concept with a phoneme in our languages, so God could make with mind and body. In this way, there would be no violation of the principle of causal closure. The reason why this philosophical school was called occasionalism is that it was thought that God would produce such a connection in the occasion of a specific human volition (in order to preserve individual freedom). But also this was unacceptable to Leibniz, since he could not support the idea of an arbitrary connection between mind and body and of God acting on specific occasions, opportunistically so to say and only by means of miracles (being this kind of God's action outside of the Nature's laws). On the other hand, he agreed with the Occasionalists that there could be no natural dynamical interconnection between mind and body. The solution that he could find was the pre-established harmony[91]: mind and body (and monads more in general, which, according to him, are kind of spiritual or energetic point-like systems constituting any body and mind[92]) show no dynamical interconnection at all but are only *statically* (and isomorphically)

[88] Descartes (1641a, pp. 25–26), and relative objections of Gassendi with responses.

[89] Leibniz (1686, 1695a, b).

[90] de La Forge (1666), Geulincx (1669), Malebranche (1674–1678). See also Auletta (1992).

[91] "Leibniz an des Bosses", 1st Sept. 1706: Leibniz (1875, II, 313–314); "Leibniz an des Bosses", 26th May 1712: Leibniz (1875, II, 444–445).

[92] Leibniz (1710–1712, Sects. 13–15), Leibniz (1712–1714, Sects. 7 and 56–57).

correlated. This connection, according to Leibniz, is again due to God, but is not arbitrary and would be established one time for all before the creation of the world.[93]

Symphonia Panta

There is a momentous consequence here, which was well understood by Einstein, as already mentioned: in such a universe, no physical system could be isolated, so that the very definition of physical system would be destroyed. There would be only *a single whole* universe, and it is not by chance that Leibniz used the Greek expression for denoting such universal harmony: *symphonia panta*.[94] Moreover, there is also a second consequence: to speak of an universe in which only static relations are real and not dynamical ones, makes such an universe indistinguishable from a pure thought–universe or even from a dream (representing the idealistic but objectivist view that we have already mentioned in Sect. 2.2.4).[95] In fact, it would be a universe without dynamical exchanges, a kind of representation or image (presumably a kind of representation in God's mind). The question then is the following: can QM allow for non-classical interconnections without destroying the notion of physical system and even of reality? Or should we embrace a consequent idealism, whether subjective or objective?

On the other hand, to admit the existence of correlations as the only interconnection among things has another crucial consequence that was drawn by Leibniz himself: there can be no space-time independently from the relations that exists between systems, so that also space and time are nothing else that kinds of order relations among things or monads.[96] This problem was at the hearth of another major intellectual battle, this time against the great British physicist and theologian Isaac Newton, who supported the idea of absolute space (and absolute time). The dispute broke out at the beginning of the eighteenth century.[97] Leibniz pointed out that, since several configurations of our universe (through rotation or displacement of the whole) can happen in the same absolute space and no difference could be observed, then absolute space must be a fiction. Similar considerations were developed for time (which is not trivial when considering that Galilean relativity assumed absolute time). Implicitly, Leibniz assumes here that there is no physical operation or action without some observable effect. In such a context note that Newton was unwilling to accept gravitation as a natural force since this would imply an action-at-a-distance (a violation of both locality and separability). Again, he resorted to a

[93] Leibniz (1702).

[94] As expressed in the famous treatise "Monadologie" Leibniz (1712–1714, Sect. 61). See also Leibniz (1710, I, Sect. 9).

[95] A problem pointed out in Peirce (1886). Peirce calls this the iconic side of signs and representations.

[96] On this difficult subject see the well-documented book (Earman 1989). Although I do not always agree with some of the conclusions it is a fundamental reference work.

[97] Leibniz (1715–1716). See also the historical reconstruction in Koyré (1957, Chaps. 7–11).

Deus ex machina and thought that gravitation could be conceived as an effect of the action of God in the world.[98] In fact, according to Newton, God acts in the world in order to keep bodies in motion that otherwise would sooner or later come to rest due to inertia (we see here how relevant was the later consideration of energy and of its conservation). Such an action required a ceaseless control on the universe and this happened through space and time, which were thought by Newton to represent the *sensorium* (or rather, *sensoria*) through which God embraces everything and therefore the context in which he acts in Nature. Leibniz rejected this conception because in his opinion introducing a direct action of God did not solve the scientific problem of the action-at-a-distance posed by gravitation and thought that its solution could only be found in a relational understanding of space and time guaranteeing local (although phenomenal) transmission of effects.

We might recall that Leibniz's interpretation of space and time in terms of relations seems to have been the first hint to Einstein's theory of relativity through the work of the German epistemologist and historian of science Ernst Mach (1838–1926).[99] It is interesting to note that Einstein was well aware that admitting a (non-dynamical) causal influx on a distant system would imply to consider physical systems as kind of "(quasi-)closed systems".[100] In other words, introducing static correlations would imply the existence of entities like the Leibnizian monads. So, Leibniz was finally very coherent.

Block Universe

Now, historical issues (even in the history of science) are always more tangled as they appear at first. Although Einstein had conceived space and time as relations, his fierce opposition to any kind of violation of separability apparently made him sympathetic to an interpretation of general relativity that many still consider the only acceptable one: the so-called *block universe*.[101] In fact, if in Nature only dynamical (causal) relations exist, space and time need to be conceived as manifestations of this basic causal structure. But if so, time (and space-time) is a derived concept and the universe could be conceived as a single atemporal block of 'causal' interconnections that exists from ever to ever: everything would be predetermined since ever. In the following, I shall show that the first part of the statement does not imply the second one.

---

[98] Newton (1687). See also Koyré (1966).

[99] Mach (1883). However, later on Einstein took more and more distance from E. Mach (Home and Whitaker 2007, Chap. 1), for reasons that are not very different from those that led him to took the distance from Bridgman's philosophy, as we shall see.

[100] Einstein (1948).

[101] An interpretation that it is likely to have been formulated by one of the fathers of the theory of symmetries, the German mathematician and physicist Hermann Weyl (1885–1955) for the first time (Weyl 1949). Interesting are the conversation between Popper and Einstein in which the former defined this vision as Parmenidean (Popper 1974, pp. 148–150). For a brief and untechnical examination of the problem see Auletta (2011a, Sect. 3.3.4).

It is not difficult to consider this view as a kind of re-edition of the understanding of Nature of the famous Dutch–Jew philosopher Baruch Spinoza (1632–1677),[102] a philosophical stance that (at least in the case of the Dutch philosopher) could be considered as a kind of 'monistic' occasionalism. This is really amazing, since in the 1930s Einstein had connected realism and locality (together with separability) precisely for avoiding the nullification of the individuality (isolability) of physical systems, and in the 1940s he seems to join the vision of a single block universe! This would make Einstein's and Leibniz's view very close and both dealing with an atemporal universe. The only difference between this view and Leibniz's one is that in the first case we deal with a 'causal' block. This shows that there is a potential conflict between the tenets of classical causality and of locality-separability, and, in fact, as we shall see, Einstein's own position seems more complex than what has been reported so far.

2.3.4 Possible Replies to EPR

In conclusion, the issues of reality and separability are the only two assumptions that do not depend on QM, and represent the premises from which we can infer the incompleteness of such a theory. It is not by chance that in the same year (i.e 1935) in which they published their historical paper, two fathers of quantum theory, Bohr and Schrödinger, tried to show the inconsistency of the argument by attacking either of these two assumptions, respectively. This will be the object of the next chapter. As we shall see, this involves very much Einstein, Bohr and Schrödinger in the examination and discussion of the same kind of problems. To these three giants we need to add Heisenberg to complete a poker that invites us to deeply rethink the way in which QM suggests us to deal with reality.

2.4 The Problem of Causality

2.4.1 Classically

Again on Continuity

The previous examination has raised the problem of causality many times. I have pointed out the causal structure of SR and shown that EPR aimed at proving that QM, if violating the criterion of reality, would also violate the relativistic principle of the local–causal transmission of signals. On the other hand, measurement raises the problem whether there are breaks in physical causality. Of course, CM is very differ-

[102] Spinoza (1677). The late Einstein was indeed very much influenced by Spinoza: on this subject and on the general religious background of Einstein see Jammer (1999).

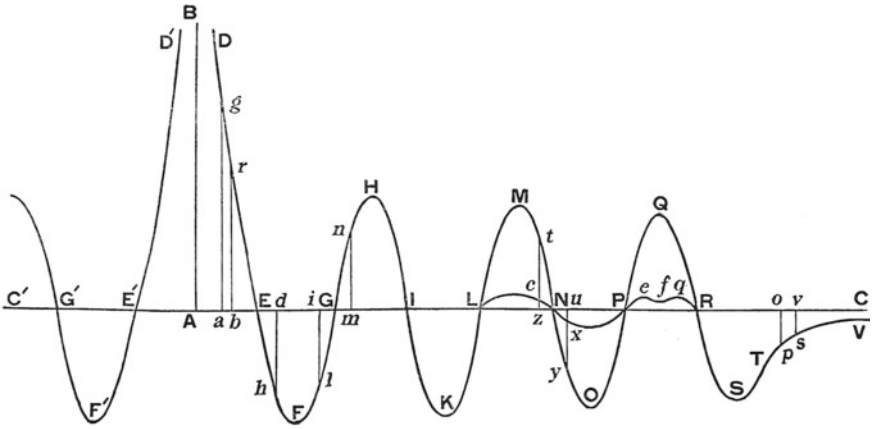


Fig. 2.14 Boscovich's model of the alternation of forces. The x axis represents the distance while the y axis the force (attractive for $y < 0$ and repulsive for $y > 0$). Adapted from Boscovich (1763, p. 41)

ent under this respect. As I have mentioned (in Sect. 1.1.2), from its beginning CM assumed a principle of continuity: nobody among the classical scholars raised the minimal doubt about the validity of this assumption. In the words of Leibniz, *natura non facit saltum*.[103] For these reasons, Leibniz introduced the adimensional monads (as recalled in Sect. 2.3.3) and the Croatian physicist and philosopher Rogerius Boscovich (1711–1787) the atoms of force[104]: these are ways to avoid the basic discontinuity that would result from a world constituted by indivisible quanta of matter. Of course, at that time many physicists were atomist or corpuscularist[105] but none had dared to assume that there could be breaks or jumps in physical processes (and few had apperceived the potential conflict between these two views). Boscovich, parallel to, and independently from, I. Kant, developed a theory according to which matter is constituted by some adimensional centres of force that were attractive or repulsive according to distance in a kind of wave-like alternation (Fig. 2.14).[106] Clearly, these are speculations, but we may recall here that the strong force is attractive at a distance of about one femtometre (10^{-15} m) losing its power at a distance of 2.5 fm but becomes repulsive at a distance of about 0.7 fm. It may be further recalled that up to that time matter was essentially considered inert, and these were the first attempts (on the line of Leibniz) at enabling nature with some internal activity.

[103] Leibniz (1715–1716, p. 25): "Continuatis autem in tempore, extensione, qualitatibus, motives, omnique naturae transitu reperitur, qui numquam fit per saltum".

[104] Leibniz (1712–1714), Boscovich (1754).

[105] See e.g. Locke (1689, B. II, Chap. VIII, parr. 9–13). See also Mccann (1994).

[106] Boscovich (1763), Kant (1747, 1756).

Sufficient Reason

The ground for assuming the principle of continuity was that admitting some break in continuity would represent a violation of the principle of sufficient reason (again formulated by Leibniz), and likely many scholars, even in our days, think so. The principle can be formulated as: of everything there is a reason why it so and not otherwise (note that the principle is much stronger than the no-action-without-effect assumption).[107] In fact, if a physical system can, e.g. jump from a location to another (as it happens when electron jump from one level to another in an atom: Sect. 1.1.2), we could not rationally explain how it came to the new location, provided that we cannot trace a trajectory (see also Sect. 2.2.2). This would be a necessary condition for violation of causality, as a particular manifestation of the principle of sufficient reason: in fact, any gap in continuity would certainly impair any causal explanation of classical kind. It is not by chance that N. Bohr, in his famous presentation of the complementarity principle (Sect. 1.3.3), considered as a major threat to causal explanations quantum discontinuities (or quantum jumps),[108] the so-called *quantum of action*, that is, the minimal quantity of energy (and momentum) that, following Planck's and Einstein's results, Bohr assumed to determine the basic discontinuous processes of the quantum world. In fact, whenever there is a causal relation among physical systems it is necessary that they locally interact, what means that they need to exchange dynamical quantities like energy and momentum.[109] In this classical consideration of the problem, it is interesting to note two crucial issues:

- Causality is treated exclusively in terms of *mechanical* causality. In fact, classical physicists expressed an abiding faith into universal determinism. *Determinism* could be expressed saying that it consists in assuming that, for every state of any system (included the whole universe), there is a *single* future state at any time t and a single past state at any time $-t$ into which and from which the system evolves, respectively. As we shall see, this seems to be a fundamental requirement of any physical theory. Note that this problem needs to be kept distinct from the issue of reversibility, as we shall see later. However, CM further assumes that a system evolves with *all its properties* in a way that is fully predictable, according to the *omnimoda determinatio* assumption (Sects. 1.2.1 and 1.2.4).[110] In the words of Planck, an event is causally determined if it can be predicted with certainty.[111]

---

[107] Leibniz (1712–1714, Sect. 32): "Nous considerons qu'aucun fait ne sauroit se trover vray ou existant, aucune Enontiation veritable, sans qu'il y ait une raison suffisante, pourquoy il en soit ainsi et non pas autrement, quoyque ces raisons le plus souvent ne puissent point nous tre connues". See also Leibniz (1677). About the potential conflict between the classical principle of sufficient reason and QM see Auletta (2006).

[108] Bohr (1928).

[109] Kistler (2006, p. 9).

[110] On this the work of the French physics and mathematician Laplace is paradigmatic Laplace (1825).

[111] Quoted in D'ariano et al. (2017, Sect. 5.1). I fully agree with what the authors say here about the confusion causality–determinism.

This is the essence of what can be called *classical* or mechanical determinism.[112] Although the issue of determinism deserves a critical analysis, it becomes universal custom that the term *causality* refers to mechanical (dynamical) causality.

- The connection between causality and the principle of sufficient reason is the manifestation of a tenet again shared by everybody until the beginning of the twentieth century. According to this view, causality and Nature's laws essentially coincide because to causally explain a phenomenon means to lead it back to the iron and rational laws that rule everything.[113] Such a view is partly justified by the fact that the laws of Nature provide a regular bridge between some antecedent and some consequent state of the involved systems.[114] This standpoint is still expressed by Bohr in his famous (and quoted) paper on complementarity in 1928. At the opposite, Max Born has clearly refuted the identification of causality with laws,[115] and this is the standpoint that I shall support here.

Of course, these two aspects are closely related. In fact, to assume that laws and causal-mechanical explanations coincide means to say that laws completely describe *single* happenings in nature, what would certainly satisfy EPR requirement of the reality of observables and properties and thus also of the completeness of a theory in a classical sense (Sect. 2.3.1). As a consequence, the theory would also be deterministic in a classical sense, provided that continuity is satisfied.[116]

Are There Causes at All?

Before going on in our argumentation, allow me to stress that the notion of causality is far from being obvious. It may be recalled that some philosophers have put this notion into question. The first to have done it systematically it is assumed to have been the Scottish historian and philosopher David Hume (1711–1776). He pointed out that the notion of cause is derived (inferred) from our experience and is not the result of a direct perception.[117] For this reason, he affirmed that it is the result of a pure association (among direct perceptions or their impressions in memory) whose "necessity" cannot be grounded. I prefer to speak of *regularity* instead of necessity. Nevertheless, it should be stressed that the target of Hume's criticism is not the notion of causality as such but rather the contribution of Nature's *laws* (the "necessity" component) to causal explanations. However, the fact remains that

---

[112] D'ariano et al. (2017, Sect. 5.1).

[113] In Mehra and Rechenberg (1982–2001, VI, p. 678), a definition of a physics dictionary has been reported: "The physicist considers causality as identical with determinism, that is, with the unique fixing of the future events by the present ones according to the laws of nature".

[114] Poincaré (1911, p. 49).

[115] Born (1949b, Chaps. 2–4).

[116] Although a deterministic mechanics is still possible that does not satisfy continuity: see Auletta (2004).

[117] Hume (1777, pp. 32–47).

according to him laws of Nature be the result of frequent associations, a point that deserves a specific analysis below.

Also in more recent times, the British philosopher Bertrand Russell (1872–1970) has criticised this notion[118] again on the ground of a strict connection between laws and causal explanations, followed in this by many philosophers still today, like the American philosopher of science John Norton, who supports Russell's idea that science deals with relations among quantities,[119] which in my parlance are Nature's *laws*.

The causal structure of both special and general relativity demands further scrutiny of this problem. First of all, let me clarify with simple examples what in scientific context is meant with causality. When we speak of *causal mechanism*, we may think at something like the photoelectric or Compton effects. In the first case, e.g. a photon is absorbed by an electron and the latter evades the atomic bond. In cases of this kind, we say that the photon's absorption has *caused* such and such behaviour of the electron as well as the scattered electron, if detected, has caused such and such effect on our detectors. However, such processes could be perfectly probabilistic. It is only the classical confusion between law and causality that expresses unsatisfaction with probabilities and additionally requires that it should be also classically deterministic.

2.4.2 Quantum-Mechanically

Random Events

The previous examination makes more understandable the resistance to accept, and even the overt opposition to, the new view of the physical world suggested by QM. With the words of Bohr[120] QM suggests us to accept a *complementarity* between (see also Sect. 1.3.3)

(i) The spatial–temporal localisation of the system as it results in a detection event, which is discontinuous (Sect. 2.2.4) and
(ii) An analogue of classical causal description according to the quantum-mechanical laws (where law and causality were considered as identical), which needs to be continuous.

Bohr was aware that to accept the former means to introduce a kind of irrationality into the heart of physics. As a matter of fact, when QM was finally accepted by the community of physicist (what can be said to have definitively happened after the Second World War), a significant part of this community shifted from supporting a universal classical determinism towards a position that makes of chance the fundamental explanation of many physical processes. It may be mentioned here that

[118] Russell (1912–1913).
[119] Norton (2003).
[120] Bohr (1928).

the Darwinian theory of evolution, and in particular the stress on chance mutations and chance selection events by the so-called Neo-Darwinian synthesis, may have contributed to the affirmation of this new standpoint.[121] This could be considered one of the major conceptual changes never happened in science.

Before entering some details, let us stress that, in a quantum-mechanical context, to admit this kind of universal chance is strictly connected to admit (detection) events as the only form of reality, an idea that has been supported, at least for a certain time, by the Austrian experimental physicist Anton Zeilinger.[122] The reason is quite clear and was also evident to Bohr: events are the only reality out of the quantum world that we can make experience of, although, as stressed, they represent precisely the source of some of the major quantum discontinuities. They happen randomly by definition. In the following, we need a careful examination of the consequences of this standpoint in order to ascertain whether detection events can in fact be considered as the only form of reality.

A Probabilistic Determinism

There is also a second source of the perceived difficulty to deal with causal explanations in QM: the multipath dynamics that is a consequence of superposition (Fig. 2.7, Sect. 2.2.2), an aspect that appears to have been underestimated by Bohr (but stressed by Schrödinger). In fact, if we are unable to tell which particular path a system can take from an initial position to a final one, again no causal explanation seems possible.

Thus, it seems that both the basic principles of QM (quantisation and superposition: Sect. 1.2) lead us to a conflict with a causal understanding of physical phenomena. Nevertheless, it is likely that we can find here not only the source of troubles but also a hint to a possible solution. In fact, such a multipath dynamics is ruled by the laws of Nature, in particular, by the Schrödinger (or Heisenberg) equation. However, this does not yet clarify how is it possible to have detection events that seem to represent a break with those laws (Sect. 2.2). This conflict could be solved if we are willing to admit that laws have in QM a general or global value and not a particular or local one. In this sense, we could still assert that QM is both deterministic and complete but not in a classical sense, since claiming the impossibility to deal with single events (and therefore attribution of properties) at the level of the laws, what means acknowledging that there is not a single actualised future (neither a single actualised past) for any state of a quantum system as far as we are concerned with the intrinsic multipath, superposition-like, dynamics. In this perspective, quantum laws only ruled the probabilities to obtain events (including the interference terms in those calculations), according to Born's point of view (Sect. 2.1.2) but *not the*

[121] Darwin (1859). On these problems, see also Auletta (2011b, Chap. 9), Auletta and Pons (2013).
[122] Zeilinger (1999, 2000).

single events themselves.[123] In fact, Born tells us (see Sect. 1.2.1)[124]: "The motion of particles conforms to the law of probability, but the probability itself is propagated in accordance with the law of causality", meaning with this that probability amplitude spread and evolve according to deterministic physical processes. In other words, we would have a determinism (mapping univocally a quantum state into another, past or future, quantum state) that is also probabilistic (since no event can be predicted or retrodicted on the basis of this evolution), a probabilistic determinism in short. I call *quantum determinism* this determinism without *omnimoda determinatio*. The following examination will tell us if this interpretation is viable and whether it can solve in fact the completeness problem raised by EPR.

Laws of Nature

In other words, I am supporting a dissociation between laws of Nature and causal explanations. Although laws of Nature clearly enter in causal explanations as components of the latter, those explanations cannot be reduced to laws. To show this, let us focus on some fundamental characters of the laws of Nature. First of all, although what we call laws of Nature are clearly a theoretical construct, they should nevertheless express regularities that are objective and independent from the way in which we formulate them. However, they cannot be reduced to occasionally observed regularities either, according to a Humean frequentist standpoint. In fact, it can be shown that Humean uniformities are neither sufficient nor necessary condition of the laws of Nature.[125] There are many regularities that do not deserve the name *law of Nature*. On the other hand, on that basis factually excluded possibilities would be interdicted by the laws of Nature, what appears absurd. In particular, the Humean standpoint is not able to deal with intrinsically probabilistic laws of Nature, what is especially relevant for QM. In fact, probabilistic laws allow any possible distribution of the facts that they govern: although some are more probable than others, no distribution is excluded and thus a probability cannot be identified with a limiting relative frequency.[126] As mentioned, statistics and probability are two different concepts.

Thus, the ground of the laws of Nature must be found elsewhere. It has been proposed that laws of Nature express relations among properties of the physical systems.[127] However, I have mentioned that properties in QM are *prima facie* not ruled by a law-like dynamical evolution (Sects. 2.1.3 and 2.2). Moreover, laws do not express connection among static elements but rather they relate certain *variations*

[123] The reader may also have a look at Auletta (2011b, Chap. 1), Auletta and Wang (2014, Sects. 7.3 and 9.5).

[124] Born (1926). For historical reconstruction see Mehra and Rechenberg (1982–2001, VI, p. 678 ff.).

[125] Armstrong (1983, Chaps. 2–3). This philosopher has contributed very much to our understanding of the laws of Nature.

[126] Armstrong (1983, p. 31).

[127] Bird (2007). See also Carroll (1994) for further references to the philosophical discussion on the laws of Nature.

with other variations. Therefore, it would be more appropriate to say that they express relations *among observables*. Consider the Schrödinger picture: in particular, the von Neumann equation (2.17) tells us that the density matrix of a system (which represents a state but is also observable) stands in a commutation relation with the Hamiltonian, and in the particular case in which it is a stationary state (Sect. 1.2.3), commutes with it. On the other hand, the Heisenberg picture as expressed by Eq. (2.13) also tells us that the evolution of observables stand in a commutation relation with the Hamiltonian and, in the particular case in which it is a conserved quantity, the observable commutes with the Hamiltonian (Eq. (1.320)). It is evident that energy, and in particular the conservation of energy, needs to play a central role for the dynamical evolution of (quantum) systems.[128] A further example is represented by the uncertainty relations (1.286) that again express relations among observables. The same is true, as we shall see, for the complementarity principle and also for the CPT theorem (Sect. 1.2.5) and even to a certain extent for the correspondence principle (Sect. 1.3.2). Also classically, the third Kepler's law expresses a relation between a period and a distance as well as Newton's gravitation law a relation between a force and distance, and more generally the dynamics of classical systems is ruled by the Poisson brackets (1.216). I recall that also E. Mach and the German philosopher Ernst Cassirer (1874–1945) supported such a broad functionalist point of view.[129] In conclusion, we can say that laws ultimately express correlations among observables and the variations of their possible values or probabilities to obtain those values. Since the laws of Nature impose specific invariances, like the temporal symmetry of the quantum dynamics, they ultimately rely on or are connected with symmetries.[130] A remark is opportune here. The notion of symmetry is wider than that of correlation (and order). In fact, systems in a maximal entropic state (like a uniform distribution of gas molecules in a cavity) also show at a certain level a symmetry, although by definition the molecules are not correlated. However, since a uniform distribution is again what we expect if sufficient time elapses, a deviation from this symmetric state represented an anomaly.

Additional Factors

The standpoint introduced here is less innocent than it seems. In fact, we were forced to admit that causal explanations deal with additional factors that are not (and cannot be) taken into consideration in the formulation of the laws of Nature. These additional factors could be the spatial or temporal localisation of the system under consideration (as Bohr said) or the interactions that it can have with other systems (as assumed by Heisenberg), and especially the mentioned measurement procedures, but also the physical context that these procedures determine. In other words, *boundary* and

[128] See also Kistler (2006, p. 67).

[129] Mach (1905), Cassirer (1910).

[130] See van Fraassen (1989), especially p. 275.

initial conditions would be a fundamental ingredient of causal explanations that need to be kept distinct from laws and regularities.[131]

Note that there is a certain ambiguity in the way in which causal mechanisms are understood. They could be understood as a general explanation of how a kind of physical system reacted when there are some general conditions, like an electron jumping an orbital level in an atom due to the absorption of a photon. At this level, we could be tempted to say that in this mechanism there is nothing else apart from the laws of Nature.[132] Here, we do not need to introduce initial or boundary conditions if not again in the idealised case (at the level of the general model). However, a causal explanation dealing with quantum systems (as for CM) need often to account for the singular physical context and the singular causal effect, given the role that irreversible (detection) events can play. For instance, we could be interested in knowing which one of a couple of photons has impinged on a detector (in the following, we shall consider several experiments dealing with this problem). In such cases, initial and boundary conditions are an essential ingredient. Note that SR suggests the same conclusion, as far as it is essentially a theory of the causal relations that can be established from an event (and to an event) thanks to future (past) light cones.

The consequence drawn here could be understood as a reformulation of the complementarity principle presented previously, but with a fundamental difference. In fact, I shall suggest the opportunity to broaden our understanding of causality beyond the classical dynamical-mechanical explanation, in order to take into account some of the involved factors that, being typical of quantum-mechanical systems, can help us to make compatible the happening of irreversible events and the existence of reversible laws. Since the only non-classical factor that here intervenes in quantum systems is represented by the evolving probability amplitudes and the mentioned interference terms giving rise to correlations (see also Sect. 2.1.2), I am claiming that it is on this issue that we need to focus our attention if we like to solve this problem in a satisfactory way.

Causal Processes Are Not Reversible

I shall show now that also classically, laws and events (it does not matter here whether in first instance they are considered random or not) represent two distinct aspects of the problem, and therefore there must be a dissociation between reversible laws and irreversible causal processes. If the classical laws ruled all the specific conditions at contour or all of the initial conditions (i.e. the details of the physical world, the physical context in which a certain dynamical process happens), this would imply that also any causal chain could be developed in the reversed direction. Indeed, many assume that it should be so (Sect. 2.4.1). Nevertheless, this inference is not

[131] It seems that E. Anscombe has been the first scholar to have proposed a particularist view of causality (Anscombe 1971).

[132] This why Margenau tells us that causality is nothing else than a law ruling the transformation of a physical state into another minus the temporal dimension (Margenau 1950, Chap. 19).

correct.[133] Suppose, for instance, that somebody shots a ball with a billiard cue and this gives rise to several collisions among the balls. If we look at these shock events themselves, we could indeed see the 'movie' both ways. However, by reversing the ordinary sequence, we would reach a point (the motion of the last ball in the sequence in the real world) in which something amazing happens: this motion is caused by nothing (in some cases, we would even see the ball coming out of a hole)! In fact, the initial cause of the motion is determined by ourselves playing billiard, and the initial event represented by a shot stands at one end (actually, at the start) of the chain, and not the other way around. Also, this shot would appear bizarre if seen in a reversed way: we would observe a number of balls converging on a single one able to push the player's stick. In other words, reading the sequence in reversed way is possible only for some of its segments, that is, it is based on *extrapolation*, but if we ask about the further causes of certain events, sooner or later we reach a baffling situation of the kind depicted: in this case, the whole Nature "appeared as delivered to chance".[134] In other words, in any causal chain also *localised events* are involved which determine local irreversibility, and these classical events need to be taken as rooted in quantum events happening randomly. Thus, not only laws but both physical contexts and events, which represent boundary and initial conditions, are inherently relevant to *any* causal process, and we should consider the classical approach to causality as a simplification. This also explains that the classical deterministic model of causality works only when we abstract the causal process without consideration of those conditions that introduce probability in every physical process. Of course, that abstraction is viable and helpful in very elementary situations but to take it as a piece of reality is both physically and epistemologically wrong. Similar considerations are true for complex and chaotic systems. Here, small fluctuations in the initial conditions may have major effects with the time.[135] And, when we try to increase the resolution or the analysis of the initial conditions, we often incur in situations that are not soluble with finite means.

It is still true that one could assume that the chain of (classical) causal events, if fully reconstituted (I mean for the whole universe), would re-establish a full equipollency between nomological and causal explanations (the idea of the block universe is precisely this, as seen in Sect. 2.3.3). So, the argument of extrapolation could be retorted by a potential opponent. However, there are serious doubts to cast on such an argument: if the original event out of which everything happened in our universe, i.e. the Big Bang, is itself a quantum event[136] (which, as a consequence, should have happened randomly), how can we find that guess reasonable at all? Moreover, our visible universe displays a directional evolution: we observe indeed that galaxies and stars were formed after the first chemical element was produced.

The fact is that (at any point of the "chain") any "initial" cause is an event, and as such it always represents some breaking of symmetry or source of disturbance

[133] Auletta (2005).
[134] Poincaré (1907).
[135] Poincaré (1907).
[136] Tryon (1973).

(no ball moves by itself into a given direction from a rest position in equilibrium in the gravitational field), and therefore it is irreversible by its own nature.[137] The main conclusion is that while the laws of Nature ultimately rely on symmetries, causal explanations are relevant especially when there is some symmetry break or at least when there are local anomalies. To a certain extent, this was also understood by classical physics, as far as Laplace tells us that the quasi-circularity of planetary orbits (of the solar system), being very improbable, demands a causal explanation.[138] Thus, from improbability Laplace infers the necessity of a cause. We can dispute whether the notion of improbability matches that of loss of symmetry (circular orbits appear more symmetric). Certainly, it denotes an anomaly relative to what we expect to be the ordinary course of Nature as described by our laws.

2.4.3 Delayed Choice

Bohr's Version of Young's Experiment

I shall present now an experiment that raises significant problems about how to understand causality in QM but helps us in being a good test of the solution advanced here. In 1949, Bohr proposed a quantum variant of Young's double-slit experiment (Fig. 1.3, Sect. 1.1.3), which in turn served as basis for the delayed-choice experiment.[139] Bohr's variant consists in casting a boson beam through a slit in a screen I, and before it reaches the plate, it is made to pass, e.g. through two slits in another screen, II. This arrangement can be inverted. The problem is, however, whether the detection of the particles "trajectory" wipes out the interference bands typical of the quantum entities or whether the detection of a quantum wave annihilates the trajectory. With this experiment Bohr concluded, in agreement with the complementarity principle (Sect. 1.3.3), that the choice of a measuring instrument forced the photon to 'choose' between going either through the two slits or through one only. In other words, the physicist had to make the choice: either he drew the path of the particle, or he just observed the interference effects.[140] Let us call this kind of procedure a *complementarity experiment*. What is interesting is that Bohr pointed out that such a choice can be postponed and made in due course, so that we do not need to fix beforehand a particular experimental arrangement.

[137] See Shimony (1965) for interesting comments on Whitehead relative to this issue.
[138] Laplace (1796, pp. 541–544).
[139] Bohr (1949a, p. 230).
[140] See Auletta (2000, p. 131).

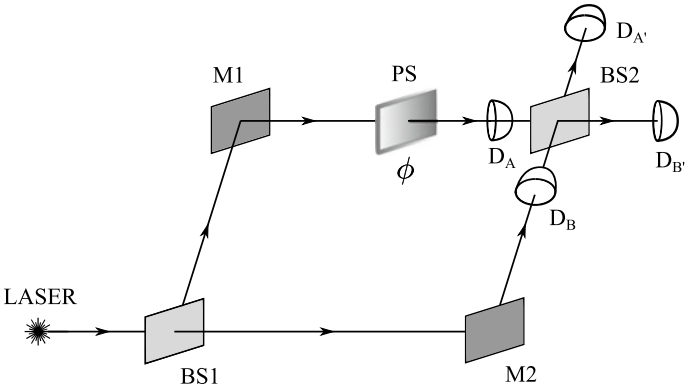


Fig. 2.15 In a delayed-choice experiment we can choose at the last moment whether to place the detectors before the beam splitter BS2 (as D_A and D_B) or after the beam splitter BS2 (as $D_{A'}$ and $D_{B'}$). Adapted from Auletta and Wang (2014, p. 101)

The Delayed-Choice Experiment

Following the suggestion of N. Bohr, the American physicist John A. Wheeler, wondered whether researchers could somehow wait until the beam of light had passed through the slits before deciding how to observe it.[141] And so he drafted what he later called the *Delayed-choice* experiment. However, for the purpose at hand, I will not describe the original two-slit formulation,[142] but rather a variant based on interferometry. The photon detectors can be moved and placed before BS2 in order to assess the path followed by an energy quantum, according to Fig. 2.15. Therefore, after the beam has actually passed BS1, it is possible to decide whether we want to observe the interference—detectors placed after the BS2—or instead to detect and establish which is the path the quantum followed—detectors placed before the BS2.[143] The question is: is there any observable difference between a delayed-choice experiment and an ordinary complementary one?

To test this issue, experiments were performed already in 1985 and 1986.[144] Initially, the experiment's layout made it impossible for researchers to check if a single photon "had taken" at BS1 the right or the left path. Consequently, the photon had taken both ways at the same time after the splitting, and later interfering with one another at the detector. The experimenters deployed a customised crystal, called pockel cell (PC), in midway, followed by a polarisation prism (Fig. 2.16). When an

[141] Wheeler (1978). See also the summaries in Auletta et al. (2009a, Sect. 9.6), Auletta and Wang (2014, Sect. 5.6).

[142] Such a description can be found in Auletta (2000, p. 448).

[143] See Auletta (2000, pp. 449–451).

[144] Hellmuth et al. (1987). Recently, a delayed–choice experiment has been performed with single atoms Manning et al. (2015). Again, we may note that very different kinds of quantum systems essentially have the same behaviour in similar physical contexts.

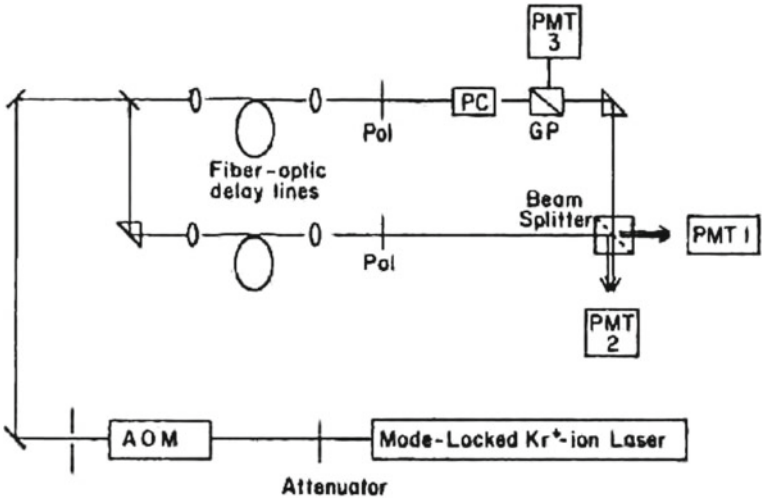


Fig. 2.16 The realised delayed-choice experiment. Photons come from the right on the bottom and are "split" at the top on the left. Transit time is approximately 24 nsec (=24 $\times 10^{-9}$ s). A Pockel cell (PC) is placed in one arm. By means of the appropriate voltage, a phase shift of one-half wave is introduced between the two orthogonal polarisation components of the incident light. The polarisation is thereby rotated of 90°. The PC is followed by a Galn polarising prism (GP) which reflects light polarised in the plane of the figure and transmits the orthogonal polarisation. The switching time of the PC is 4–5 nsec. The two beams are recombined by the second BS, a cube having BM function, before going to the cooled photomultipliers (PMT1 and PMT2). Adapted from Hellmuth et al. (1987, 419)

electrical current is applied to the PC, it can diffract the photons to an auxiliary detector (PMT3). Without the electrical current, the photons are not absorbed and go through the aforementioned cell and the prism. A random signal generator allows the cell to switch itself on and off after the photon has passed BS1 but before it reaches the detector, just as Wheeler wanted. When the cell detector is on, the photon should behave like a particle that either follows one path or the other, but never both at the same time. But if the cell detector is off, an interference ought to be shown at the final detectors (PMT1-2) indicating that the photon 'had taken' both paths at the same time. In other words, by means of an appropriate voltage we introduce a phase gap of half a wave (Fig. 1.1, Sect. 1.1.1) between the two orthogonally polarised components of the impinging light. Through a polarisation prism, the reflected light is polarised in the figure plane and then the orthogonal polarisation is transmitted. The time the PC cell takes to turn itself on is of about 4–5 nsec. If the cell is turned on five nanoseconds before the quantum of light reaches the first device, BS1, the experiment has a fixed configuration: and we refer to it as the ordinary experiment. But, if the PC cell, being off, can switch itself on just five nanoseconds after the quantum of light has passed through the first device, BS1, the experiment then does not have a fixed configuration and we refer to it as the delayed-choice experiment.

Now, comparing the data as regards both detections we find out that no experimental difference whatsoever between the results of these two kinds of experiment exists (Fig. 2.17). I stress the particularity of this result: an experiment at proving that in fact there is no difference, whereas we expect from most experiments that they show precisely a difference. In this way, we realise that the delayed-choice experiment does not alter the quantum-mechanical predictions. Put otherwise, the experiment's sophistication and ingenuity come to prove that it is not possible to know what has actually happened in the interferometer before the results of the detection, since these are identical in both experiments.

Retro-Causation?

The experiment was meant to represent evidence for Bohr's standpoint about the primacy of observation events or of observed events (as well as complementarity). However, the delayed-choice experiment is much more. This is evident when applied to astrophysical problems, as pointed out by Wheeler. Again, an observer on the Earth has the possibility to choose between observing (i) an interference phenomenon and (ii) a simple trajectory of light. In the model, the latter originates from an extremely bright quasar that came across a massive stellar object, acting as a gravitational lens, either aside or as split by the latter and merging into the direction of the observer on Earth. This means that at a later time the observer seems to determine the behaviour of something that we expect to have actually happened many millions (or even billions)

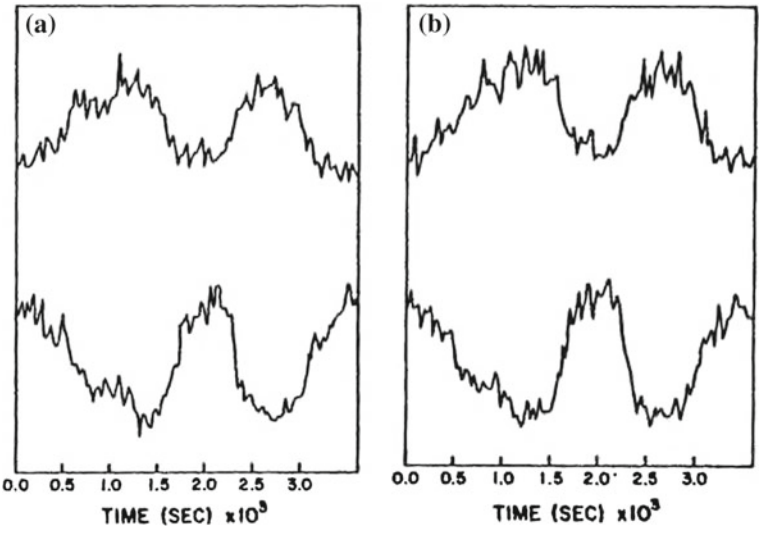


Fig. 2.17 Results of **a** the 'ordinary' experiment and of **b** the delayed-choice one. On the *y* axis the number of photodetections. Adapted from Hellmuth et al. (1987, 421)

of years before. In other words, how we decide to observe the quasar photons in the present would determine which of the two possible behaviours had the photons as they passed near the galaxy in a very remote time. How can this be brought in harmony with SR (Sect. 2.3.2)? Should such a thought experiment imply that there is a kind of retro-causation? This possibility has been raised by some scholars in the context of relativity or classical causal chains.[145] I do not wish to enter now a difficult debate on this problem. What is clear is that this possibility was meant by the proponents to be rather an exception and not the ordinary way in which causal processes happen in our world.

At the opposite, if we admit retro-causation to be the explanation of the delayed-choice experiment, we are forced to assume that the whole of the quantum-mechanical world (i.e. the basic structure of our universe) is ruled by retro-causation. In fact, although specific, the delayed-choice experiment shows a general aspect of quantum theory, i.e. the fact that anytime that we detect we seem to retro-act on some physical processes already happened since we seem to determine the state of the system that should already pre-exist our detection (as pointed out also by EPR: see Sect. 2.3.1). Now, it would be absurd to assume that we have both a quantum-mechanical arrow of time and classical arrow of time that go into opposite directions, or at least there is no evident reason to us for justifying this hypothesis,[146] being the question of why we phenomenologically observe the opposite direction in the macroscopic world unanswered to date. It would be much more reasonable to assume that either there is no time arrow at all or also the macroscopic phenomena conform to the time direction of the microscopic ones (or vice versa, if somebody thinks in a classical form). In the latter case, however, we would not speak of retro-causation since this would be the "ordinary" way in which we consider causal phenomena to happen. So we are left with the former option, i.e. that there is no time arrow. However, again we risk to be fooled by overseeing the difference between reversible laws of Nature and causal processes.

Time Asymmetry

From a relativistic point of view, as mentioned, for each presently occurring event there are a past and a future light cone, determining how time-ordered causal effects can happen (Sect. 2.3.2). For this reason, one introduces in relativity the notion of *temporal orientation* or time arrow.[147] But QM? Does quantum physics fulfil the requirement on causality imposed by SR? In the discussion of the delayed-choice experiment, I have implicitly assumed a time arrow or time asymmetry: in whatever way we like to interpret this experiment, it is grounded on the fact that our detections happens at a *later* time relative to what may have or may have not happened a

[145] First by the Austrian mathematician Gödel (1949) and then more recently in philosophy: see Lewis (1976). See also Maudlin (1994, Chaps. 4–5).

[146] A similar point has been discussed in Kistler (2006, pp. 34–35).

[147] Malament (2012, Sect. 2.2). See also Malament (1977).

long time ago. Thus, also QM apparently displays a kind of irreversibility when events happen. Is this a specificity of the delayed-choice experiment? By now, I stress that there is further evidence that choices made at whatever time can only influence future behaviour of quantum systems. In fact, in a study on this point, the conclusion of the authors was: "According to our model, if you could travel into the past quantum-mechanically, you would only see those alternatives consistent with the world you left behind you. In other words, while you are aware of the past, you cannot change it. No matter how unlikely the events are that could have led to your present circumstances, once they have actually occurred, they cannot be changed."[148] As a matter of fact, detection events represent 'choices' out of a multitude of possibilities. We shall discover, that such a sequence cannot be altered (we cannot have those choices unmade). Thus, on the basis of the previous examination we need to assume the causal (temporally asymmetric) structure of SR as valid also for QM and hope that this can help us to solve the problem raised by the delayed-choice experiment. However, the big question is: which are the changes in the notion of causation that we need to introduce to make compatible a time arrow with the reversible laws of QM?

2.5 Summary of the Chapter

In this chapter, the reader was introduced to the main problems of QM and to some basic requirements that the quantum theory needs to satisfy:

- There is a fundamental necessity for every physical theory to make ontological ascriptions.
- Probability amplitudes show a thread among a system's components that is real but apparently not physically real in an ordinary sense.
- The notions of state, observable and property should denote referents of the theory. Moreover, in opposition to CM, they are different notions in a quantum context.
- The basic distinction between detection and measurement has been introduced, where detection seems to be an uncontrollable operation, and shown that uncertainty relations forbid cumulating information. Moreover, uncertainty relations depend on the presence of interference terms. Having two different kinds of dynamics (linear and unitary on the one hand, nonlinear and irreversible on the other) appears a weakness and represent a difficulty for quantum theory.
- According to von Neumann, the only non-physical factor in measurement is the observer's mind, although he preferred psychophysical parallelism.
- A relevant consequence of the general issue of quantum measurement is that knowledge cannot be understood as a pure receptive activity detached from the observed objects.
- The thought experiment of the Schrödinger's cat has been presented.

[148] Greenberger and Svozil (2005).

- The EPR paper has been summarised. The distinction between locality and separability has also been introduced. Einstein supported the latter assumption because quantum interdependencies put in danger the notion of physical system. The idea of a block universe has been presented.
- QM does not satisfy the classical principle of sufficient reason. For this reason, the distinction between laws (that rule probabilities) and causal connections has been made. Moreover, the necessity to broaden the notion of causality has been stressed.
- The delayed-choice experiment has been summarised. This raises the question of why there seems to be a time arrow and how to cope with the problem of causation.

References

Anscombe, G. Elizabeth M. 1971. *Causality and Determination*. Cambridge: Cambridge University Press; rep. in [Sosa/Tooley 1993, 88–104].

Armstrong, David M. 1983. *What is a Law of Nature*. Cambridge: Cambridge University Press.

Auletta, Gennaro. 1992. Sémiotique et problématique corps-esprit chez Leibniz et les Occasionnalistes. *Histoire Epistémologie Language* 14 (2): 85–106.

Auletta, Gennaro. 2000. *Foundations and Interpretation of Quantum Mechanics: In the Light of a Critical-Historical Analysis of the Problems and of a Synthesis of the Results*. Singapore: World Scientific; rev. ed. 2001.

Auletta, Gennaro. 2004. Critical Examination of the Conceptual Foundations of Classical Mechanics in the Light of Quantum Physics. *Epistemologia* 27: 55–82.

Auletta, Gennaro. 2005. Quantum Information and Inferential Reasoning. *Foundations of Physics* 35: 155–169.

Auletta, Gennaro. 2006. The Ontology Suggested by Quantum Mechanics. In *Topics on General and Formal Ontology*, ed. Valore, Paolo, 161–179. Monza: Polimetrica International Scientific Publisher.

Auletta, Gennaro. 2011a. (in collaboration with I. Colagè, P. D'ambrosio, and L. Torcal), *Integrated Cognitive Strategies in a Changing World*. Rome: G and B Press.

Auletta, Gennaro. 2011b. *Cognitive Biology: Dealing With Information from Bacteria to Minds*. Oxford: Oxford University Press.

Auletta, Gennaro. 2011c. Correlations and Hyper-Correlations. *Journal of Modern Physics* 2: 958–961. https://doi.org/10.4236/jmp.2011.29114, http://www.scirp.org/journal/PaperInformation.aspx?paperID=7135.

Auletta, Gennaro. 2016. Critical Examination of Peirce's Theory of Categories. *SCIO: Journal of Philosophy* 12: 23–49.

Auletta, G., and S. Pons Domenech (eds.). 2013. *Si può parlare oggi di una finalità dell'evoluzione? Riflessioni filosofiche e teologiche alla luce della scienza contemporanea*. Rome: G and B Press.

Auletta, G., and S.-Y. Wang. 2014. *Quantum Mechanics for Thinkers*. Singapore: Pan Stanford Pub. https://doi.org/10.4032/9789814411721.

Auletta, G., M. Fortunato, and G. Parisi. 2009a. *Quantum Mechanics*. Cambridge: Cambridge University Press; rev. paperback ed. 2013.

Auletta, G., M. Alai, and G. Tarozzi. 2009b. Einstein's Local Realism and the Realistic Interpretation of the Wave Function C. ALUNNI – M. CASTELLANA – D. RIA– A. ROSSI eds. *A. Einstein et H. Weyl (1955–2005). Questions épistémologiques ouvertes* Paris–Maglie 33–49.

Bell, John S. 1981. Quantum Mechanics for Cosmologists. In *Quantum Gravity 2*, ed. Isham, C.J., R. Penrose, and D. Sciama. Oxford: Clarendon; rep. in [Bell 1987, 117–138].
Bell, John S. 1987. *Speakable and Unspeakable in Quantum Mechanics*. Cambridge: Cambridge University Press.
Bird, Alexander. 2007. *Nature's Metaphysics: Laws and Properties*. Oxford: Oxford University Press.
Bohm, David. 1951. *Quantum Theory*. New York: Prentice-Hall.
Bokulich, Alisa. 2008. *Reexamining the Quantum-Classical Relation: Beyond Reductionism and Pluralism*. Cambridge: Cambridge University Press.
Born, Max. 1926. Quantenmechanik der Stoßvorgänge. *Zeitschrift für Physik* 38: 803–827.
Born, Max. 1949a. Discussion With Einstein on Epistemological Problems in Atomic Physics, in [Schilpp 1949, 201–241].
Born, Max. 1949b. *Natural Philosophy of Chance and Cause*. New York, Oxford: Clarendon.
Bohr, Niels. 1928. The Quantum Postulate and the Recent Development of Atomic Theory. *Nature* 121: 580–590; rep. in [Bohr CW, VI, 148–158].
Bohr, Niels. 1985–1996. *Collected Works*. Amsterdam, North Holland.
Boscovich, Rogerius J. 1754. *De continuitatis lege and ejus consectariis, pertinentibus ad prima materiae elementa eorumque vires*. Roma: Collegio Romano.
Boscovich, Rogerius J. 1763. *Theoria philosophiæ naturalis redacta ad unicam legem virium in natura existentium* Venezia; Vienna.
Carroll, John W. 1994. *Laws of Nature*. Cambridge: Cambridge University Press.
Cassirer, Ernst. 1910. *Substanzbegriff und Funktionsbegriff*, Berlin. Darmstadt: Wissenschaftliche Buchgesellschaft, 1994.
Chaitin, G.J., N. da Costa, and F.A. Doria. 2011. *GödelÕs Way: Exploits into an Undecidable World*. Milton Park: Taylor and Francis.
Chappell, Vere (ed.). 1994. *The Cambridge Companion to Locke*. Cambridge: Cambridge University Press.
Clauser, J.F., M.A. Horne, A. Shimony, and R.A. Holt. 1969. Proposed Experiment to Test Local Hidden-Variable Theories. *Physical Review Letters* 23: 880–884; rep. in [Wheeler/Zurek 1983, 409–413].
Conway, J., and S. Kochen. 2006. The Free Will Theorem. *Foundations of Physics* 36: 1441–1473. https://doi.org/10.1007/s10701-006-9068-6.
Conway, J., and S. Kochen. 2009. The Strong Free Will Theorem. *Notices of the AMS* 56 (2): 226–232.
D'Ariano, G.M., and H.P. Yuen. 1996. Impossibility of Measuring the Wave Function of a Single Quantum System. *Physical Review Letters* 76: 2832–2835.
D'Ariano, M., G. Chiribella, and P. Perinotti. 2017. *Quantum Theory from First Principles: An informational approach*. Cambridge: Cambridge University Press.
Darwin, Charles R. 1859. *The Origin of Species*, London, Murray. 6th ed., 2009. Cambridge: Cambridge University Press.
de La Forge, Louis. 1666. *Traité de l'Esprit de l'Homme, de ses faculté ou fonctions et de son union avec le corps. Suivant les principes de René Descartes*. Amsterdam: A. Wolfgang; Hildesheim: Olms.
Descartes, René. 1641a. *Meditationes de prima philosophia*, Paris; in: [Descartes Œ, VII].
Descartes, René. 1641b. Meditationes de prima philosophia. In *Œuvres*, ed. C. Adam. 1907–1913 and P. Tannery. 1964–1974. Paris: Vrin.
Deutsch, David. 1997. *The Fabric of Reality: The Science of Parallel Universes and its Implications*. London: Penguin Books.
Deutsch, David. 2011. *The Beginning of Infinity: Explanations that Transform the World*. London: Penguin Books.
Dieks, Dennis. 1982. Communication by EPR Devices. *Physics Letters* 92A: 271–272.
Earman, John. 1986. *A Primer on Determinism*. Dordrecht: Reidel.

Earman, John. 1989. *World Enough and Space-Time: Absolute vs. Relational Theories of Space and Time*. Cambridge, MA: The MIT Press.
Eddington, Arthur Stanley. 1934. On the Method of Theoretical Physics, *Philosophy of Science* 1: 163–169; eng. tr. of [Eddington (1930)].
Eddington. 1939. *Philosophy of Physical Science*. Cambridge: Cambridge University Press.
Einstein, Albert. 1905. Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt. *Annalen der Physik* 17: 132–148.
Einstein, Albert. 1909. Über die Entwicklung unserer Anschauungen über das Wesen und die Konstitution der Strahlung. *Physikalische Zeitschrift* 10: 817–825.
Einstein, Albert. 1917. *Über die spezielle und die allgemeine Relativitätstheorie*. Braunschweig: Vieweg.
Einstein, Albert. 1922. Vier Vorlesungen über Relativitätstheorie. Braunschweig: Viewer; 2nd ed. with the title *Grundzüge der Relativitätstheorie*, 1956.
Einstein, Albert. 1930. Zur Methodik der Theoretischen Physik, rep. in [Einstein 1993, 113–119].
Einstein, Albert. 1948. Quantenmechanik und Wirklichkeit. *Dialectica* 2: 320–324.
Einstein, Albert. 1993. *Mein Weltbild*. Frankfurt: Ullstein.
Einstein, Albert. 2006. *Relativity: The Special and The General Theory*, Penguin, eng. tr. of [Einstein 1917]
Einstein, A., B. Podolsky, and N. Rosen. 1935. Can Quantum-Mechanical Description of Physical Reality be Considered Complete? *Physical Review* 47: 777–780; rep. in [Wheeler/Zurek 1983, 138–141].
Elitzur, A.C., and L. Vaidman. 1993. Quantum Mechanical Interaction-Free Measurements. *Foundations of Physics* 23: 987–997.
Elitzur, A.C., S. Dolev, and N. Kolenda (eds.). 2005. *Quo Vadis Quantum Mechanics?*. Berlin: Springer.
Ferrero, M., and J.L. Sánchez-Gómez. 2013. Heuristic Classification of Physical Theories based on Quantum Correlations. *EuroPhysics Journal Plus* 128: 65. https://doi.org/10.1140/epjp/i2013-13065-5.
Feynman, R.P., and A.R. Hibbs. 1965. *Quantum Mechanics and Path Integrals*. New York: McGraw-Hill; emended edition by Daniel F. New York: Dover: Styer, 2005.
French, A.P. 1968. *Special relativity*. New York: W.W. Norton and Company.
Friedman, Michael. 1983. *Foundations of Space–Time Theories. Relativistic Physics and Philosophy of Science*. Princeton: Princeton University Press.
Galilei, Galileo. 1632. *Dialogo sui Massimi sistemi*. In ed. L. Sosio. 1970. Torino: Einaudi.
Gerlach, W., and O. Stern. 1968. Sämtliche Schriften. In ed. J.P.N. Land. 1891–1893. Frommann: Den Haag; rep. Stuttgart-Bad Cannstatt.
Geroch, Robert. 1978. *General Relativity from A to B*. Chicago: University of Chicago Press.
Geulincx, Arnold. 1669. *Metaphysica vera*; in [Geulincx SS, II, 137–198].
Gnedenko, Boris V. 1969. *The Theory of Probability*. Moscow: Mir.
Gödel, Kurt. 1949. A Remark About the Relationship Between Relativity Theory and Idealistic Philosophy, in [Schilpp 1949, 557–562].
Greenberger, D.M., and K. Svozil. 2005. Quantum Theory Looks at Time Travel, in [Elitzur et al. 2005, 63–71].
Heisenberg, Werner. 1927. Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. *Zeitschrift für Physik* 43: 172–198; rep. in [Heisenberg GW, A.I, 478–504].
Heisenberg, Werner. 1984. *Gesammelte Werke*. München/Berlin: Piper/Springer.
Hellmuth, T., H. Walther, A.G. Zajonc, and W.P. Schleich. 1987. Delayed-Choice Experiments in Quantum Interference. *Physical Review* A35: 2532–2541.
Home, D., and M.A.B. Whitaker. 2007. *Einstein's Struggles with Quantum Theory: A Reappraisal*. Berlin: Springer.
Howard, Don. 1985. Einstein on Locality and Separability. *Studies in History and Philosophy of Science* 16: 171–201.

Hume, David. 1777. *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*, 2nd ed. 1902; 3rd ed. 1975. London, Oxford: Clarendon.
Jammer, Max. 1974. *The Philosophy of Quantum Mechanics. The Interpretation of Quantum Mechanics in Historical Perspective*. New York: Wiley.
Jammer, Max. 1999. *Einstein and Religion*. Princeton: Princeton University Press.
Kant, Immanuel. 1747. *Gedanken von der wahren Schätzung der lebendigen Kräfte ...*; in [Kant A, I, 1–182].
Kant, Immanuel. 1756. *Metaphysicæcum geometria iunctæ usus in philosophia naturali, cuius Specimen I continet Monadologiam Physicam*; in [Kant A, I, 1–182].
Kant, Immanuel. 1968. *Kants Werke. Akademie Textausgabe*. Berlin: W. de Gruyter.
Kistler, Max. 2006. *Causation and Laws of Nature*. Abingdon: Routledge.
Kolmogorov, Andrei N. 1933. *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Berlin: Julius Springer.
Koyré, Alexandre. 1957. *From the Closed World to the Infinite Universe*. Baltimore: John Hopkins Press.
Koyré. 1966. *Etudes galiléennes*. Paris: Hermann.
Landau, Lev D., and E.M. Liftshitz. 1976. *Mechanics* (eng. tr.), I vol. of *The Course of Theoretical Physics*. Oxford: Pergamon.
Landsman, Klaas. 2017. *Foundations of Quantum Theory: From Classical Concepts to Operator Algebras*. Berlin: Springer.
Laplace, Pierre–Simon de. 1796. *Exposition du système du monde*. Paris: Fayard.
Laplace, Pierre–Simon de. 1825. *Essai philosophique sur les probabilités*, 5th ed. 2009. Cambridge: Cambridge University Press.
Leibniz, Gottfried W. 1666. *Philosophische Schriften*. In ed., Gerhardt. 1875. Halle; rep. Hildesheim: Olms 1978.
Leibniz, Gottfried W. 1677. Specimen inventorum de admirandis naturæ generaliz arcanis [Leibniz PS, VII, 309–318].
Leibniz, Gottfried. 1686. Brevis Demonstratio Erroris Mirabilis Cartesii et aliorum .... *Acta Eruditorum Lipsiensium*, in [Leibniz MS, VI, 117–123].
Leibniz, Gottfried W. 1695a. Systeme nouveau de la nature et de la communication des substances, aussi bien que de l'union qu'il y a entre l'ame et le corps. *Journal des Savants*; [Leibniz PS, IV, 477–487].
Leibniz, Gottfried W. 1695b. Specimen Dynamicum pro admirandis NaturæLegibus circa corporum vires et mutuas actiones detegendis et ad suas causas revocandis. Pars I. *Acta Eruditorum Lipsiensium*; rep. in [Leibniz MS, VI, 234–246].
Leibniz, Gottfried W. 1702. Extrait du Dictionnaire de M. Bayle article Rorarius p. 2599 sqq. de l'Edition de l'an 1702 avec mes remarques, in [Leibniz PS, IV, 524–554].
Leibniz, Gottfried W. 1710. *Essais de Théodicée*; in [Leibniz PS, VI, 21–436].
Leibniz, Gottfried W. 1710–1712. Principes de la Nature et de la Grace, in [Leibniz PS, VI, 598–606].
Leibniz, Gottfried W. 1712–1714. Monadologie, in [Leibniz PS, VI, 607–623].
Leibniz, Gottfried W. 1715–1716. Streitschriften zwischen Leibniz und Clarke, in [Leibniz PS, VII, 345–440]. Published before as *A Collection of Papers, Which Passed Between the Late Learned Mr. Leibniz and Dr. Clarke in the Years 1715 and 1716, Relating to the Principles of Natural Philosophy and Religion*, London, 1717.
Leibniz, Gottfried W. 1875. *Philosophische Schriften*, In ed. Gerhardt. Halle; rep. Olms: Hildesheim, 1978.
Lewis, David. 1976. The Paradoxes of Time Travel. *American Philosophical Quarterly* 13: 145–152; rep. in [Lewis PP, II, 67–80].

Lewis, David. 1983. *Philosophical Papers*. Oxford: Oxford University Press.
Lindblad. 1999. A General No-Cloning Theorem. *Letters in Mathematical Physics* 47: 189–196. https://doi.org/10.1023/a:1007581027660.
Locke, John. 1689. *An Essay Concerning Human Understanding*, 2nd ed. 1694; 4th ed. 1700. Oxford: Oxford University Press.
Lorentz, Hendrik A. 1904. Electromagnetic Phenomena in a System Moving with any Velocity Smaller than that of Light. *Proceedings of the Royal Netherlands Academy of Arts and Sciences* 6: 809–831.
Ludwig, Güunther. 1954. *Die Grundlagen der Quantenmechanik*. Berlin: Springer.
Ludwig, Günther. 1978. *Die Grundstrukturen einer physikalischen Theorie*, 2nd ed. 1990. Berlin: Springer.
Ludwig, Günther. 1983. *Foundations of Quantum Mechanics*. Berlin: Springer; eng. tr. of [Ludwig 1954].
Mach, Ernst. 1883. *Die Mechanik in Ihrer Entwicklung. Historisch-kritisch dargestellt*, 9th ed. 1933. Leipzig, Darmstadt: WBG.
Mach, Ernst. 1905. *Erkenntnis und Irrtum*, 2nd ed. 1906; 5th ed. 1926. Leipzig, Darmstadt: WBG.
Malament, David B. 1977. The Class of Continuous Timelike Curves Determines the Topology of Spacetime. *Journal of Mathematical Physics* 18: 1399–1404. https://doi.org/10.1063/1.523436.
Malament, David B. 1979. Œuvres. In ed. G. Rodis–Lewis. Paris: Pleiade.
Malament, David B. 2012. *Topics in the Foundations of General Relativity and Newtonian Gravitation Theory*. Chicago: Chicago University Press.
Malebranche, Nicolas. 1674–1678. *De la recherche de la verité*, 6th ed. 1712. Paris; in [Malebranche Œ, I, 1–1126].
Manning, A.G., R.I. Khakimov, R.G. Dall, and A.G. Truscott. 2015. Wheeler's Delayed-Choice Gedanken Experiment with a Single Atom. *Nature Physics* 11: 539–542. https://doi.org/10.1038/nphys3343.
Margenau, Henry. 1950. *The Nature of Physical Reality: A Philosophy of Modern Physics*. New York: McGraw-Hill; Woodbridge, CT.: Ox Bow Press. 1977.
Maudlin, Tim. 1994. *Quantum Non-Locality and Relativity: Metaphysical Intimations of Modern Physics*, 2nd ed. 2002. Oxford: Blackwell.
Marlow, A.R. (ed.). 1978. *Mathematical Foundations of Quantum Theory*. New York: Academic.
McCann, Edwin. 1994. Locke's Philosophy of Body, in [Chappel 1994, 56–88].
Mehra, J., and H. Rechenberg. 1982–2001. *The Historical Development of Quantum Theory*. New York: Springer.
Minkowski, Hermann. 1907–1908. Die Grundgleichungen fũr die elektromagnetischen Vorgänge in bewegten Körpern. *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse* 1907–1908: 53–111.
Misner, C.W., K.S. Thorne, and J.A. Wheeler. 1970. *Gravitation*. San Francisco: Freeman and Co.
Newton, Isaac. 1687. Philosophiæ Naturalis Principia Mathematica, London; Cambridge, 1713; London, 1726; Harvard: Harvard University Press. 1972. (ed. Cohen): A. Koyré and I. B.
Norton, John D. 2003. Causation as Folk Science. *Philosophers' Imprint* 3 (4): 1–22.
Peirce, Charles S. 1886. An Elementary Account of the Logic of Relatives, in [Peirce W, V, 379–387].
Peirce, Charles S. 1903a. Sundry Logical Conceptions, in [Peirce EP, II, 267–288].
Peirce, Charles S. 1903b. The Categories Defended, in [Peirce EP, II, 160–178].
Peirce, Charles S. 1903c. Thirdness — The Reality of Thirdness, in [Peirce CP, 1.343–349].
Peirce, C.P. 1958. *The Collected Papers*, vol. I–VI (ed. Charles Hartshorne and Paul Weiss). Cambridge: Harvard University Press. 1931–1935; vol. VII–VIII (ed. Arthur W. Burks), Cambridge: Harvard University Press.
Peirce, Charles S. 1982. *Writings*, vol. I. Bloomington: Indiana University Press.
Peirce, Charles S. 1998. *The Essential Peirce*, vols. I–II. Bloomington: Indiana University Press.
Penrose, Roger. 2004. *The Road to Reality: A Complete Guide to the Laws of the Universe*. J. Cape, London: Vintage.

Poincaré, Henri. 1907. Le hasard. *Revue du mois* 3: 257–276; rep. in [Poincaré 1991,135–156].
Poincaré, Henri. 1911. L'évolution des lois. *Scientia. Rivista di scienza* 9 (4): 275–292; rep. in [?, 48–67].
Poincaré, Henri. 1991. *L'analyse et la recherche*. Paris: Hermann.
Popescu, S., and D. Rohrlich. 1994. Quantum Nonlocality as an Axiom. *Foundations of Physics* 24: 379–385.
Popper, Karl R. 1934. *Logik der Forschung*. Wien: Springer; 8th ed. 1984. Tübingen: Mohr.
Popper, Karl R. 1974. *Unended Quest: An Intellectual Autobiography*. London: Routledge.
Renninger, Mauritius K. 1960. Messungen ohne Störung des Meßobjekts. *Zeitschrift für Physik* 158: 417–421.
Rindler, Wolfgang. 2001. *Relativity: Special, General, and Cosmological*. Oxford: Oxford University Press.
Ruelle, David. 1990. The Claude Bernard Lecture, 1989. Deterministic Chaos: The Science and the Fiction. *Proceedings of the Royal Society of London* A427: 241–248.
Russell, Bertrand. 1912–1913. On the Notion of Cause. *Proceedings of the Aristotelian Society* 13: 1–26.
Schilpp, Arthur (ed.). 1949. *Albert Einstein. Philosopher-Scientist*. La Salle: Open Court. 3rd ed., 1988.
Schrödinger, Erwin. 1935. Die gegenwärtige Situation in der Quantenmechanick. I–III. *Naturwissenschaften* 23: 807–812, 823–828, 844–849; rep. in [Schrödinger GA, III, 484–501].
Schrödinger, Erwin. 1936. Probability Relations Between Separated Systems. *Proceedings of the Cambridge Philosophical Society* 31: 555–563; rep. in [Schrödinger GA, III, 143–165].
Schrödinger, Erwin. 1984. *Gesammelte Abhandlungen*. Wien: Verlag der österreichischen Akademie der Wissenschaften.
Schwartz, Matthew D. 2014. *Quantum Field Theory and the Standard Model*. Cambridge: Cambridge University Press.
Shimony, Abner. 1965. Quantum Physics and the Philosophy of Whitehead. In *Philosophy in America*, ed. Max Black, London: Allen and Unwin; rep. in [Shimony 1993, I, 183–273].
Shimony, Abner. 1993. *Search for a Naturalistic Point of View*. Cambridge: Cambridge University Press.
Sosa, E., and M. Tooley (eds.). 1993. *Causation*. Oxford: Oxford University Press.
Spinoza, Baruch. 1677. *Ethica*, Amsterdam; in [Spinoza O, II, 41–308].
Spinoza, Baruch. 1925. *Opera*. In ed. C. Gebhardt, vol. 4. Heidelberg, rep. Heidelberg: Carl Winter 1972.
Tarozzi, Gino. 1988. Science, Metaphysics and Meaningful Philosophical Principles. *Epistemologia* 1197–1104.
Tarozzi, Gino. 1996. Quantum Measurements and Macroscopical Reality: Epistemological Implications of a Proposed Paradox. *Foundations of Physics* 26: 907–917.
'T Hooft, Gerard. 2016. *The Cellular Automaton Interpretation of Quantum Mechanics*. Berlin: Springer.
Tryon, Edward P. 1973. Is the Universe a Vacuum Fluctuation? *Nature* 246: 396–397.
van Fraassen, Bas C. 1989. *Laws and Symmetry*. Oxford: Clarendon.
von Neumann, John. 1932. *Mathematische Grundlagen der Quantenmechanik*. Berlin: Springer.
Wallace, David. 2012. *The Emergent Multiverse: Quantum Theory according to the Everett Interpretation*. Oxford: Oxford University Press.
Weyl, Hermann. 1949. *Philosophy of Mathematics and Natural Science*. Princeton: Princeton University Press.
Wheeler, John A. 1978. The 'Past' and the 'Delayed-Choice' Double-Slit Experiment, in [Marlow 1978, 9–48].
Wheeler, J.A., and W. Zurek (eds.). 1983. *Quantum Theory and Measurement*. Princeton: Princeton University Press.
Wootters, W.K., and W.H. Zurek. 1982. A Single Quantum Cannot Be Cloned. *Nature* 299: 802–803.

Zeilinger, Anton. 1999. A Foundational Principle for Quantum Mechanics. *Foundations of Physics* 29: 631–643.
Zeilinger, Anton. 2000. Quantum Entangled Bits Step Closer to It. *Science* 289: 405–406.

Chapter 3
The Main Interpretations



I have grouped the main problems in interpreting QM in the previous chapter. In the present chapter, we shall deal again with these four groups of problems (formalism, measurement, non-locality and causality) but by introducing the reader to the main solutions that have given to them and critically evaluating them. In fact, only a careful examination of what has been said on the subject could help us to find an original way to see the problems. It helps us to restrict the range of the viable pieces of interpretation by excluding those hypotheses that, for one reason or the other, cannot work and thus to attribute the right values to those hypotheses that resist critical analysis. I stress that no interpretation or hypothesis can be dismissed in few words for at least two good reasons: (i) QM looks like a complex array of puzzles, what requires extreme care in making any assertion, and (ii) all of these interpretations have been provided by excellent scholars who have tried to deeply penetrate that riddle.

Note that for the sake of argumentation I invert the order of the exposition starting with the measurement problem: it is the source of the main (and, for the theory, most dangerous) difficulties and in fact it also affects the other problems. After an analysis of the subjectivist and objectivist interpretations, the originality of the decoherence solution will be shown, and then going to the issue of the significance of the formalism, where we shall discover the relevance of the notion of information. The EPR argument, re-presented in logical form, is the starting point of the analysis of hidden-variable theories aiming at building a full classical deterministic theory overcoming QM. This is a crucial test for the theory. After presenting Bell's historical accomplishment, several important developments on non-locality and information are presented (entanglement swapping, teleportation, quantum cryptography) for arriving at a definition of the notion of non-separability. The final section will essentially deal with the notion of causal constraints, enlarging our understanding of causality in a way that is consistent with QM, as well as the notion of potentiality, initially introduced in this theory by W. Heisenberg.

© Springer Nature Switzerland AG 2019

G. Auletta, *The Quantum Mechanics Conundrum*,
https://doi.org/10.1007/978-3-030-16649-6_3

3.1 Solutions to the Measurement Problem

3.1.1 Subjectivist Interpretation

Reasons for this Standpoint

We need to understand the reason for the subjectivist interpretation: recalling our presentation of the measurement problem in Sect. 2.2 one could be tempted to say that it is only our lack of knowledge (about the state of the object system) that makes such a process so special. Therefore, it is the transition from lack of knowledge to knowledge that should represent the solution to the problem.

The subjectivist interpretation was supported by some prominent scholars.[1] After von Neumann's seminal work, it is necessary to take into account the joint contribution of the French physicist Edmond Bauer (1880–1963) and the German physical chemist Fritz London (1900–1954), representing the first detailed analysis of the problem.[2] It was clear to these two scholars that, by coupling an object system and an apparatus (in the measurement step that is called *premeasurement*, since it comes before detection), we obtain an extraordinary consequence that we have already noted when discussing the Schrödinger's cat in Sect. 2.2.5: if the object system is assumed to be in some superposition state relative to the observable that we wish to measure (what is the ordinary situation), such a superposition will also 'affect' the measurement apparatus itself producing an entanglement between object system and apparatus (Sect. 1.4.2).

Two Different Kinds of Cross-Terms

In this chapter we shall focus on such a coupling and to this purpose come back to Eq. (2.43). I attract the attention of the reader on the fact that the "cross" terms

$$|d\rangle\,|a_u\rangle \text{ and } |u\rangle\,|a_d\rangle \tag{3.1}$$

are absent in this equation. These terms are not to mix with interference terms (Sect. 2.1.2). At the opposite, their addition to the two terms on the right side of expression (2.43) would give rise to a classical state for two systems that no longer interact and are separated (see Eqs. (1.386) and (1.388)). In fact, the presence of all possible combinations (four as whole) of the components of the state of the apparatus and the components of the state of the object system would tell us that we can get any of the four possible joint outputs randomly: it is a product state. Thus, the above terms express rather the *failure* of the apparatus to correctly read the result, and, for this reason, are associated to equivocation or noise. In such a context, I shall not take

[1] For a quick summary see Auletta (2004b).

[2] London and Bauer (1939).

these problems into consideration. At the opposite, the terms

$$|d\rangle\,|a_d\rangle \text{ and } |u\rangle\,|a_u\rangle \tag{3.2}$$

are present and represent the correlation between object and apparatus that is wanted. It is on these terms that we need now to focus our analysis. By now, keep in mind that entanglement represents a *reduction* of the possible (joint) outcomes (here, from four to two). Note finally that coupling, i.e. transformations (2.43) and (2.44), is unitary, and the change of the states of the object system $\mathcal{S}$ and the apparatus $\mathcal{A}$, induced by coupling or premeasurement, takes the general form

$$\hat{U}_t^{\mathcal{SA}} = e^{-\frac{i}{\hbar}t\hat{H}_{\mathcal{SA}}}, \tag{3.3}$$

in accordance with Eq. (1.16) and where $\hat{H}_{\mathcal{SA}}$ is the interaction Hamiltonian between these two systems (Eq. (2.21)). Note also that the total Hamiltonian of the compound system in general also comprehends terms that are independent of this interaction. In the following, we shall focus on an example of similar interaction Hamiltonians.

A Model

Let us consider only the metre $\mathcal{M}$ (that is, a microscopic system) instead of the whole apparatus–detector (equipped with amplification devices and reading devices), and assume that both the system $\mathcal{S}$ and the metre $\mathcal{M}$ are 2D systems.[3] Let us also assume that both are fermions. In particular, I choose the operator $\hat{\sigma}_z^{\mathcal{S}}$ for the observable of $\mathcal{S}$ and $\hat{\sigma}_x^{\mathcal{M}}$ for the observable of the metre (see the matrices (1.324)). The system $\mathcal{S}$ is initially prepared in a superposition of the two eigenstates of $\hat{\sigma}_z^{\mathcal{S}}$, i.e. $|\uparrow\rangle_{\mathcal{S}}, |\downarrow\rangle_{\mathcal{S}}$, given by Eq. (1.322). As previously agreed, I have dropped the subscript for the z-direction. The metre is assumed to be initially in the z spin-down state, $|\downarrow\rangle_{\mathcal{M}}$, so that if, after the interaction, the system is in $|\downarrow\rangle_{\mathcal{S}}$, the state of $\mathcal{M}$ remains unchanged. Otherwise, it will become $|\uparrow\rangle_{\mathcal{M}}$. The interaction Hamiltonian can then be explicitly written as

$$\hat{H}_{\mathcal{SM}} = \varepsilon_{\mathcal{SM}}\left(1 + \hat{\sigma}_z^{\mathcal{S}}\right)\hat{\sigma}_x^{\mathcal{M}}\,, \tag{3.4}$$

where $\varepsilon_{\mathcal{SM}}$ is some coupling function between the metre and the object system. The first step we need to make in order to calculate the action of the unitary operator

$$\hat{U}_\tau^{\mathcal{SM}} = e^{-\frac{i}{\hbar}\tau\hat{H}_{\mathcal{SM}}} \tag{3.5}$$

on the initial (uncoupled) state $|\Psi(0)\rangle_{\mathcal{SM}}$ of the joint system (where τ is the time interval of the interaction between $\mathcal{S}$ and $\mathcal{M}$), is to diagonalise the matrix $\hat{H}_{\mathcal{SM}}$. The operator $\hat{\sigma}_z^{\mathcal{S}}$ is already diagonal with respect to the basis states (1.322), according to

[3] For this example see Auletta et al. (2009, Sect. 9.1).

Eqs. (1.321). On the other hand, the two eigenvalues of $\hat{\sigma}_x^{\mathcal{M}}$ are ± 1 and the eigenkets, expressed in terms of the z spin-up and spin-down states, are given by Eqs. (1.328a). Now, in order to find the time evolution of the quantum state of the compound system, we only need to write its initial state in terms of the eigenkets (1.322) and (1.328a) of $\hat{H}_{\mathcal{SM}}$

$$\begin{aligned} |\,\Psi(\tau)\rangle_{\mathcal{SM}} &= \hat{U}_{\tau}^{\mathcal{SM}}\,|\,\Psi(0)\rangle_{\mathcal{SM}} \\ &= e^{-\frac{i}{\hbar}\tau\varepsilon_{\mathcal{SM}}\left(1+\hat{\sigma}_z^{\mathcal{S}}\right)\hat{\sigma}_x^{\mathcal{M}}}\left[\left(c_{\uparrow}\,|\uparrow\rangle_{\mathcal{S}}+c_{\downarrow}\,|\downarrow\rangle_{\mathcal{S}}\right)|\downarrow\rangle_{\mathcal{M}}\right]\,. \end{aligned} \tag{3.6}$$

Thus, we are assuming that the interaction is switched on at $t=0$ and off at $t=\tau$. Alternatively, one might envisage a situation where the interaction is smoothly turned on and off. As we know from Sect. 1.2.3, the action of the Hamiltonian onto its eigenkets simply returns the corresponding eigenvalues. Therefore, since

$$|\downarrow\rangle_{\mathcal{M}}=\frac{1}{\sqrt{2}}\left(|\uparrow_x\rangle_{\mathcal{M}}-|\downarrow_x\rangle_{\mathcal{M}}\right)\,, \tag{3.7}$$

we have

$$\begin{aligned} |\,\Psi(\tau)\rangle_{\mathcal{SM}} = \frac{1}{\sqrt{2}}\Big(&c_{\uparrow}e^{-\frac{2i}{\hbar}\tau\varepsilon_{\mathcal{SM}}}\,|\uparrow\rangle_{\mathcal{S}}\,|\uparrow_x\rangle_{\mathcal{M}}+c_{\downarrow}\,|\downarrow\rangle_{\mathcal{S}}\,|\uparrow_x\rangle_{\mathcal{M}} \\ &-c_{\uparrow}e^{+\frac{2i}{\hbar}\tau\varepsilon_{\mathcal{SM}}}\,|\uparrow\rangle_{\mathcal{S}}\,|\downarrow_x\rangle_{\mathcal{M}}-c_{\downarrow}\,|\downarrow\rangle_{\mathcal{S}}\,|\downarrow_x\rangle_{\mathcal{M}}\Big)\,, \end{aligned} \tag{3.8}$$

where the exponential $e^{-\frac{i}{\hbar}\tau\varepsilon_{\mathcal{SM}}\left(1+\hat{\sigma}_z^{\mathcal{S}}\right)\hat{\sigma}_x^{\mathcal{M}}}$ is 1 when the eigenvalue of $\hat{\sigma}_z^{\mathcal{S}}$ is -1, i.e. for the state $|\downarrow\rangle_{\mathcal{S}}$. Substituting Eqs. (1.328a) for the metre into Eq. (3.8) and using expressions (2.34), we obtain

$$\begin{aligned} |\,\Psi(\tau)\rangle_{\mathcal{SM}} = \frac{1}{\sqrt{2}}\Big(&-c_{\uparrow}\frac{2i}{\sqrt{2}}\sin\frac{2\tau\varepsilon_{\mathcal{SM}}}{\hbar}\,|\uparrow\rangle_{\mathcal{S}}\,|\uparrow\rangle_{\mathcal{M}}+\sqrt{2}c_{\downarrow}\,|\downarrow\rangle_{\mathcal{S}}\,|\downarrow\rangle_{\mathcal{M}} \\ &+c_{\uparrow}\frac{2i}{\sqrt{2}}\cos\frac{2\tau\varepsilon_{\mathcal{SM}}}{\hbar}\,|\uparrow\rangle_{\mathcal{S}}\,|\downarrow\rangle_{\mathcal{M}}\Big)\,. \end{aligned} \tag{3.9}$$

Choosing now

$$\frac{2\tau\varepsilon_{\mathcal{SM}}}{\hbar}=\frac{\pi}{2}\,,\qquad \text{or}\qquad \tau=\frac{\pi\hbar}{4\varepsilon_{\mathcal{SM}}}\,, \tag{3.10}$$

i.e. by fine-tuning the interaction time, we kill the last term and finally obtain

$$|\,\Psi(\tau)\rangle_{\mathcal{SM}}=c_{\downarrow}\,|\downarrow\rangle_{\mathcal{S}}\,|\downarrow\rangle_{\mathcal{M}}-ic_{\uparrow}\,|\uparrow\rangle_{\mathcal{S}}\,|\uparrow\rangle_{\mathcal{M}}\,, \tag{3.11}$$

which is the required coupling between the system and the metre. This model can be easily treated with the density matrix formalism.

Some Lessons

The main result is that, as a consequence of premeasurement, the metre and the object systems become entangled, as displayed in the previous equation. Now, with an entanglement of this kind there are two relevant aspects that need to be considered:

- The object system and the metre, or, more in general, the apparatus, are brought into some *correspondence*, that is, any component of the apparatus corresponds to a component of the superposition of the object system, so that we can associate a specific value of the apparatus' pointer to any measurement outcome: this is a consequence of the fact that not all pairings between components of the two systems are allowed but terms like (3.1) are not present. This is what we expect from a measuring device and it is also what we in fact experience in our laboratories.
- However, as long as the object system and the apparatus are in such an entangled state, *none of these outcomes* (and eigenvalues) can be produced. This is a restatement of von Neumann's problem (Sect. 2.2.3). Indeed, it is like a situation in which the pointer points at 'any' possible value, a situation that nobody has observed in any known circumstance (and which, to a certain extent, recalls again the Schrödinger's cat: Sect. 2.2.5). In such a situation we would even consider such a device as not working as a detection apparatus at all.

Therefore, it was clear to London and Bauer (as it was already to von Neumann) that something should happen once the coupling between object system and apparatus or the premeasurement step is accomplished in order to allow a proper detection. Now, these two scholars knew that, by somehow erasing these interference terms, we could get a particular mixture that would show precisely the required interconnection between object system and apparatus without displaying any interference or entanglement, allowing in this way the reading of a final determined value. In other words, recalling Eq. (2.48), we need to obtain a state of this kind out of the entanglement (3.11):

$$\hat{\tilde{\rho}}^{\mathcal{SM}} = \left|c_{\uparrow}\right|^2 \hat{P}^{\mathcal{S}}_{\uparrow} \hat{P}^{\mathcal{M}}_{\uparrow} + \left|c_{\downarrow}\right|^2 \hat{P}^{\mathcal{S}}_{\downarrow} \hat{P}^{\mathcal{M}}_{\downarrow}, \tag{3.12}$$

where the projectors $\hat{P}^{\mathcal{S}}_{\uparrow}$, $\hat{P}^{\mathcal{S}}_{\downarrow}$ act on state vectors in $\mathcal{H}_{\mathcal{S}}$ while the projectors $\hat{P}^{\mathcal{M}}_{\uparrow}$, $\hat{P}^{\mathcal{M}}_{\downarrow}$ on state vectors in $\mathcal{H}_{\mathcal{M}}$. Note that the quantum interference terms that have been dropped here are

$$c_{\uparrow} c^{*}_{\downarrow} \left|\uparrow\right\rangle \left\langle\downarrow\right|_{\mathcal{S}} \otimes \left|\uparrow\right\rangle \left\langle\downarrow\right|_{\mathcal{M}} \quad \text{and} \quad c^{*}_{\uparrow} c_{\downarrow} \left|\downarrow\right\rangle \left\langle\uparrow\right|_{\mathcal{S}} \otimes \left|\downarrow\right\rangle \left\langle\uparrow\right|_{\mathcal{M}}. \tag{3.13}$$

These terms, as well as the terms (3.2), represent cross-terms *inside* a single subsystem and not *between* the two subsystems, as it occurs for the classical, although weird, terms (3.1). The expression (3.12) is an instance of *Lüders mixture*, after the name of the German physicist Gerhart Lüders (1920–1995).[4] In such mixture, all the components of the two systems need to be coupled in a 'parallel' way.

[4] See Lüders (1951).

Mixture and Detection

Although representing the necessary condition for detection, a mixture like the (3.12) is still insufficient. In fact, London and Bauer pointed out that detection does not consist in a classical coupling between *each* of the component states of the apparatus and *each* of the component states of the object system, but in a *single* coupling, as witnessed by the pointer being in *a certain* specific position, while a mixture displays all possible connections (although without interference terms). In such a case, we need to make use of the famous von Neumann's formula for computing the output state $\hat{\rho}_j$ after detection of the object system:

$$\hat{\rho}_j = \frac{\hat{P}_j \hat{\rho} \hat{P}_j}{\mathrm{Tr}(\hat{\rho} \hat{P}_j)}, \tag{3.14}$$

where the denominator ensuring the normalisation of the state $\hat{\rho}_j$ is given by Eq. (1.379), while the numerator in the 2D case is given by an analogue of Eq. (2.48). If the initial state of the object system is $\hat{\rho} = |\psi\rangle \langle\psi|$, where $|\psi\rangle = c_\uparrow |\uparrow\rangle + c_\downarrow |\downarrow\rangle$, we shall get, e.g.:

$$\hat{\rho}_\uparrow = \frac{|c_\uparrow|^2 \hat{P}_\uparrow}{|c_\uparrow|^2} = \hat{P}_\uparrow. \tag{3.15}$$

The view of London and Bauer in part is justified, in part is not. In fact, a mixture is already a kind of classical state, and therefore, we have already got what appears to be the true substantial step: dropping the quantum–mechanical interference terms that hinder a true detection (and make the true difference between classical and quantum measurement). For instance, a mixture like (3.12) could represent the probabilities to get one of the faces of a coin before throwing it. As we shall see, in the transition from a pure state to a mixture, there is in fact a difference (of which the two scholars are aware) in terms of growth of entropy and loss of information (this deserves further analysis). On the other hand, detection is the only irreversible step of the whole process of which we can have experimental evidence. Detection (or the reduction of the wave packet) can indeed be understood as a kind of symmetry breaking: starting from a state in which all couplings are possible, the final detection represents a collapse on one of them (Sect. 2.2). Now, it seemed to the two scholars that only the observer can bring to this extraordinary result: when he/she, e.g. says "We perceive the apparatus as being in the state $|a_\uparrow\rangle$", he/she is justified in inferring that the object system must in fact be in the physical state $|\uparrow\rangle$. According to London and Bauer, this should not be understood in solipsistic terms either, since any other member of the scientific community could verify the same result and become aware of the same situation. Summing up, once that a mixture is produced (although the authors did not fully clarify the specific mechanism through which it should happen), it is the observer who in fact observes only one of its possible components.

Wigner's Standpoint

A fundamental development of this school is due to Eugene Wigner,[5] who, as already mentioned, gave substantial contributions to QM.[6] According to him, the attribution of physical reality independently from the subject is untenable. I recall that the opposite was precisely one of the main assumptions of classical physics, and we might consider as a progress induced by QM the acknowledgment of this (epistemologically riper) standpoint that can be called Kantian in a broad sense. From this point of view, Wigner can be considered a follower of Bohr (and of the Copenhagen interpretation: Sects. 1.3.3 and 2.4.3) who, as we shall see, after 1935, in the discussion with EPR, had introduced the idea that it is impossible to make any ontological ascription without considering the epistemic conditions in which this is feasible at all. What we can say at the most is that we can compute the probability for certain outcomes. This is why Wheeler, following Bohr, said: "No elementary phenomenon is a phenomenon until it is a registered (observed) phenomenon".[7] There is certainly some truth in such a statement, and it is not by chance that Wheeler was the proponent of the delayed-choice experiment (Sect. 2.4.3). Let us call this standpoint *quantum phenomenism* (or phenomenism in short when no ambiguity arises). Provisionally, we can say that it consists in affirming that ascription of reality needs to be made on the basis of observation–registration.

However, Wigner seems to go a step further in his analysis of the problem.[8] According to him, any quantum-mechanical description of a physical system (the wave function or the state vector) is simply a mathematical tool that connects subsequent possible detection events expressing their probability to happen (Sect. 2.4.2).[9] In itself it describes no reality. These detection events are real and actually are *observational experiences* involving apperception and memory (i.e. registration of the experience in our mind). Now, it is clear to Wigner that, by definition, an experience is an act of (not necessarily voluntary) decision: we can perceive (and keep in memory) either this or that reality but certainly not a superposition of two alternatives.[10] So, according to Wigner, a correct explanation of the measurement problem is rooted in the fact that, in our, at least conscious, experience, we always perceive a *single* component of some superposition or entangled state. In other words, an entanglement between object system and apparatus is not compatible with the nature of the mind and the observer. An interesting question is: is this state of affairs a specificity

[5] Wigner (1961). See also Stapp (1982, 1993).

[6] Apart from the Wigner's theorem (Sect. 1.2.5), I recall here the quasi-probability Wigner function (of both position and momentum), important especially in quantum optics. I also recall his crucial contributions to group theory and its applications to QM.

[7] Wheeler (1983, p. 184). It has been pointed out that from 1930s on Bohr considered the term *phenomenon* as referred to a dynamical process and not to a state (Stachel 2017).

[8] Although we can find some sparse statements of Bohr that go in this direction (Bohr 1948; Home and Whitaker 2007, p. 61).

[9] For a similar statement see Eddington (1939, p. 50).

[10] For perception this was shown already by the *Gestalt* school. For a short summary see Auletta (2011a, Sect. 4.4.4).

of QM or applies to the whole of physical reality? Wigner seemed inclined to think that QM is in fact potentially able to describe the whole of the physical reality in terms of probabilities.

So, according to Wigner, the problem of measurement is solved by simply denying that anything can exist (or that it makes sense to say that it exists) without being experienced. Note that in such a way Wigner gives a new turn to the subjectivist interpretation: instead of assuming that the mind can have an influence on the physical world or at least on our relation with the latter, he explicitly rejects the notion of a reality as independent of us as not meaningful. On this line was apparently also the Austrian physicist A. Zeilinger in his previous work when he affirmed that the distinction between reality and our knowledge of reality or also between reality and information cannot be made.[11] Thus, we face here an alternative. We could, in fact, admit that, without the contribution of a subject (able to build theories for conferring certain structural characters to both experience and what we call reality), we could not speak of any reality, according to a Kantian epistemological framework broadly understood (a weak form of idealism).[12] To a certain extent this needs to be postulated. However, this could also mean something much stronger, i.e. that only the mind exists, which would brings us to a form of extreme idealism on the line of the eighteenth century Anglo-Irish philosopher and Bishop George Berkeley (1685–1753), who affirmed that we cannot say that an object exists if not perceived: *esse est percipi*.[13] However, if this was the case, we would likely not fulfil the minimal requirement of objectivism that is necessary for scientific research (Sect. 2.1.1).[14] Physics (and science in general) cannot exist without some kind of separation, in what we call experience, between what must be attributed to the person (observer) and what need to be ascribed to Nature.[15] Otherwise, what is a physical theory for? Why we do not transform our whole science in a hyper-*epoché*[16] (putting everything else between parentheses) and reduce any problem to a psychological one? In fact, the problem would no longer be how we *describe* a physical reality, but it should become how we *construct* external reality (the Kantian or gnoseological–epistemological part) and come to *attribute* to it an objective or independent existence (the psychological part), a question that, in some measure, already concerned von Neumann's standpoint (Sect. 2.2.3). There is however, a significant difference between Wigner's and von Neumann's approaches: while for the latter there was a disagreement between mind and reality (the observer perceives something while the laws of physics describe another state of affairs), for the former there is full agreement (only what the subject observes can be called reality).

---

[11] Zeilinger (2005).

[12] Kant (1787). A point on which also Einstein agreed (Einstein 1934, p. 166).

[13] Berkeley (1710).

[14] It is not by chance that the idealist philosopher G. W. F. Hegel vindicated the supremacy of the 'dialectical method', that is, of the procedures over the facts.

[15] See Geroch (1978, p. 35).

[16] The epoché is a notion introduced in philosophy by the father of phenomenology (Edmund Husserl 1859–1938; Husserl 1931).

A Weak Idealism Is Insufficient

I am not sure that a coherent strong idealism was really Wigner's option. However, the alternative option (i.e. to assume that reality cannot be reduced to a construction of the mind although we know it through our sensory channels) leads us to immediate problem in such a context. The crucial problem for this *weak* idealism is: how to account for the fact that the fathers of QM succeeded to formulate such a complex mathematics able to make the correct predictions (or to compute the correct probabilities)? Indeed, we would be inclined to support a kind of agnosticism, and consequently we should affirm that our formalism works for reasons that we are finally unable to understand in our theoretical framework.

Summarising, we have here two different interrogatives:

1. How to account for the *agreement* between mind and reality if we like to exclude extreme idealism (denial of reality as such) and without resorting to a perpetual miracle?
2. Beyond this problem, there is another one: does it make sense at all to affirm that our access to reality does consist only of spot-like experiences connected through an abstract formalism when such a formalism has no ontological implications?

We might remind here that the first objection was the essence of Leibniz's criticism to the Cartesian school (Sect. 2.3.3). As a matter of fact, Wigner was very much worried by the problem of this agreement (a worry that he also shared with Einstein, as we shall see).[17] The second objection points out that our notion of physical world could risk to become very poor by leaving such a fundamental questions unanswered (if not negatively answered). In fact, this issue goes a long way beyond than the problem of isolated observations (and it is not by chance that Einstein criticised the reduction of a theory to observation). Of course, everything with which we deal in our experience goes in fact through our sensory or mental processes, so that we cannot assume anything about the world if not in dependence on us and our constructions.[18] This form of Kantism or weak idealism is clearly true, as already remarked: no observer is without bias.[19] However, it should be rather understood in the sense that we, e.g. perceive objects (understood in a wide sense) through *their effects on us* and are then able to somehow frame them in certain theories.[20] And to say that objects have an effect on us means that we *interact* with them (as stressed by Heisenberg), what presupposes that they are real. Moreover, since in the course of this interaction we are forced to certain conclusions that are 'imposed' to us as evidence, and this can only happen if reality exerts on us objective constraints, it seems that we cannot consider the external reality as a formless magma on which we would 'impose' our

[17] Wigner (1960).

[18] Carnap (1928), Margenau (1950, Sect. 4.5).

[19] Munowitz (2005, p. 43).

[20] Deutsch (1997, p. 57), Auletta (2011a, Chaps. 2 and 4). It seems to me that Einstein goes in the same direction when he says that there are conceptual constructions but they need to be validated in their reference to the "real" (Einstein 1949b).

theoretical structures. With the words of Peirce: "For what is observation? What is experience? It is the enforced element in the history of our lives. It is that which we are constrained to be conscious of by an occult force residing in an object which we contemplate."[21] This is the basis of the objectivism of any scientific enterprise, as remarked in Sect. 2.1.1.

In other words, from the correct point of view that we can have access to real facts and events only through our experience, it would be incorrect to infer that the only regularities that we find in the external world are due to the imposition of the structure of our mind. This is quite the other way around: we are able to discover these regularities thanks to a natural order that exists without us (a crucial tenet of Einstein too: see Sect. 2.3) and even pre-existed to those facts and data. This is so true that reality continuously *correct* us in our exchanges with it.[22] Indeed, what Kant considered as forms of a priori knowledge and thus (universally) fixed once and forever have in fact been shown to depend on the way we progress in empirical sciences. This is clearly true for the form of space that Kant considered as Euclidean (Sect. 2.3.2).[23] The Kantian approach could be considered as a consequence of the classical vision of the observer as detached spectator of the world (Sect. 2.2.2), with the difference that, after Kant, he/she no longer naïvely mirrors external reality but the latter is hidden (*noumenon*) behind the representations that the subject imposes on it, and so in fact it is quite the reverse.

Heisenberg's Instrumentalism

There is a further interpretation (that is broadly in the Copenhagen framework) that, not by chance, has some connections with Wigner's one. It is the interpretation endorsed by the young W. Heisenberg (and later on abandoned), according to which mathematics makes predictions about reality that we need to verify experimentally in the interaction with the object system.[24] Any other ontological question is here simply dropped and the fact that we observe the quantities predicted by the theory exhausts what we call physical reality. The reason for such an interpretation is the difficulty to make any ontological ascription to quantum systems independently of what is strictly observable (see also Sect. 2.2.2). There are certainly many differences between Heisenberg's and Wigner's standpoints: the major one seems to be that the latter scholar is a kind of idealist (the stress is on observing) whereas the young Heisenberg is not and seems rather an instrumentalist (the stress is on interacting). However, they converge on the fact that the relation between experimenter and object system (plus the formal aspects of the theory) exhausts what we call reality. I recall

[21] Peirce (1898, p. 170). A synthetic examination of this issue can be found in Auletta (2011c Sect. 2.1.4). See also Popper (1982, Preface).

[22] As pointed out in Peirce (1877). This was very well understood by the Scottish philosopher Thomas Reid (1710–1796) (Reid 1764).

[23] As also remarked in Heisenberg (1952, p. 22, 30).

[24] Heisenberg (1927). On the problem see also Auletta (2000, Sect. 6.2).

here that an instrumentalist standpoint was in fact disproved by the interaction-free experiment introduced in Sects. 1.2.4 and 2.1.2. It is not by chance that his proponent, M. Renninger, considered this ideal experiment as a falsification of Heisenberg's own interpretation of the uncertainty relations.[25]

Wigner's Friend

A well-known problem that deals with the difficulty of the von Neumann's cut (Sect. 2.2.3) is the so-called paradox of the *Wigner's friend*, which could be summarised as follows[26]: suppose that we describe the situation in which a friend observes a quantum system. Also this compound system (friend + object system + apparatus), which is clearly macroscopic, can be described by means of a state vector, for instance,

$$|\Psi\rangle_{SAF} = \sum_j c_j \, |j\rangle \, \left|a_j\right\rangle \left|f_j\right\rangle , \tag{3.16}$$

where the c_j's $\in \mathbb{C}$ satisfy the normalisation condition, the $|j\rangle$'s are the states of the object system, the $\left|a_j\right\rangle$ the states of the apparatus, and the $\left|f_j\right\rangle$ the states of the friend. I attract the reader's attention on the similarity of this equation with Eq. (2.49), although the physical model is very different. Now, it seems that it is only for an additional observer that this compound system acquires a particular reality and not for the friend that is here a described object (he/she is in an entangled state with both the apparatus and the object system). Therefore, we seem to fall back into solipsism: only from the point of view of *one* subject who is actually observing the compound system, it makes sense to speak of a determined output of measurement like $|k\rangle \, |a_k\rangle \, |f_k\rangle$. Although admitting that solipsism is compatible with quantum theory, Wigner answered in the spirit of London and Bauer: the observer could always ask the friend and obtain a determined answer, e.g. $|k\rangle \, |a_k\rangle$ or could compare what he observes and the friend observes. In fact, a theory in which the two descriptions were different, would not be a viable physical theory. So, we can definitively say that the subjectivist interpretation does not imply solipsism by itself. However, I stress that excluding solipsism does not well agree with a strong idealism, especially if the latter is grounded on personal–individual experience. This may explain some incertitude in Wigner's position.

[25] Renninger (1960). In his reply, annexed to Renninger's paper, Heisenberg essentially reiterates his instrumentalist standpoint.

[26] See Auletta (2000, Sect. 14.2.1).

3.1.2 *Objectivist Interpretation*

Statistics and Probability

The opposite interpretation seems to have its root in Einstein's work. In fact, EPR considered QM as a kind of provisional statistical theory (Sect. 2.3.1), and this represents somehow the bridge to the objectivist interpretation,[27] so that somebody has said that this is the statistical interpretation "taken literarily".[28] In fact, this statement is partly misleading. We have already noted the probabilistic character of QM (Sect. 2.4.2). As I shall explain in more details in the following, this probabilistic character does not depend on the number of particles involved, as it seems to be the case for statistical mechanics. In fact, even in the case in which we have a single photon in an interferometer, the two components between the two beam splitters still represent probability amplitudes or probabilities. This means that I accept the frequency interpretation of probability (Sect. 2.1.2) only in the sense that we expect that a statistics on a sufficiently large sample will approximate probability but not in the sense that it grounds the latter, which should be rather conceived as fundamental and irreducible. In fact, I have pointed out that quantum interference terms do not depend on the way in which we sample the event space, and thus, they need to be taken as rather *intrinsic* to any quantum state. Since the interpretation that we shall discuss now focuses on such an irreducible probabilistic character, it would be more accurate to consider it as a combination of the classical-realist stance of Einstein and the approach of Schrödinger that focuses on quantum correlations. Let us consider this question.

Everett's Contribution

The objectivist interpretation of measurement was formulated for the first time in 1957 by the American physicist Hugh Everett (1930–1982).[29] According to him, what does exist is what is described by the wave function. When we measure, what we physically obtain is a certain coupling between apparatus and object system, whose result is called by Everett *relative state* and is essentially the entanglement analysed by London and Bauer, as displayed on the RHS of Eq. (2.43) or (2.44). According to him, this *exhausts* the measurement process. As mentioned, Schrödinger had indirectly contributed to this solution. In fact, in his famous paper on the situation of QM, he clearly says that to measure is simply to couple two systems in such a way that we can make *conditional predictions* (if the apparatus is a certain state, the object system must be in a corresponding state).[30] However, conditional predictions

[27] As pointed out in Ballentine (1970). On the history of the problem see Jammer (1974). See also Adler (2004) and Auletta (2000, Sects. 6.5, 23.1) for a summary of these interpretations.

[28] Peres and Zurek (1982).

[29] Everett (1957).

[30] Schrödinger (1935a).

do not require detection events (and this clarifies London and Bauer's standpoint). In fact, Schrödinger considered that the whole worry of measurement derives from the isolation of the object system that, in a final state of the kind (3.15), is considered *as not* in entanglement with the apparatus (what, in his opinion, would make measurement meaningless). This is the way in which he also solves the so-called cat paradox (Sect. 2.2.5): according to Schrödinger, the apparent weirdness here only arises through the (local) extrapolation of the object system (here, the cat) out of a physical situation described by Eq. (2.49). It is clear from this, that, for Schrödinger, any object of the universe is in fact a quantum-mechanical system interconnected with all other quantum systems, and also the mind (of the observer and the observation act itself) is (are) part of this universal entanglement. This explains why Everett's interpretation (and Schrödinger's own position), although being to a certain extent near to Einstein's realistic tenet, are very far away from it on a crucial point: the issue of the separability of the systems discussed in Sects. 2.3.2–2.3.3.

Now, our experience tells us that an observer only observes a single component of this objective superposition, what seems to clash with Everett's explanation of measurement so far. In fact, if we assume with Everett that only the entangled state described by QM exists, we cannot admit that *only one* of its components is actually observed: according to him, this would introduce an unexplainable asymmetry and even a kind of irrationality in the heart of physics (Sects. 2.4.1–2.4.2). Now, we are well aware that different observers in alternative detection acts *could* perceive different components of this entanglement (a counterfactual statement). Coherently, Everett assumes that all components *are in fact observed*, and all these observation acts, as far as their physical reality is concerned, are already included in the same description of the quantum reality (what is not far away from the Wigner's friend, as far as the issue at stake is the object + observer description only). The problem perceived by Everett was well-known in the history of the philosophical thought: many supporters of a classical deterministic understanding of our world (the most famous one is likely to be the already mentioned B. Spinoza) have assumed that everything that can exist also exists precisely with the aim to avoid that only some state of affairs happen among many possible ones, what would be not rationally explainable.[31]

D. Deutsch points out that, in his Dublin conference held in 1952, Schrödinger admitted that his equation described several histories and therefore could be considered as the first scholar to have envisaged this solution.[32] Without denying that he gave substantial hints in this sense, as mentioned, I am more inclined to think that he preferred the idea of extrapolation, what would make Schrödinger rather a precursor of the decoherence solution (to be developed in the next subsection), or at least his standpoint could be located before the split between decoherence and Everett's own approach (in fact, decoherence took its first steps from a criticism to the objectivist interpretation).

[31] Spinoza (1677). The reader interested in this kind of subtleties may consult (Auletta 1994, Chap. 2).

[32] Deutsch (2011, Chap. 12).

Many Worlds

Everett's solution is quite amazing because it is a kind of solipsism but objective! It is even a multiple solipsism as far as each observer is 'trapped' in a component of the wave function. In fact, none of these alternative observers communicate with any of the others and likely none is even aware of their existence. Such hypothesis appears not fully justifiable in the framework of our physical universe or at least of our common experience of our universe: what does it mean to say that different observers measure simultaneously the same wave function? What is the reason for that? Where are these different observers 'located'? This explains why, later on, the American physicist Bryce DeWitt (1923–2004) introduced the idea that each observer of the same system actually represents a kind of counterpart of the other observers, inhabiting each one a *different world* and getting a different outcome when measuring the *same* wave function.[33] So, what exists is actually a *multi-universe* or a multiversal wave function, and the counterpart observers themselves are part of it but each one 'inhabiting' a different universe. In this way, the so-called many-worlds interpretation (MWI) was born.

Bases Degeneracy

The MWI has some advantages on the subjectivist interpretation, or at least meets the physicists' feeling that there must be an independent reality (as stressed, e.g. by Einstein). Moreover, it takes QM in full account. In fact, it is followed by a substantial group of scholars, for instance, many working on quantum information. However, it has its flaws when it is understood as an explanation of measurement. In fact, if to measure would coincide with the premeasurement step (described by Everett's relative state), a weird consequence followed: the formalism of QM shows that the same apparatus–object system state, as displayed by Eq. (2.43) or (2.44), can be expanded in very different bases, as pointed out by the Polish-American physicist Wojciech Zurek, what is called *bases degeneracy*.[34] Let us suppose that we wish to measure a generic observable (see Eq. (1.96))

$$\hat{O} = \sum_j o_j \left| o_j \right\rangle \left\langle o_j \right| , \tag{3.17}$$

where $\left| o_j \right\rangle$ are its eigenstates with o_j the corresponding eigenvalues. Suppose that system and apparatus as a whole evolves according to an analogous of transformation (2.44):

$$|\psi\rangle|A_0\rangle \longmapsto \sum_j c_j |o_j\rangle|a_j\rangle, \tag{3.18}$$

[33] Dewitt (1970, 1971). See also Everett (1973).

[34] Zurek (1981). See also Auletta and Wang (2014, Sect. 9.4).

where, as before, $|\psi\rangle$ is the initial state of the object system and is a superposition of the eigenstates $\left|o_j\right\rangle$, $|A_0\rangle$ is the initially ready state of the apparatus, $\left|a_j\right\rangle$ are the pointer states of the apparatus corresponding to $\left|o_j\right\rangle$, and the c_j's $\in \mathbb{C}$ satisfy the normalisation condition. As we have seen, Everett calls such transformed state a relative state and assumes that measurement is accomplished already at this stage and nothing happens thereafter. However, we know that the same state can be expanded in different bases. Indeed, using the completeness relation $\forall \left|a'_k\right\rangle$

$$\sum_k \left|a'_k\right\rangle\left\langle a'_k\right| = \hat{I}, \tag{3.19}$$

we can now write the states of the apparatus as (see Eqs. (1.114))

$$|a_j\rangle = \sum_k \left\langle a'_k \mid a_j\right\rangle |a'_k\rangle. \tag{3.20}$$

Therefore, the laws of QM do not forbid us to write the RHS of transformation (3.18) as

$$\begin{aligned}\sum_j c_j \left|o_j\right\rangle \left|a_j\right\rangle &= \sum_k \left(\sum_j c_j \left\langle a'_k \mid a_j\right\rangle \left|o_j\right\rangle\right) \left|a'_k\right\rangle \\ &= \sum_k \left|o'_k\right\rangle \left|a'_k\right\rangle, \end{aligned} \tag{3.21}$$

where the states $\left|o'_k\right\rangle$ are given by

$$|o'_k\rangle = \sum_j c_{jk} \left|o_j\right\rangle, \tag{3.22}$$

and where (see also Eq. (1.400))

$$c_{jk} = c_j \left\langle a'_k \mid a_j\right\rangle. \tag{3.23}$$

Note that, in the last line of Eq. (3.21), we obtain two different bases for both the object system and the apparatus. This seems to imply, in Everett's terminology, that the relative state on the RHS of transformation (3.18) represents not only the measurement of the observable $\hat{O}$ but also of the observable

$$\hat{O}' = \sum_j o'_j \left|o'_j\right\rangle\left\langle o'_j\right|, \tag{3.24}$$

which in general does not commute with $\hat{O}$. Of course, also the inverse transformation can be written. Indeed, since we have

$$\left|o'_k\right\rangle = \sum_j c_j \left\langle a'_k \mid a_j \right\rangle \left|o_j\right\rangle \text{, and } \left|a'_k\right\rangle = \sum_j \left|a_j\right\rangle \left\langle a_j \mid a'_k \right\rangle, \tag{3.25}$$

we can also write

$$\begin{aligned}
\sum_k \left|o'_k\right\rangle \left|a'_k\right\rangle &= \sum_k \sum_j c_j \left\langle a_j \mid a'_k\right\rangle \left\langle a'_k \mid a_j\right\rangle \left|o_j\right\rangle \left|a_j\right\rangle \\
&= \sum_j c_j \left\langle a_j \right| \left(\sum_k \left|a'_k\right\rangle \left\langle a'_k\right| \right) \left|a_j\right\rangle \left|o_j\right\rangle \left|a_j\right\rangle \\
&= \sum_j c_j \left\langle a_j \mid a_j\right\rangle \left|o_j\right\rangle \left|a_j\right\rangle \\
&= \sum_j c_j \left|o_j\right\rangle \left|a_j\right\rangle ,
\end{aligned} \tag{3.26}$$

where in both the forelast and last steps I have used the completeness condition. It is interesting to note that we got the new object observable's eigenbasis $\{\left|o'_k\right\rangle\}$ through the overlap terms $\left\langle a'_k \mid a_j\right\rangle$ between two different apparatus observables' bases.

Multiple Observables

Summarising, according to the above calculus, it is like to measure different observables simultaneously (each of them having its own eigenbasis). However, most of these observables do not commute. Therefore, there seems to be some confusion here between two very different issues:

- The *formal* possibility to expand the joint state of the object system and apparatus in different bases and
- The *experimental* set-up for measuring a particular observable (the premeasurement step), a physical context that is chosen by the observer.

If we interpret the experimental set-up as instantiating that possibility, such a multiple "measurement" would imply a violation of the uncertainty relations, and this cannot be the case. If, on the contrary, we interpret the formal expansions as exhausting the description of measurement, this would force us to accept a definition of measurement that seems to deprive it of any significance, since to affirm that any possible observables would be jointly measured would dissolve the very act of decision that is connected with any experimental operation (this is also true for classical physics). Such a situation would represent the simultaneous presence of any experimental set-up procedure that the system can undergo. In fact, there is a connection between bases degeneracy and the character of the physical process involved here: as mentioned in the previous subsection, at least in principle, such a premeasurement step, which is described by unitary operators, can be made reversible, a statement that need to be taken for granted by now, because it needs sophisticated experiments to

be factually and positively proved.[35] This shows that up to this point Everett is right with his theory of the relative state. However, the fact remains that, even in reversible couplings, we always deal with a specific experimental set-up that selects a certain observable and not others, and therefore selects also a certain basis (in fact, we speak of complementary set-ups or settings); and the latter imposes particular *physical conditions* that cannot be changed without doing simply another kind of experiment. For instance, the set-up $D_A - D_B$ or $D_{A'} - D_{B'}$ in the delayed-choice experiment, as displayed by Fig. 2.15, Sect. 2.4.3, represent two alternative and complementary experiments.[36] In other words, measurement is by definition an interaction that deals with a certain specific observable (and relative set-up), and we shall show that the choice of measuring a certain observable is irreducible because it is purely local and cannot be manipulated from the exterior of the experimental context. This is an issue that is much more fundamental than the question whether we obtain only one of the eigenstates of the measured observable or not (von Neumann's problem: Sect. 2.2.3).

Selection Acts and Probabilities

It seems that by trying to cope with the impossibility to get (through detection) an eigenstate of a certain observable through unitary evolution, Everett and the other objectivists might have incurred a far bigger problem, a possible confusion between dynamics and formalism. With the words of W. Zurek,[37] "in a completely quantum universe one is forced to seek sets of preferred, effectively classical but ultimately quantum states that define branches of the universal state vector, and allow observers to keep reliable records. Without such a preferred basis relative states are 'too relative', and the relative state approach suffers from basis ambiguity." Obviously, the supporter of the MWI might deny that these acts of selections or decisions that we call measurements occur at all. In such a case, there would be no "preferred" or chosen basis[38] and the problem of measurement would disappear or better would be explained away.[39] Thus, the crucial issue here is the following: are we authorised to admit (or to deny) that in our universe selections of experimental contexts occur? And since the MWI also denies the singularity of the detection act, and detections are only a particular (i.e. controlled) case of physical interactions, are we authorised or not to admit that in our universe interactions occurs at all? Schrödinger seemed inclined to deny such occurrences. But if only *detection* events do not occur, what is the impediment here? Moreover, if the probability amplitudes are no longer connected with the probabilities of possible detection events, what is their meaning at all? Note that these probabilities are what connects the theory with experimental

[35] Basic references are Wang et al. (1991), Scully et al. (1991).

[36] For the relevance of the context see also Epperson and Zafiris (2013, p. 92).

[37] Zurek (2010, p. 410).

[38] As affirmed in Wallace (2012, Sect. 1.9).

[39] A certain reject of the measurement problem by the scientific community is well expressed by the title "Against'Measurement"' of reference (Bell 1990).

evidence. On the other hand, if in a multiverse every possible event happens (with certitude), does it make sense to assign such probabilities?[40] The issue at stake is not how we interpret probability[41] but how we bring in agreement the fact that possible quantum events have *specific* probabilities amplitudes (and therefore probabilities to occur) with the MWI's assumption that *none* or *every* outcome does occur.[42] If so, it seems to me that we should always have equiprobability of events, but this is falsified from what we know about the dynamics of quantum systems.

Alter Ego

It may be recalled that the MWI is strongly reminiscent of Leibniz's theory of the possible worlds. And this not by chance, because both Leibniz and the supporters of the MWI were lead to this conclusion by assuming the universality of correlations (Sect. 2.3.3). In fact, there are reasons to think that also Leibniz tacitly assumed that all of these worlds in fact exist or should exist.[43] This is an issue that in analytic philosophy was brought to attention through the work of the American philosopher David Lewis (1941–2001).[44] To recall Leibniz in such a context is particularly instructive for the following examination, since he came to the conclusion that the different possible instantiations of myself in different worlds would be rather literarily alter egos or lookalikes, i.e. *resembling* me but not being me or alternative "myselves" (characterised by different properties). Leibniz was brought to this consequence during his correspondence with the French Catholic theologian and philosopher Antoine Arnauld (1612–1694), a follower of Descartes who could not admit that there can be (even at the level of a pure possibility) different "myselves" due to the fact that any of us can say"myself" only through an act of self-reference that includes the whole of his/her properties and his/her biography.[45] It is likely that Leibniz finally agreed with this conclusion precisely starting from the idea of the universal entrenchment: in fact, if everything is entrenched with everything else, as recalled in Sects. 2.3.2 and 2.3.3, any minute change in our universe will determine trillions of other changes, and therefore, we cannot have two identical persons displaying a *single* difference, namely that one observes this component and its alter ego another one.

Let us come back to the problem of measurement. D. Deutsch, one of the most authoritative proponents of the MWI, is well aware that in any experiment of, e.g. interferometry many of the detection results (depending on the experimental arrange-

[40] The problem was discussed in Graham (1973) where the solution offered appears to be *ad hoc*. See also Maudlin (1994, p. 5). For a collection of different points of view on such an issue see Saunders et al. (2010, Parts 3–4).

[41] As in Wallace (2012, Chap. 4).

[42] On this stuff see also Norsen (2017, Sect. 10.2).

[43] Leibniz (1686, 1710). For an examination of the problem see also Auletta (1994, Chap. 4).

[44] Lewis (1986).

[45] Leibniz an Arnauld, 14th July 1686, in Leibniz (2019, II, pp. 53–54): "Si dans la vie de quelque personne et même dans tout cet univers quelque chose alloit autrement qu'elle ne va, rien nous empêcheroit de dire que se seroit une autre personne ou un autre universe possible".

ment) can be explained only in the case in which something does interfere with the measured particle. Now, Deutsch tells us that either we accept that, for any real photon that undergoes a certain process and is finally detected, e.g. on a screen, there are trillions of "shadows photons" that exist in other worlds, are diversely captured by the same screen (but in different universes) and thus interfere with 'our' photon, or we should resort to an action at a distance or even admit that nothing at all is responsible for such an interference.[46] Is this alternative well posed? Is it not thinkable that there is in fact a 'real' interference but in the framework of our universe? Let us consider more carefully this problem.

We should agree with Schrödinger that all quantum systems of our universe are more or less entangled. However, according to Leibniz's lesson about the "myselves", it is reasonable to say that, if we like to minimally change whatever event in our universe, this will affect many other possible events elsewhere in the same physical universe. Note that, due to entanglement, this happens in the same moment in which a detection event occurs (a statement that will be fully understood below), so that the idea of a subsequent ramification of the different universes is of no use here. However, if so, we cannot establish a continuum such that any alternative to each contingent event of our universe happens in another universe or in another possible ramification of our universe that would be equal to our own but different only for *this single happening*.[47] In other words, we cannot say of the *same* photon that here in our universe behaves in a way and in another universe or branch behaves slightly differently. Neither that the *same* screen in our universe catches a photon at a certain location and in another one, say, a nanometre further. It is thus quite likely that our photon and our screen only exist in our universe, and we come back to the issue of the real and particular physical conditions that we consider when dealing with measurement as such. Summarising, due to quantum entanglement (and likely also to classical correlations) the infinite sea of possibilities cannot be taken as an explanation of the measurement puzzle. It seems that the supporters of the MWI have underestimated the role of quantum correlations inside a *single* universe or branch of the universe (as we shall see, they have in fact converged with a theory of local observables in the sense of EPR *within a single* universe).

The Fabric of Reality

It seems also clear that the MWI presents a wider aspect that goes much further than an explanation of measurement. If the MWI is taken seriously in all its generality, it appears to be an explanation that tells us that the only thing to really exist is a huge superposition of all possibles (or possible events). Any conceivable detection or even

[46] Deutsch (1997, p. 93).

[47] As pointed out in Auletta (2009). Once, addressing the audience of a symposium, D. Deutsch has said: "I'll start with a simple fact: in this room, in some nearby universes, Hugh Everett is here with us, celebrating. Perhaps he's there, in that seat where Simon is. And therefore, in those universes, Simon is somewhere else" (Deutsch 2010, p. 542).

premeasurement or whatever interaction is in fact only 'located' in an alternative stream of this 'sea' of possibilities, the "fabric of reality", of which D. Deutsch says that no physical object in the ordinary sense of the world could "scratch the surface".[48] This is a very significant widening of the perspective. Under this point of view, as it was already true for the subjectivist party, the problem of measurement is connected with more general issues about the fundamental characters of reality.

In fact, if all possibilities are instantiated in some world, where is the difference between the notions of actual and possible? To be actual would become here only a question of contextual reference (or indexicality): to say "this object exists or is real" becomes like to say "here" or "now" in other contexts, since, to that expression, the crucial specification "in our world" should be (or is implicitly) added. In such a case, QM would be essentially a theory of relative frames and their connections, as SR is, but this time extended over different universes. To be able to do this we would need a whole physics allowing us a new understanding of physical problems (as it was in fact the case for SR). I think that is helpful for supporters of the MWI to consider the problem in this way.

On the other hand, Deutsch's attribution of reality to possibilities is to be taken seriously.[49] However, in the following I would try to disentangle the issue of both the status of all possibles and the possible existence of other worlds (that deserves a specific analysis later on) from the measurement problem. Clearly, if these possible worlds are not the explanation of the measurement process, it becomes natural to ask what are they for, and what is the relation of these universes with the whole universal fabric of reality. Avoiding by now to answer the second question, I mention that other universes could be conceived as gross alternatives to our universe such that they instantiate different laws of physics. In such a case, we could no longer speak of alternative events since it is likely that the quanta of another universe (supposing that there are elsewhere) are, e.g. of different kinds or are characterised by different properties. In fact, this possibility has been really explored. As is well-known, the fundamental aspects of our world (the structure of matter, the constitution and life cycle of stars, the emergence of life, to quotes some of them) depend on very fine-tuned values of twenty basic physical constants.[50] This is one of the major puzzles of science and a true problem also for philosophy. Some scientists have tried to deal with this problem by assuming that there are in fact many other universes apart from our own, each one displaying a specific combination of these values.[51] The number of these universes can be really huge: perhaps 10^{200} or even more. This is obviously matter of further research. Other, more economic, solutions of the above problem

[48] Deutsch (1997, p. 126).

[49] Wallace (2010, p. 54).

[50] The problem was first raised in Barrow and Tipler (1986). For a different point of view on the subject see Rees (1999). See also Tegmark et al. (2006). The existence of natural constants does not seem to be in contrast with the relational standpoint that I support. With the words of Robert Laughlin, "all the fundamental constants require an environmental context to make sense" (Laughlin 2005, p. 19).

[51] Tegmark (2003).

have been also proposed, like a mechanism of universal Darwinian selection.[52] These solutions are not necessarily in contradiction with the assumption that all possibilities are somehow real and can be even well combined with the idea of a fabric of reality, and in the following, we shall explore this possibility.

It is finally worth mentioning here that some philosophers have made a merging between the MWI and the subjectivist interpretation: scholars like the British philosopher Michael Lockwood have interpreted the multiverse wave function as describing an universal Mind, of which any finite observer is only a part, an idea that also Schrödinger could have shared.[53] However, this seems to be a strong speculation that leads us very far away from the epistemological requirement of the Occam's razor.

3.1.3 Role of the Environment

Local Interactions and Global Interconnections

A different idea about measurement has been developed step by step between the 1960s and 1980s, especially as a consequence of the introduction of quantum optical devices such as lasers, that allowed to consider the relation between a quantum system and its environment for the first time.[54] It can be considered as a standpoint overcoming the unilaterality of both the subjectivist and objectivist interpretations. In fact, in this perspective, the so-called reduction of the wave packet could be the consequence of two different facts:

(i) A spontaneous coupling of the object system with the apparatus and the *environment*, on the one hand, and
(ii) The *local interaction* between object system and apparatus which occurs when measuring, including detecting, on the other hand.

The main idea, which is called *decoherence*, is that, when we locally couple object system and apparatus, we are unaware of the entanglement that these two systems have with the rest of the world.[55] Then, when we consider these two systems alone and in particular their local interaction, we are in fact doing an extrapolation out of a global system (see the previous subsection). Mathematically speaking, such a procedure is obtained by performing a partial trace (excluding the environment from our considerations) of the kind (1.409) out of a state with the form (1.408).

---

[52] Zurek (2004, 2007, 2009). See also Smolin (1997).

[53] There is a whole debate on this subject: see Lockwood (1996), Deutsch (1996), Chalmers (1996). See also Auletta (2004b). This was actually an important idea in oriental philosophy: see Schrodinger (1958, Chap. 4) and also Shimony (1981).

[54] The first seeds of this new approach can be found in Zeh (1970). For a wide but rather technical account see Auletta et al. (2009, Sect. 9.4).

[55] This was the subject of what is become now a classical study (Zurek 1982). For a good and extensive introduction to decoherence see Schlosshauer (2007).

The Solution in Formal Terms

The measurement process, after the initial preparation of the object system, can be schematically divided into two distinct steps followed by the partial trace and eventually detection, which completes the process. The initial state (at time $t_0 = 0$) is for the sake of simplicity a factorised state of the object system $\mathcal{S}$, the environment $\mathcal{E}$, and the apparatus $\mathcal{A}$. Then, at time $t = t_1$, due to the interaction between $\mathcal{S}$ and $\mathcal{A}$, these become entangled (we have a premeasurement). In the time interval $t_1 \leq t \leq t_2$, also the environment entangles with $\mathcal{S}$ and $\mathcal{A}$. At time t_2, the interaction is switched off. Formally, we need to expand the transformation (2.44) as follows:

$$\begin{aligned} |\Psi_{SAE}(0)\rangle &= |\psi\rangle|A_0\rangle|E_0\rangle \\ &\longmapsto |\Psi_{SAE}(t = t_1)\rangle = \left[\sum_j c_j(|j\rangle|a_j\rangle)\right]|E(t_1)\rangle \\ &\longmapsto |\Psi_{SAE}(t \geq t_2)\rangle = \sum_j c'_j|j\rangle|a_j\rangle|e_j\rangle\ , \end{aligned} \tag{3.27}$$

where $\{|e_j\rangle\}$ is some basis for the environment, the $|j\rangle$'s are eigenstate of the observable that we have selected, and the c_j, c'_j's $\in \mathbb{C}$ satisfy the normalisation condition. At time t_1, the connection between $\mathcal{A}$ and $\mathcal{S}$ that we have called premeasurement has been provided and is described by a unitary operator of the kind (3.3). The last line shows the action of the environment, and can be described again by a unitary operator which couples $\mathcal{E}$ to $\mathcal{S} + \mathcal{A}$, which takes the general form:

$$\hat{U}_t^{SA,E} = e^{-\frac{i}{\hbar}t\hat{H}_{SA,E}} \tag{3.28}$$

for $t_1 < t < t_2$, where $\hat{H}_{SA,E}$ is the coupling Hamiltonian between object–apparatus and environment. Now we can trace the environment out by writing the corresponding reduced density matrix

$$\begin{aligned} \hat{\tilde{\varrho}}_{SA} &= \mathrm{Tr}_{\mathcal{E}}\left(\hat{\rho}_{SAE}\right) \\ &= \mathrm{Tr}_{\mathcal{E}}\left[|\Psi_{SAE}(\tau)\rangle\langle\Psi_{SAE}(\tau)|\right] \\ &= \sum_j |c_j|^2|j\rangle\langle j| \otimes |a_j\rangle\langle a_j|, \end{aligned} \tag{3.29}$$

where $\tau \geq t_2$. Equation (3.29) can be considered both as a generalisation of Eq. (3.12) as well as an instance of Eq. (1.409), and represents the required Lüders mixture. Note that it has been obtained under the simplified hypothesis represented by the last line of Eq. (3.27). Therefore, we have found something that, at least on a formal level, describes what in fact happens during a measurement. This explanation has been called *decoherence* because the interference terms that connect the apparatus

Table 3.1 Dust grains have about the size of 10^{-3} cm while large molecules of about 10^{-6} cm. The quantity considered in the left columns is the characteristic rate (computed in $\text{cm}^{-2}\text{s}^{-1}$) at which spatial coherence becomes suppressed as a consequence of spontaneous decoherence. The quantity considered in the right columns is the relative estimate of decoherence timescales. Adapted from Schlosshauer (2007, p. 135)

	Suppression rate		Decoherence timescale	
Environment	Dust grain	Large molecule	Dust grain	Large molecule
Cosmic background radiation	10^{6}	10^{-12}	1	10^{24}
Photons at room temperature	10^{24}	10^{6}	10^{-18}	10^{6}
Best laboratory vacuum	10^{20}	10^{14}	10^{-14}	10^{-2}
Air at normal pressure	10^{37}	10^{31}	10^{-31}	10^{-19}

and the object system go rapidly down and the two systems are no longer in the coherent state represented by an entanglement (although in the literature the term *coherent state* denotes a specific state of light, relevant for lasers, as we shall see).

Controlled and Uncontrolled Processes

It is worth noticing that here measurement becomes a special case of a wider class of interactions among *open systems*, that is, among those physical systems that are open to the environment.[56] In fact, it is assumed here that any quantum system is always connected with the environment, according to Schrödinger's view reported in the previous subsection. As a matter of fact, Schrödinger had already the insight that measurement can induce a kind of loss of correlation (decoherence) among the different components of a superposition or entangled state.[57] Therefore, it is likely that what we observe in a *controlled form* in our laboratory is to a certain extent also what spontaneously happens in nature several times in *uncontrolled form* when open systems interact. Table 3.1 shows some calculation of decoherence rate and time for different physical systems and in different environments.[58] A simple (idealised) model of such kind of process is represented by two wave packets of width l centred at $x = \pm x_0$:

[56] A point of view strongly supported in Joos and Zeh (1985). See also Giulini et al. (1996), Schlosshauer (2007, Chap. 3).

[57] Quoted in Mehra and Rechenberg (1982, VI, p. 754).

[58] Joos and Zeh (1985), Schlosshauer (2007, p. 135).

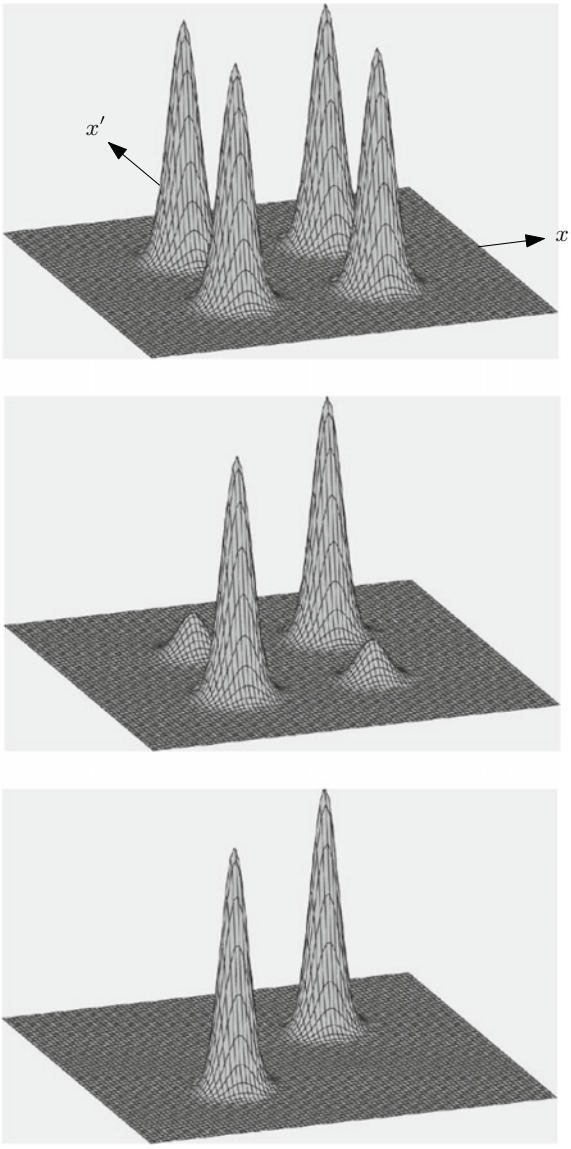


Fig. 3.1 Time evolution of the initial density matrix $\hat{\rho} = \psi(x, t_0)\psi^{\star}(x', t_0)$ of two wave packets described by the wave function (3.30). Interference terms along the off-diagonal $x = -x'$ become progressively damped due to the scattering of environmental quanta. Adapted from Schlosshauer (2007, p. 149)

$$\begin{aligned}\Psi(x, t_0) &= \psi_{x_0}(x, t_0) + \psi_{-x_0}(x, t_0) \\ &= \left(\frac{1}{2\pi l}\right)^{\frac{1}{4}} \left(e^{-\frac{(x-x_0)^2}{4l^2}} + e^{-\frac{(x+x_0)^2}{4l^2}}\right). \end{aligned} \tag{3.30}$$

Computing the density matrix $\hat{\rho} = \psi(x, t_0)\psi^{\star}(x', t_0)$, its subsequent time evolution is displayed by the plots in Fig. 3.1.

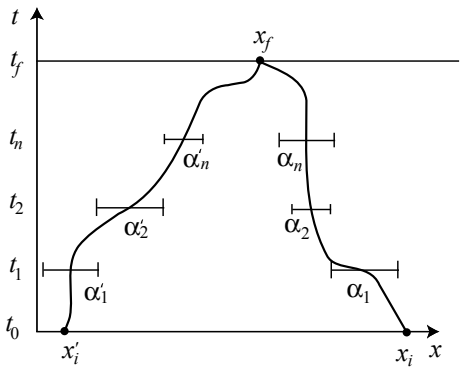


Fig. 3.2 Decohering histories. Here two histories, α and α' are shown, leading to a common final event $\hat{P}_{x_f}$, where x is some observable (e.g. position). Adapted from Auletta et al. (2009, p. 290)

The importance of this standpoint lies in the fact that for the first time we have found a connection between measurement and spontaneous quantum dynamics (of open systems), and therefore, the former might no longer appear as something not only mysterious but even conflicting with the latter (Sects. 2.2.3–2.2.4). Thus, the issue at stake is not only a question of interpretation but concerns the theoretical coherence and even consistency of the theory.

I mention here also a related approach: the decoherent histories, an approach, which, understood as general branching of universes, can be very helpful here.[59] Suppose that at each (discrete) time $t_1, t_2, \ldots$, the quantum dynamics of the whole universe projects into (binary, for simplicity) alternatives, and that the history that collects these two alternatives are called α, α'. Then, we can build a decoherence functional in this way (Fig. 3.2)

$$\mathcal{D}(\alpha, \alpha') = \mathrm{Tr}\left[\hat{P}_{\alpha_n}(t_n)\hat{P}_{\alpha_{n-1}}(t_{n-1})\ldots\hat{P}_{\alpha_1}(t_1)\hat{\rho}_U\hat{P}_{\alpha'_1}(t_1)\ldots\hat{P}_{\alpha'_{n-1}}(t_{n-1})\hat{P}_{\alpha'_n}(t_n)\right], \tag{3.31}$$

where $\hat{\rho}_U$ is the state of the universe at $t = 0$. It may be noted that the decoherent-history approach acknowledges a cosmic time arrow, in accordance with the conclusions of Sect. 2.4.3. The new approaches to decoherent histories, are focused on 'endogenous' models of decoherence, since there is no possible environment for the entire universe.[60]

Linear Transfer of Information

Although attracting, this interpretation presents a problem: since the state of the object system can be prepared in any arbitrary superposition, we may wonder that, when we couple it with the apparatus, the latter spontaneously evolves into a state

[59] See Hartle and Hawking (1983), Griffiths (1984), Halliwell (1993). See also Wallace (2012, Sect. 3.9).

[60] Halliwell (2010).

whose components are precisely the eigenstates of the observable that we wish to measure. The worry is well grounded, since we know that, before such a coupling,

(i) The object system itself, when being prepared in a certain state and before having chosen a certain coupling set-up, can be indeed assumed to be expanded in any possible basis, as displayed by the discussion in the previous subsection, and
(ii) The apparatus is in a default state that has no relations whatsoever with the eigenstates of the observable we are measuring (or, if its is entagled, it could be so in ways that are not necessarily Lüders mixtures, having the form of the second (or third) line of Eq. (3.27)).

Therefore, we have in fact two problems: how the object system is driven (and maintained) into a state whose components are the eigenstates of the measured observable (the issue of the preferred basis) and how an appropriate coupling arises at all. In the following, we shall clarify that entanglement (here between apparatus and object system) is actually a form of mutual information,[61] what means that, when the system and the apparatus are entangled, they share information. This process will bring the apparatus from an initial steady state into a state in which there is such coupling with the components of the object system. Then, we can say that premeasurement allows for a *linear transfer of information* from the object system to the apparatus. Postponing by now the analysis of this process, let us ask which are the conditions that allow a linear transfer of information at all. It can be proved[62] that, if the apparatus should discriminate among the components of the superposition state of the object system (as we expect to happen) allowing them to survive intact a successful information transfer (so that the immediate re-measurement can yield the same result), these components need to be *orthogonal*. In fact, a linear transfer of information, in the binary case displayed in Eq. (2.43), implies:

$$|u\rangle\,|A_0\rangle \longmapsto |u\rangle\,|A_u\rangle \text{ and } |d\rangle\,|A_0\rangle \longmapsto |d\rangle\,|A_d\rangle\,. \tag{3.32}$$

If the states $|u\rangle, |d\rangle$ of the object system remain untouched (guaranteeing repeatability of measurement), we have preservation of the scalar products (under unitary transformation: Sect. 1.2.5):

$$\langle u \mid d\rangle = \langle u \mid d\rangle\,\langle A_u \mid A_d\rangle\,. \tag{3.33}$$

Now, it is evident that when $\langle u \mid d\rangle \neq 0$, this implies $\langle A_u \mid A_d\rangle = 1$. In other words, the states of the apparatus $|A_u\rangle$ and $|A_d\rangle$ can differ from each other and therefore *discriminate* between the two possible outcomes of detection only when $\langle u \mid d\rangle = 0$, that is, when the states of the object system are orthogonal. In other words,

1. The latter need to represent an orthonormal basis and in particular be eigenstates of a single observable.

[61] Zurek (1989a, b), Barnett and Phoenix (1989), Cerf and Adami (1997). See also Auletta et al. (2009, Sect. 17.2), Auletta and Wang (2014, Sect. 11.7).

[62] Zurek (2007, 2013).

2. Where we assume that the latter is the observable that in fact has been selected through the premeasurement of the system.

In such a way, the conditions are established for detection as a finally selection of one of them. If it is so, this would mean that the required orthogonality among the components of the state of the object system allows us to bridge between the *linearity* of the information acquisition of premasurement and the *nonlinearity* of detection. Therefore, in appropriate environmental conditions, the selection of a given observable (and thus the existence of a preferred basis) has to be, at least locally, effective, what makes decoherence satisfy Born's rule for probability (Sect. 1.1.3) that has no citizenship in the MWI.[63]

Let us first show that there is always a preferred basis (no basis degeneracy) when we include the environment as third system. This is known as the uniqueness of triorthogonal decomposition.[64] First, note that sets of linearly independent vectors cannot differ only trivially, i.e. only for a global phase factor. Now, consider the state vector

$$|\Psi\rangle = \sum_j c_j |a_j\rangle|b_j\rangle|e_j\rangle, \tag{3.34}$$

where $\{|a_j\rangle\}$, $\{|b_j\rangle\}$, $\{|e_j\rangle\}$ are orthogonal sets of vectors, respectively, in the Hilbert spaces $\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3$ for three generic systems $\mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_3$, respectively. Now, I claim that, even if some of the $|c_j|$'s are equal, no alternative orthogonal sets $\{|a'_j\rangle\}$, $\{|b'_j\rangle\}$, $\{|e'_j\rangle\}$ exist such that $|\Psi\rangle = \sum_j c'_j |a'_j\rangle|b'_j\rangle|e'_j\rangle$, unless each alternative set of vectors differs only trivially from the set it replaces. In fact, assume, without loss of generality, that $\{|e_j\rangle\}$ differs not trivially from $\{|e'_j\rangle\}$, and let us write

$$|\Psi\rangle = \sum_j c_j |w_j\rangle|e_j\rangle, \tag{3.35}$$

where $|w_j\rangle = |a_j\rangle|b_j\rangle$. Now, suppose that

$$|\Psi\rangle = \sum_j c'_j |w'_j\rangle|e'_j\rangle, \tag{3.36}$$

where $|w'_j\rangle = |a'_j\rangle|b'_j\rangle$. Then, it is clear that we cannot rewrite the factorised state $|a'_j\rangle|b'_j\rangle$ as an entangled state. But, since we have Eq. (3.35) and since $\{|e_j\rangle\}$ differs not trivially from $\{|e'_j\rangle\}$, then the expansion (3.36) is correct only if $|w'_k\rangle = \sum_j c_j^{(k)} |w_j\rangle$ Eq. (3.20), where at least two of the $c_j^{(k)}$'s are non-zero. But from $|w_j\rangle = |a_j\rangle|b_j\rangle$ it follows that $|w'_k\rangle$ is an entangled state, i.e.

[63] Zurek (2010).

[64] Elby and Bub (1994). See also Auletta et al. (2009, Sect. 9.4.1).

$$|w'_k\rangle = \sum_j c_j^{(k)} |a_j\rangle |b_j\rangle, \tag{3.37}$$

which is in contradiction with the fact that $|w'_k\rangle$ was not entangled.

FAPP and Ontology

Many have considered decoherence as a pure technical trick devoid of any fundamental significance. The Irish physicist John Bell (1928–1990) has spoken of For All Practical Purposes (FAPP) explanations[65]: to involve the 'unknown' or at least uncontrolled environment to many seems a way to do not aggress what occurs in the local interaction in our labs. On the other hand, the idea that, globally, everything goes on *as if* the local measurement never happened gives the feeling that the latter is only illusionary, and so it appears to be not completely clear what decoherence tells us more than the MWI. So, two questions naturally arise:

- Can a kind of ontological substrate of decoherence be found? In other words, does decoherence refer to a process that can be said to be physically real?
- Is it true that what really does matter here is only the local interaction that we control in our lab? Or, said otherwise: can we dismiss the rest of the world? And if not, what is the relation between our local experience and our quantum universe?

As a matter of fact, a consistent answer to both questions (which will result to be interconnected) can be provided, but we must also be willing to introduce changes in our notion of *physical reality*. I have said that the system and the apparatus, when are entangled, share information. But decoherence tells us that they also share information with the environment. Then, we can explain the passage to a mixture by assuming that, during a local interaction, the correlation (interference) terms of the kind (3.13) between apparatus and object system become definitively lost from our local point of view so that the information that they represent is no longer locally accessible to us.[66] This would justify London and Bauer's analysis, who were puzzled by the fact that to get a mixture means to lose some information (Sect. 3.1.1). This would also explain von Neumann's starting point (Sect. 2.2.3): that to measure implies an irreversible process. In fact, local growth of entropy, due to local loss of information (and especially of correlation terms), as such means an irreversible process. In other words, we get a mixture and a final outcome every time that a part of the information is transferred into the rest of the world: it is precisely such a local transfer that maintains the same state of the whole universe (since what is involved in such a process is ultimately only a part of the universe, that is, our immediate environment, the so-called *Umwelt*). This solves the apparent contradiction noted in Sect. 2.2.2 that we appear to be free in our local choice and at the same time are

[65] Bell (1990).

[66] For a deeper understanding of this issue I recommend (Zwolak and Zurek 2013). I shall come back on these problems.

not free. In fact, the 'information download' into the environment is an uncontrolled process as the detection event itself (they are like two sides of the same coin). For dealing more deeply with this we need to deal with nature and role of the information, the object of the next section.

Decoherence Time

Let us note that, locally, the interference terms are never fully washed out.[67] It has been shown that to let quantum interference terms go to zero would require an infinite amount of time, although, as mentioned, the timescale of (approximate) decoherence is very short.[68] This can be shown for the case of a spatial superposition of two Gaussian wave functions: Zurek and co-workers proposed the definition of a decoherence time parameter τ_d as

$$\tau_d \simeq \gamma^{-1} \left(\frac{\lambda_T}{\Delta x} \right)^2 , \tag{3.38}$$

where γ is the relaxation rate, i.e. the coupling constant of the system–environment interaction, Δx is the separation between the two Gaussian peaks representing the system's superposition, and, thanks to formula (1.6), λ_T is the *de Broglie thermal wavelength*

$$\lambda_T = \frac{h}{p_x} = \frac{h}{\sqrt{2\pi m k_B T}} , \tag{3.39}$$

where T is the temperature, m the mass of the particles, and k_B is the Boltzmann constant (1.355), which, I recall, has the same dimension of entropy: $[E][\Theta^{-1}]$. This equation can be derived by taking the statistical–mechanical form of energy

$$E = k_B T, \tag{3.40}$$

or also the so-called effective kinetic energy $E_k = \pi k_B T$, where $p_x \approx \sqrt{2mE}$, in accordance with formula (1.7). For a gas of bosons at room temperature, this wave length is much smaller than the mean distance between the bosons, and the gas is typically disordered, what implies that the relaxation time is short. The result (3.38) has been explicitly derived for a particle moving in one dimension, subject to a harmonic oscillator potential, linearly interacting with an environment described by a large number of harmonic oscillators (a *reservoir* or a *bath*).[69]

[67] As shown in Zurek (1982, 1991).

[68] Wallace (2010).

[69] See Paz et al. (1993, 489–94).

Conclusions

In conclusion, they are the same local physical conditions that should allow us to (i) entangle apparatus and object system, and in this way select a specific observable, and (ii) extrapolate these two systems from the network of their world correlations in order to locally acquire information through detection. Therefore, it seems that decoherence is something more than a pure FAPP and provides a new understanding of the relation between local and global processes. In fact, quantum systems show non-local interference factors, but, when speaking of the state of the whole system that includes our object system and apparatus as a part (and this is finally the universe itself), this whole should comprehend *both* these non-local factors and the (at least potential) *local* outcomes events.

3.1.4 Solution of the Cat's Paradox

A Short-Living State

In Sects. 2.2.5 and 3.1.2 we have already considered Schrödinger's solution to the paradox of the cat, which turns out to be similar to that proposed by the MWI for measurement but even more to decoherence. In fact, the idea that the worry arises from an extrapolation can be brought in harmony with decoherence and the distinction local–global.[70] Thus, we wish to solve this paradox by making use of the concepts introduced in the previous subsection. In fact, although the formal expression (2.49) for the Schrödinger cat state is correct and we may indeed hypothesise that it describes the real physical situation (at least at a certain stage of the process), the fact remains that it is very unlikely to have a cat in animated suspension between life and death, and, what is even more important, nobody could observe it by definition. What I am saying is that this state does not represent an impossibility as such, but its extreme unlikeliness is rather a consequence of the huge complexity of the involved systems. In fact, assuming that the whole system cat-measuring apparatus is an entangled state in accordance with that formal description, it is likely that this be a very unstable, short-living state, erased immediately by decoherence at the local level, therefore making very difficult its detection. This means that, even if the cat be in fact "between" life and death at a certain time, it would never apperceive it (nor any observer) given the extremely short decoherence time (3.38) involved here, perhaps of some femtoseconds (a femtosecond is 10^{-15} s) or even less, so totally irrelevant to biological or psychological processes (and also to detection, at least for our current standards of discrimination). In other words, decoherence helps us solve such an apparent paradox in a quite natural way.

[70] For the use of the notion of *global* see also Schlosshauer (2007, Sect. 2.3).

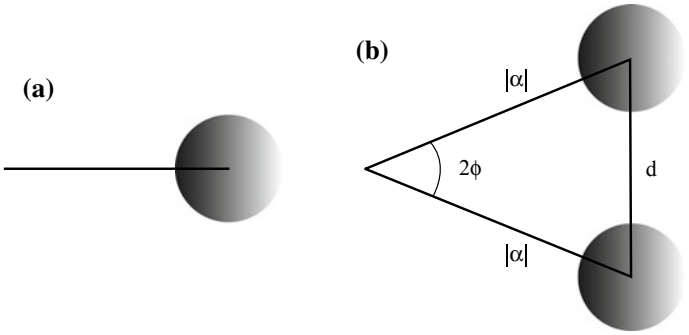


Fig. 3.3 **a** Pictorial representation in phase space of a coherent state of a quantum oscillator (Fig. 1.25, Sect. 1.2.4). **b** The two coherent-state components, when separated by a distance $d = 2\sqrt{N}\sin\phi > 1$, where $N = |\alpha|^2$ is the mean number of energy quanta or photons, form a Schrödinger cat. Adapted from Auletta et al. (2009, p. 537)

Experimental Realisations

It is worth mentioning that quantum effects at least at the mesoscopic level (that of big atoms and molecules) have been experimentally observed.[71] In the experiment performed by Monroe et al., a beryllium ion was used and states of the form

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\left(e^{i\phi}\,|e\rangle\,|\alpha\rangle + e^{-i\phi}\,|g\rangle\,|\alpha\rangle\right) \tag{3.41}$$

were produced, where $|e\rangle$ and $|g\rangle$ are its internal degrees of freedom (representing excited and ground state of energy, respectively), while $|\alpha\rangle$ and $|-\alpha\rangle$ (when $|\alpha|$ is large) are macroscopically distinguishable coherent states of the centre-of-mass motion of the ion. Coherent states will be formally introduced later on. It is interesting to note that these two different degrees of freedom, though pertaining to the same physical system (the ion), are written in a form that is completely analogous to an entanglement of two different subsystems. As a matter of fact, they pertain to two different Hilbert spaces, and this justifies the above formalism. In this context, the internal degree of freedom plays the role of the microscopic system, whereas the centre-of-mass motion the role of the macroscopic one. Intuitively, the Schrödinger's cat is produced when, in states of this form, the 'distance' between the coherent states reaches a critical threshold so that the system is somehow 'doubled' (Fig. 3.3). This is a promising field of investigation. What we need to consider is that these quantum effects become growingly less relevant when ascending to the macroscopic world, as already understood by Planck (Sect. 1.3.2). However, if there are many intermediate cases and it is a matter of relative weight of the Planck constant (together with

[71] Brune et al. (1996), Monroe et al. (1996).

other parameters like the 'distance' between states) to make the difference, there are grounds to think that there is no sharp cut between these two dimensions, as it was assumed for measurement (Sect. 2.2.3), or at the least this line appear to be movable according to the progresses in both science and technology.

Again on Complementarity

I have summarised the complementarity principle in Sect. 1.3.3 and recalled it in Sect. 2.4.2. It is worth mentioning that, in its original formulation, the complementary behaviours wave-like and particle-like were still understood as dependent on the macroscopic properties of the apparatus that we use to perform complementary experiments.[72] As recalled, this is commonly known as the Copenhagen interpretation, initially devised by Bohr and Heisenberg in the years 1924–1928. At the opposite, on the basis of the previous results, I suggest that the macroscopic bodies should be rather conceived of as *emerging* from the microscopic quantum world when certain local processes happen. It is right that a macroscopic apparatus is necessary for ascribing a specific property to a quantum system. Nevertheless, as I have stressed, these kinds of interactions should also happen spontaneously in nature, although in this case there is nobody reading, interpreting and eventually storing the result, and so, properly speaking there is an outcome of the interaction but no property ascription (on this point Bohr is obviously right). So, ontological and epistemic issues need to be kept apart as different problems, and we cannot answer issues of one kind with statements of the other kind.

Bohr's approach is quite understandable at a time in which QM was still in its infancy. Indeed, on the one hand it was not possible at that time to perform experiments using apparata that are of the same size as the measured quantum system, as it is currently the case for the many novel experiments on mesoscopic systems (although this information needs to be somehow broadcast to us). Moreover, there was no possibility to explore intermediate situations between corpuscular and undulatory behaviours as it is currently the case,[73] a problem on which we shall come back. On the other hand, quantum theory was still conceptually dependent on classical physics and therefore quantum phenomena and quantities acquired a physical meaning only in the light of the classical theory according to the correspondence principle (Sect. 1.3.2), what remains today still true but, as said, only in the sense of the mathematical limit of some processes. Therefore, it is quite normal that this kind of assumptions or simplifications were made in the early development of quantum theory.

[72] At one moment Bohr says that "our interpretation of the experimental material rests essentially upon the classical concepts" (Bohr 1928).

[73] Pioneering studies are Wootters and Zurek (1979), Mittelstaedt et al. (1987). See also De Muynck (2002).

3.1.5 Summarising

The previous explanation shows the risks of mixing physical operations with formalism as well as ontology with epistemology or of misinterpreting one or the other element of the process attributing to them characters that could clash with our experience (included our experiments). In particular, the idea of a potential conflict between detection events and quantum-mechanical laws seems to be rooted in two misunderstandings:

- It was not considered that laws could not determine single local events but only the reversible dynamics of the whole (see also Sect. 2.4.2). The philosopher C. Peirce was the first to propose that laws do not apply to "single atoms".[74]
- It was underestimated that the world is built not only by global laws but also by (some initial or final) local events that apparently are irreducible to those laws although no violation of them occurs (otherwise the theory would be rejected on the basis of some facts).

Therefore, we could solve the puzzle of measurement but by advancing along two directions:

- We need to explore the implications of the dichotomy between global reversible mappings and local events. Is this dichotomy such that in our world a dangerous gap does exist? We shall see that dynamics is in fact the entrenchment of these two dimensions. This has been already hinted at by saying that the interference terms are never fully washed out.
- We need to refine our analysis of an ontological substrate of decoherence. The next section is devoted to this examination.

3.2 Formalism and Reality: Information

3.2.1 Objectivism

The solution to the measurement problem so far advanced relies on three of Schrödinger's ideas:

1. All quantum systems of the universe are correlated.
2. What counts for the entire network counts also for each of its parts, so that no local operation can destroy the quantum character of microphysical systems.
3. The paradox of measurement disappears when we are aware that it consists in a local extrapolation (of information).

These ideas are grounded on the postulate that the quantum formalism perfectly describes the microscopic world (and therefore it is not simply a tool for making

[74] Peirce (1891).

predictions), so that it is reputed to be the only guide allowing us to discriminate between what is real and what is not. In this context, it is interesting to report what the historian of science Jagdish Mehra and H. Rechenberg observe[75]:

> Planck, von Laue, and Schrödinger certainly did not adhere to old classical theories; they did not wish to renounce any of the achievements of the modern relativity and quantum theories, but only complained about the Copenhagen interpretation of QM and proposed to retain more '*Objektivierkeit*' (objectifiability) in the sense accepted since centuries by scientists in many different fields.

Thus, let us call this standpoint *quantum objectivism*, and let us consider it a weaker form of the objectivist interpretation (of measurement) so far discussed (Sect. 3.1.2). A general form of physical objectivism, according to which the formalism of physics need to describe reality, is a standpoint that Schrödinger shared with Einstein, although the issue of separability will also mark their distance from 1935 on.[76]

Thus, the previous examination leads us to the conclusion that an objectivist standpoint seems unavoidable for QM, and quantum objectivism will be proved to be a crucial ingredient of all the solutions that have been provided to quantum-mechanical puzzles. The further examination of EPR's standpoint will tell us more on this point.

The physicist who deals with statistical mechanics or hydrodynamics or also material structures will find quantum objectivism not so understandable, because, although he/she knows that also in other fields studied by physics there are quantum effects (as displayed by the Schrödinger's cat), he/she tends to consider these as marginal aspects of his/her field of studies and thus will judge an objectivist point of view bizarre. There is here an important standpoint that cannot be denied and will be explored later. In short, what I shall suggest is that QM needs to be a sufficiently robust theory for accounting for the emergence of realities that cannot be fully described by the quantum formalism only: it could provide a general framework but it does not need to deal with the details of other fields of investigation.

3.2.2 A Formalism That Denotes a Reality?

Information Objective or Subjective?

The central ontological question could be formulated in these terms: are there specific formal aspects expressed by the mathematics of QM that we should assume to be real? Let us come back to Wigner's interpretation. It states a fundamental difference between a formal tool like the wave function and the real psychological experience that we have when we observe. Thus, as recalled, it provides a negative answer to

[75] Mehra and Rechenberg (1982, VI, p. 709). In fact, Heisenberg explicitly criticised this point: see also Heisenberg (1969, Chap. 6); Mehra and Rechenberg (1982, VI, p. 738 and ff.).

[76] In general, Einstein was quite sympathetic with the views of Schrödinger (Home and Whitaker 2007, pp. 33–34, 86), at least until 1935.

our ontological question. But what means "to observe" or "to measure"? It means to acquire information that we did not possess beforehand; otherwise, it makes no sense to observe (Sect. 2.2.2). In fact, as recalled in Sect. 3.1.1, the motivation for the subjectivist interpretation was the possibility to solve the problem of measurement in terms of a passage from lack of knowledge to knowledge. On the other hand, in Sect. 3.1.3, we have seen that there are good reasons to think that the final detection outcome is the result of a process of local displacement of information. Then, we could try to understand measurement as a difference between information nested somewhere and the piece of information that we acquire after detection.[77] The commonality with the subjectivist interpretation ends here: in fact, we wish now to disentangle the issue of the information acquisition from any subjective consideration.

In other words, I am trying to assign ontological status to information, i.e. treating it as an *objective* reality that as such is not dependent on subjective operations or conditions, although the latter could well display some common character with the former. However, the postulated objectivity of information raises immediate problem: information is a pure formal quantity, as it is evident by its dimensionless character (it cannot be expressed in terms of any physical observable or variable). How can a formal quantity play a role that has real, physical, effects? It appears that something non-physical could exert some kind of influence on something physical, in violation of Leibniz's principle of the causal closure (of the physical world): Sect. 2.3.3. In this context, I note that dimensionless physical quantities are not unknown in physics.[78] In fact, we have the *Planck time*

$$t_{\mathrm{P}} = \sqrt{\frac{\hbar G}{c^5}} \approx 5.391,160 \times 10^{-44}\,\mathrm{s}, \tag{3.42a}$$

where $G = 6.67191(99) \times 10^{-11}\,\mathrm{m^3 kg^{-1} s^{-2}}$ is Newton's gravitational constant (whose physical dimensions are $[\mathrm{L}^3][\mathrm{M}^{-1}][\mathrm{T}^{-2}]$), which is the time required for light to travel, in a vacuum, a distance of 1 *Planck length*:

$$\ell_{\mathrm{P}} = \sqrt{\frac{\hbar G}{c^3}} \approx 1.616,199 \times 10^{-35}\,\mathrm{m}. \tag{3.42b}$$

In fact, $\ell_P/t_P = c$. The Planck length is about 10^{-20} times the diameter of a proton. We have also a Planck charge, a Planck temperature, and a Planck mass (where all of these Planck units are defined by means of other natural constants)

$$m_{\mathrm{P}} = \sqrt{\frac{\hbar c}{G}} \approx 1.220910 \times 1019\ \mathrm{GeV/c^2} = 2.176470(51) \times 10^{-8}\,\mathrm{kg}. \tag{3.42c}$$

The Planck mass is Nature's maximum allowed mass for point masses (quanta), i.e. a mass capable of holding a single elementary charge. If two quanta of the Planck mass

[77] For deepening the subject see Lindblad (1973, 1983).

[78] Tegmark et al. (2006).

(or greater) met, they should spontaneously form a black hole whose Schwarzschild radius, after the name of the German physicist Karl Schwarzschild (1873–1916), equals their Compton wavelength (1.8). The Schwarzschild radius is the radius of a sphere such that, when all the mass of an object be compressed within that sphere, the escape velocity from its surface would equal the speed of light.[79]

Of course, if information is a dimensionless quantity like those just introduced, it must show a connection with other physical quantities. We shall see that this is in fact the case.

Entropies

In order to deal with this riddle, let us recall that, when two particles are entangled, they are assumed to share information. In fact, there is a specific quantity for describing this: *mutual information*. The expression of entanglement in terms of mutual information has several advantages: as I shall argue, this formal notion covers also the classical case but makes simultaneously evident which are the differences between quantum and classical correlations. To this purpose, let us first introduce the notion of entropy. The *entropy* of a variable or a set X of signals denotes the incertitudes of the values or elements of X and is given by[80]

$$S(X) := -\sum_x \wp_x \lg \wp_x, \tag{3.43}$$

where I have used binary logarithms (Fig. 3.4) and $x \in X$. Note that we call *surprisal* a single component of the many representing the entropy (3.43) of the system, e.g. $-\lg \wp_x$, so that entropy is finally a weighted sum or the expectation value of information. Entropy can be understood as the information that would be necessary in order to describe the system, and therefore, it immediately means *disorder* (since it is required a larger amount of information for describing a disordered system than an ordered one). This entropy is called the informational or *Shannon*'s entropy, after the name of the American mathematician Claude Shannon (1916–2001), the father of information theory,[81] and it is related to the thermodynamical Gibbs' or Boltzmann's entropies.[82] There is also a quantum-mechanical counterpart.[83] Note in fact that the quantum (*von Neumann*) entropy of a quantum system described by the density matrix $\hat{\rho}$ is given by

[79] Serjeant (2010, Chap. 6).

[80] For a short summary on classical information theory see Khinchin (1957). On quantum entropy see the rather technical book (Ohya and Petz 1993).

[81] Shannon (1948).

[82] For this thermodynamical quantities the reader may consult a textbook like (Huang 1963).

[83] Attempts at falsifying Shannon's measure of information in its application to QM have proved to be not very successful so far. For a relative recent approach in this sense see Brukner and Zeilinger (1999, 2001, 2005) and relative literature: (Mana 2004; Shafiee et al. 2006).

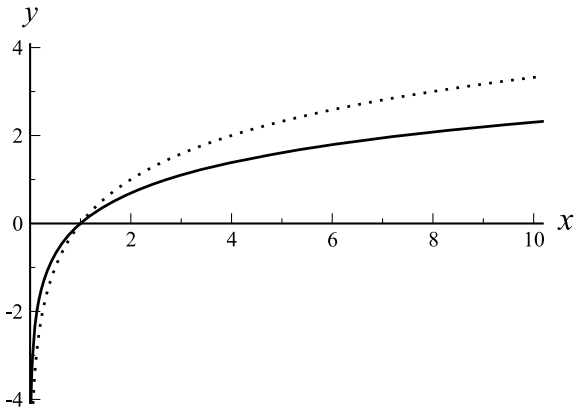


Fig. 3.4 Plot of the natural logarithm $\ln x$ (solid line) and the binary logarithm $\lg x$ (dotted line). Note that binary and natural logarithms are related by the equation $\ln x = \lg x \ln 2$ (while the natural ln and decimal logarithms log are related by $\log x = \ln x / \ln 10$). Comparing with the plot of the exponential function in Fig. 1.7, Sect. 1.2, we see that exponential increase is very fast while the logarithmic increase is very slow. Adapted from Auletta and Wang (2014, p. 274)

$$S_{\mathrm{VN}}(\hat{\rho}) = -\mathrm{Tr}(\hat{\rho} \ln \hat{\rho}) \ , \tag{3.44}$$

where ln are natural logarithms, the inverse functions relative to exponentials (i.e. $\ln(e) = 1$). A pure state is the most ordered system in nature (due to his high symmetry) and thus has zero entropy.[84] Indeed, the diagonal of a matrix describing a pure state, being a projector in its eigenbasis, shows a single element on the diagonal that is different from 0 and this is equal to 1 (see, e.g. Eqs. (1.84) and Sect. 1.4.1), and we have that $\ln(1) = 0$. In other words, when the system is so ordered there is no possible endogenous 'choice' but it covers any possible outcome: an interaction with other systems was necessary to this purpose. Note, however, that the entropy of a pure state being zero means that the state of the system is both *maximally ordered* (thanks to quantum correlations) and *maximally random* (relative to the possible outcomes). It is ordered randomness. A crucial point is that any unitary transformation preserves the entropy, since it preserves not only the scalar product but also the trace, as displayed in Eq. (1.372). This means that, if the initial state of a system has zero entropy, under unitary evolution it will always have zero entropy. It is a map from a global state to a global state (with no events occurring during this process).

Thus, London and Bauer's analysis was correct: since a mixed state shows positive entropy, the transition from a pure state to a mixture means increase in entropy and therefore local loss of information, as predicted by decoherence. In particular, both the von Neumann and the Shannon entropies reach their maximum value when they are equal to the logarithm of the number of dimensions of the system (i.e. the number of the elements of the set to which probabilities are assigned which are dimensions of

[84] Wehrl (1978). This is an extensive paper on the subject.

the Hilbert space for a quantum system). In the case of the Shannon entropy this can be proved very easily (by considering that it reaches a maximum when all possible alternatives are equiprobable):

$$S(X)^{\max} = -\sum_n \frac{1}{n} \lg \frac{1}{n} = -\lg \frac{1}{n} = \lg n \,, \tag{3.45}$$

since $\sum_n 1/n = 1$. It is also easy to see that we have the same result for the von Neumann entropy when the state of the system is the maximally mixed (1.368). Obviously, different degrees of mixing are correlated with different degrees of entropy.

Note that when the quantum states or components are orthogonal, the von Neumann and Shannon entropies coincide. For the time being we do not need to distinguish between Shannon's and von Neumann's entropies. For this reason, I shall develop considerations that are valid both for the classical and quantum case and come back to such a distinction when necessary.

Mutual Information

Let us now consider the difference between pure states and mixtures in detail. A mathematical expression for mutual information I, when considering sets of signals (or variables) X and Y, is the following

$$I(X : Y) = S(X) + S(Y) - S(X, Y), \tag{3.46}$$

where $S(X)$ and $S(Y)$ are the entropies of X and Y, respectively, and $S(X, Y)$ is the total joint entropy of the two systems, that is obviously given by (see also Eq. (2.23c))

$$S(X, Y) = S(X) + S(Y) - I(X : Y). \tag{3.47}$$

The reason for the last expression is quite clear. The more two systems described by X and Y are separated (independent), the less is their mutual information and the more their whole state is disordered, so that their joint entropy reaches its maximal value precisely when it is a sum of the entropy of the two systems taken as independent, so that we have in general

$$S(X, Y) \leq S(X) + S(Y). \tag{3.48}$$

This is called the *subadditivity* property of entropy.[85] In other words, mutual information takes its minimum value (= 0) when we have $S(X, Y) = S(X) + S(Y)$, and thus can be understood as equal to the negative of the joint entropy when the 'partial entropies' have been subtracted or also to the negative value of the whole entropy of

[85] Lieb (1975).

the system minus its disorder. Thus, Eq. (3.46) could be also written

$$I(X:Y) = -[S(X,Y) - (S(X) + S(Y))]. \tag{3.49}$$

In other words, it is mutual information to confer order to a compound system, so that its reduction is related to growth of entropy. Note that the complementary quantity of mutual information is called the information distance:

$$\delta(X,Y) = S(X,Y) - I(X:Y), \text{ with } \delta(X,Y) + \delta(Y,Z) \geq \delta(X,Z), \tag{3.50}$$

$\forall X, Y, Z$, which is positive definite and symmetric. Another character of mutual information is that it is as such independent of any signal sending or receiving or local interaction. This is ubiquitous in nature, also in the classical case. For instance, if, for instance, by using English, we interact, through some network, with somebody whom we never met personally before, this denotes that we already shared this piece of information (and likely also many others) *before exchanging the messages*. In other words, we can share information also in absence of such a local exchange.[86] Obviously, we can classically think that the information has been previously broadcast to both parties (through ordinary communication channels), for instance, both partners have learnt English through means that ultimately go back to a common source; but we shall see that this is not necessary in QM. To see this, I focus on entanglement.

Entanglement as Mutual Information

We can formalise entanglement in a way that is parallel to Eq. (3.46) or (3.49). In fact, the entanglement between systems 1 and 2 can be defined as

$$\begin{aligned} \mathrm{E}(1,2) &= -[S(1,2) - (S(1) + S(2))] \\ &= S(1) + S(2) - S(1,2)\ , \end{aligned} \tag{3.51}$$

where $S(1,2) = S(\hat{\rho}_{12})$ is the joint (total) von Neumann entropy of systems 1 and 2 (where here and in the following, I drop the subscript VN for simplicity of notation when no confusion arises) described by the density matrix $\hat{\rho}_{12}$, while

$$S(1) = S(\hat{\varrho}_1) \text{ and } S(2) = S(\hat{\varrho}_2) \tag{3.52}$$

are the entropies calculated on the reduced density matrices of the subsystems 1 and 2, respectively, relative to the compound system in state $\hat{\rho}_{12}$, which are computed by performing the partial trace of the kind (1.395) or (3.29). Now, it is evident that entanglement shows degrees, what fits very well with the assumption that all systems of our universe are more or less, directly or indirectly, entangled. In fact, taking, e.g.

[86] For some additional considerations see Auletta (2011a, Chap. 2).

the state (1.392), it is evident that, starting with a state of maximal entanglement, by rotating the state vector $|h\rangle$ we can let it finally coincide with vector $|v\rangle$, in which case we get a product state. Clearly, there are many intermediate states between these two extremes. This means that entanglement has a measure,[87] as already anticipated in the previous section. Since mutual information shows degrees, it is plausible that systems are more or less entangled in relation to whether they share more or less information.

Entropies and Entanglement of Observables

Note that the previous expressions deal with the states (density matrices) of the involved subsystems. Now, it can be shown that the entropy of the state represents an upper bound on the entropy computed on whatever observables we measure on the involved systems—this is very helpful for choosing the observables that optimally express the entanglement between subsystems.[88] In this case, in analogy with Eqs. (3.44) and (1.370), we can write the actual von Neumann entropies of observables $\hat{O}_1$ and $\hat{O}_2$ of the system 1 and 2, respectively, as

$$S(\hat{O}_1) = -\sum_j \langle j|\hat{\rho}_1|j\rangle \ln\langle j|\hat{\rho}_1|j\rangle \tag{3.53a}$$

$$S(\hat{O}_2) = -\sum_k \langle k|\hat{\rho}_2|k\rangle \ln\langle k|\hat{\rho}_2|k\rangle \ , \tag{3.53b}$$

where $\{|j\rangle\}$ ($\{|k\rangle\}$) are the eigenstates of the observable $\hat{O}_1$ ($\hat{O}_2$), i.e. the states of the system 1 (2) that will be the output states when measuring the observable $\hat{O}_1$ ($\hat{O}_2$). I recall that the expression $\langle j|\hat{\rho}_1|j\rangle$ in the previous equation can be considered as a mean value on the density matrix $\hat{\rho}_1$, and similarly for system 2. The joint entropy of the two quantum observables is given by

$$S(\hat{O}_1, \hat{O}_2) = -\sum_j \sum_k \langle j, k|\hat{\rho}|k, j\rangle \ln\langle j, k|\hat{\rho}|k, j\rangle \ , \tag{3.54}$$

where $\hat{\rho}$ is the density matrix of the global quantum system, and I have used the simplified notation $|j, k\rangle = |j\rangle \otimes |k\rangle$. Then, we may define the entanglement between these two observables as

$$\mathrm{E}(\hat{O}_1, \hat{O}_2) = S(\hat{O}_1) + S(\hat{O}_2) - S(\hat{O}_1, \hat{O}_2) \ , \tag{3.55}$$

which is formally similar to Eq. (3.51). Now, it can be proved that we always have

$$\mathrm{E}(\hat{O}_1, \hat{O}_2) \leq \mathrm{E}(1, 2), \tag{3.56}$$

[87] For a measure of entanglement see Vedral et al. (1997).

[88] Barnett and Phoenix (1989). See also Auletta et al. (2009, Sect. 17.2).

i.e. the information contained in the correlation between any two observables pertaining to the two subsystems—given by Eq. (3.55)—cannot exceed the total information content of the correlation between the two systems—given by Eq. (3.51). Of course, this is a consequence of the non-commutativity of observables that hinders the acquisition of the whole information potentially contained in the state, according to what said in Sect. 2.2.2. Moreover, it can be also proved[89] that, if the state of the compound system is pure, then

$$\mathrm{E}(\hat{O}_1, \hat{O}_2) \leq \frac{1}{2}\mathrm{E}(1,2)\ . \tag{3.57}$$

In conclusion, we can say that, with this formalism (and some other aspects that will be considered in the next chapter), we are able to fully describe entanglement.

Uncertainty Relations in Informational Terms

If observables are related to information, we can wonder whether other connections can be found. In fact, it is possible to use the von Neumann entropy for expressing uncertainty relations (Sect. 1.2.4). There is a wide literature on that but I shall focus on Deutsch's derivation.[90] Let us consider two observables

$$\hat{O} = \sum_o o\,|o\rangle\,\langle o| \quad \text{and} \quad \hat{O}' = \sum_{o'} o'\,\big|o'\big\rangle\big\langle o'\big|\,, \tag{3.58}$$

describing the same system. Let us define a function $f(\hat{O}, \hat{O}')$ that represents the maximum of the modulus of the overlap of (in general not orthogonal) eigenvectors $|o\rangle\,, \big|o'\big\rangle$:

$$f(\hat{O}, \hat{O}') = \max_{o,o'} |\,\big\langle o \,|\, o'\big\rangle\,|. \tag{3.59}$$

Suppose a quantum system prepared in the state $|\psi\rangle$ and let $\wp(o), \wp'(o')$ the probability distributions associated with measurement of $\hat{O}, \hat{O}'$, respectively, and their respective von Neumann entropies $S(\hat{O}), S(\hat{O}')$. It is clear that, in agreement with Eqs. (3.53), we have

$$\begin{aligned} S(\hat{O}) + S(\hat{O}') &= -\sum_o \wp(o)\lg(\wp(o)) - \sum_{o'} \wp(o')\lg(\wp(o')) \\ &= -\sum_o \sum_{o'} \wp(o)\wp'(o')\lg(\wp(o)\wp'(o')), \end{aligned} \tag{3.60}$$

due to the logarithm property $\lg(xy) = \lg x + \lg y$. Our aim is to bound the product of the probabilities (see Eq. (1.101))

[89] See Barnett and Phoenix (1991) for details on these proofs.

[90] Deutsch (1983). See also Nielsen and Chuang (2000, p. 503). For an overview of other proposals see Auletta (2000, Sect. 42.2).

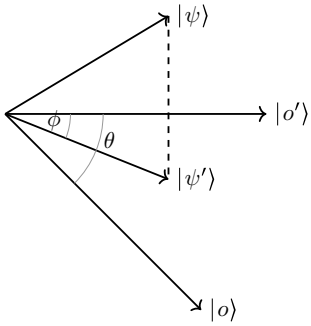


Fig. 3.5 The vector $|\psi'\rangle$ makes an angle $\theta - \phi$ with vector $|o\rangle$ and an angle ϕ with vector $|o'\rangle$. These three vectors are on the same plane

$$\wp(o)\wp'(o') = |\langle o \mid \psi\rangle|^2 \cdot \left|\langle o' \mid \psi\rangle\right|^2 \tag{3.61}$$

from above. For doing this, let $|\psi'\rangle$ be the projection of $|\psi\rangle$ into the plane spanned by $|o\rangle, |o'\rangle$, what implies that $|\psi'\rangle$ has a norm $\lambda \leq 1$, as displayed in Fig. 3.5, i.e. $|\psi'\rangle = \lambda\,|\psi\rangle$. The vector $|o'\rangle$ makes and angle θ with vector $|o\rangle$ while vector $|\psi'\rangle$ makes an angle ϕ with vector $|o'\rangle$. Then, pure geometrical considerations tell us that we have

$$\wp(o)\wp'(o') = \left|\langle o \mid \psi'\rangle\right|^2 \left|\langle o' \mid \psi'\rangle\right|^2 = \frac{1}{\lambda^2}\cos^2(\theta - \phi)\cos^2(\phi). \tag{3.62}$$

This distribution reaches its maximum when $\lambda = 1$ and $\phi = \theta/2$, which implies $\wp(o)\wp'(o') = \cos^4(\theta/2)$. Since θ expresses the separation between $|o\rangle$ and $|o'\rangle$ (i.e. $\cos\theta = \langle o \mid o'\rangle$), this in turn gives, in a loose analogy with Eq. (2.33a),

$$\wp(o)\wp'(o') = \left(\frac{1 + |\langle o \mid o'\rangle|}{2}\right)^2, \tag{3.63}$$

where I have used the trigonometric formula

$$\cos^2(x) = \frac{1 + \cos(2x)}{2}. \tag{3.64}$$

The above result, together with Eqs. (3.59) and (3.60), gives

$$\begin{aligned} S(\hat{O}) + S(\hat{O}') &= -\sum_{o,o'}\left(\frac{1 + |\langle o \mid o'\rangle|}{2}\right)^2 \lg\left(\frac{1 + |\langle o \mid o'\rangle|}{2}\right)^2 \\ &\geq -2\lg\frac{1 + f(\hat{O}, \hat{O}')}{2}, \end{aligned} \tag{3.65}$$

since $\lg x^2 = 2\lg x$.

Measurement in Informational Terms

Let us now apply this formalism to the problem of all problems: measurement. I have already anticipated that measurement can be considered in terms of local changes in accessible information accompanied by local increase (and consequent shift) in entropy. Let us now see how this can work. According to inequality (3.48), when correlations go down, as during measurement, the local entropy should grow. Let us consider[91] the initial density matrix $\hat{\rho}_{\mathcal{SAE}}(t_0)$ for the system $\mathcal{S}$ plus the apparatus $\mathcal{A}$ plus the environment $\mathcal{E}$, with $\mathcal{A}$ and $\mathcal{S}$ initially uncoupled with the environment:

$$\hat{\rho}_{\mathcal{SAE}}(t_0) = \hat{\rho}_{\mathcal{SA}}(t_0)\hat{\rho}_{\mathcal{E}}(t_0) \ , \tag{3.66}$$

in agreement with the second line of Eq. (3.27). A further property of both Shannon and von Neumann entropies is the strong subadditivity property: for arbitrary systems 1, 2, 3, we have

$$S(\hat{\rho}_{123}) + S(\hat{\rho}_2) \leq S(\hat{\rho}_{12}) + S(\hat{\rho}_{23}) \ , \tag{3.67}$$

where also here I have dropped the VN subscript for the sake of notation. This inequality is a generalisation to three systems of the (3.48). Thus, at any later time t we must have

$$S_t(\mathcal{S}, \mathcal{A}, \mathcal{E}) + S_t(\mathcal{A}) \leq S_t(\mathcal{S}, \mathcal{A}) + S_t(\mathcal{A}, \mathcal{E}) \ , \tag{3.68}$$

where $S_t(\mathcal{S}, \mathcal{A}, \mathcal{E})$ is the entropy of the whole system at time t, and similarly for the other expressions. Since the evolution of the density matrix of the total system is unitary, it follows that the total entropy from an initial time t_0 to a final time t is conserved:

$$S_{t_0}(\mathcal{S}, \mathcal{A}, \mathcal{E}) = S_t(\mathcal{S}, \mathcal{A}, \mathcal{E}) \ . \tag{3.69}$$

Since the initial state at time t_0 is by hypothesis such that both systems $\mathcal{S}$ and $\mathcal{A}$ are uncoupled with the environment $\mathcal{E}$, we have

$$S_{t_0}(\mathcal{S}, \mathcal{A}, \mathcal{E}) = S_{t_0}(\mathcal{S}, \mathcal{A}) + S_{t_0}(\mathcal{E}) \ . \tag{3.70}$$

Suppose now that $\mathcal{S}$ is a spectator while $\mathcal{A}$ and $\mathcal{E}$ interact at time t. Thus we have

$$S_t(\mathcal{S}) = S_{t_0}(\mathcal{S}) \tag{3.71a}$$

and, as a consequence of this equation together with Eqs. (3.69)–(3.70), we also have

$$S_t(\mathcal{A}, \mathcal{E}) = S_{t_0}(\mathcal{A}, \mathcal{E}) \ . \tag{3.71b}$$

[91] I synthesise here the results of Partovi (1989). See also Auletta et al. (2009, Sect. 17.3).

On the other hand, the lack of entanglement or of correlation between $\mathcal{A}$ and $\mathcal{E}$ at t_0 can be expressed as

$$S_{t_0}(\mathcal{A}, \mathcal{E}) = S_{t_0}(\mathcal{A}) + S_{t_0}(\mathcal{E}) \ . \tag{3.72}$$

Combining all these relations we show now that it is possible to derive

$$\mathrm{E}_t(\mathcal{S}, \mathcal{A}) \leq \mathrm{E}_{t_0}(\mathcal{S}, \mathcal{A}) \ , \tag{3.73}$$

where $\mathrm{E}(\mathcal{S}, \mathcal{A})$, in agreement with Eq. (3.51), expresses the complex of correlations between $\mathcal{S}$ and $\mathcal{A}$. To this purpose, let us add $S_t(\mathcal{S})$ to both sides of Eq. (3.68) and rewrite it as

$$S_t(\mathcal{S}) + S_t(\mathcal{A}) - S_t(\mathcal{S}, \mathcal{A}) \leq S_t(\mathcal{S}) + S_t(\mathcal{A}, \mathcal{E}) - S_t(\mathcal{S}, \mathcal{A}, \mathcal{E}) \ . \tag{3.74}$$

By using Eqs. (3.69)–(3.72), we obtain

$$S_t(\mathcal{S}) + S_t(\mathcal{A}) - S_t(\mathcal{S}, \mathcal{A}) \leq S_{t_0}(\mathcal{S}) + S_{t_0}(\mathcal{A}) - S_{t_0}(\mathcal{S}, \mathcal{A}) \ , \tag{3.75}$$

which, thanks to Eq. (3.51), is indeed Eq. (3.73) in explicit form. In other words, from a measurement we expect a local decrease in the quantum correlations between the system and the apparatus. As a consequence, we have formally proved that there can be a local increase in entropy and a swap of information from the local context (determined by the experiment's protocol) to the environment, in agreement with the results of Sect. 3.1.3.

Araki–Lieb Inequality

Note that, when systems are entangled and in a pure state, the fact that the whole has zero entropy does not imply that the two subsystems have also zero entropy. This means that, although the whole state is maximally ordered, the subsystems can be disordered, in agreement with decoherence. In fact, they can have some positive entropy, the only requirement being that this entropy must be equal. This is proved by the fact that any entropy (whether classical or quantum-mechanical) of systems 1 and 2 needs to satisfy the *Araki–Lieb inequality*, after the names of the Japanese mathematical physicist and mathematician Huzihiro Araki and the American mathematical physicist Elliott Lieb:

$$|S(1) - S(2)| \leq S(12) \leq S(1) + S(2), \tag{3.76}$$

where $S(12)$ is the entropy of the compound system.[92] We have already derived the last part of the inequality (see Eq. (3.48)), i.e.

[92] Araki and Lieb (1970). See also Wehrl (1978).

$$S(12) \leq S(1) + S(2). \tag{3.77}$$

The first part

$$|S(1) - S(2)| \leq S(12), \tag{3.78}$$

where the modulus is due to the fact that we do not know which of the two marginal entropies (i.e. the entropies computed on the reduced or marginal density matrices) is larger, tells us that, if $S(12) = 0$, then we must have $S(1) = S(2)$. This inequality is of great importance for understanding the mechanisms of decoherence. In particular, it tells us that, by partitioning a large environment that could have zero entropy in parts that locally possess positive entropy, these parts must possess the same degree of order and disorder. This means, that shifts in entropy and relative changes in information accessibility should be local, as already anticipated, and even happen at a quite *small scale*, in accordance with the very short timescale of decoherence.

Irreversible Processes

When we deal with irreversible processes we necessarily have an increase in entropy also in terms of the *Boltzmann entropy*

$$S_B = k_{\mathrm{B}} N \ln W, \tag{3.79}$$

which has dimension $[\mathrm{E}][\Theta^{-1}]$ and where I recall that k_{B} is the Boltzmann constant (1.355), N is the number of particles (paradigmatically, molecules of a gas), and W is the number of the possible configurations of the system. Often it is said that entropy is not a fundamental property but it is a kind of macroscopic–surface property averaging on systems as described by thermodynamics. As we shall discover, this is not true and there is a pure statistical–mechanical notion of entropy that is more fundamental than the thermodynamical one.[93]

Nevertheless, as seen, quantum information also covers the case of maximal order (when the von Neumann entropy is zero). Thus, quantum information can be considered to be independent of the statistical character of entropy, and this is precisely what constitutes its objectivity. In fact, we shall deal with single quanta of information.

What Causal Role Does Information Play?

We have seen that such a formalism is capable to deal with all issues of quantum theory and even adds something to our understanding of the previous problems. Nevertheless, the crucial questions have not yet been raised and are the following:

[93] Although there is often a difference between Boltzmann and Shannon entropy related to the degrees of freedom (which in general are higher in the former case), when we deal with elementary particles they converge (Bekenstein 2003).

what is the additional value of information relative to the ordinary tools of QM? How is information sharing causally effective in the case of quantum systems? Let us first deal with the last problem and consider the following example: suppose that the state of two particles is the singlet state (2.54). In such a case, they show a spin correlation such that when one of the two particles is found to be in a spin-up state along an experimentally chosen direction, the other one will be necessarily in a spin-down state along the same direction or vice versa. In other words, we expect to obtain either up-down or down-up but never up–up or down–down, in accordance with what Eq. (2.54) tells us. Now, the crucial point is that, if the world consisted of random events only (Sect. 2.4.2) and therefore were maximally disordered, we would expect to obtain any one of these four possible outcomes with equiprobability, according to the discussion of Eq. (3.45). The fact that we can obtain only two (either up–down or down–up) out of four cases (and not necessarily with equiprobability) represents a reduction of the space of possible events, as anticipated in Sect. 3.1.1. In other words, quantum-mechanical correlations represents *constraints* limiting the space of the events that we can obtain.[94] Note that such a restriction is a restriction of the space of the *joint* events. In fact, from the local "point of view" of the single subsystems all outcomes (for instance, both up and down for both entangled systems) are still possible and their probability remains unchanged (i.e. the probabilities computed when considering the system as separated and the probabilities computed by summing over the other system, i.e. the *marginal* probabilities, are the same), so that this result remains untouched by such a restriction. However, since it represents a restriction of the space of the joint events, entanglement is a source of order, in agreement with our definition in relation with entropy, and thus the two entangled systems must share information.

Now, if information sharing can limit the space of the possible events by connecting them (thus blocking certain possible joint outcomes), then, it shows to have a causal influence on the results that we can obtain through joint measurements, and consequently must be real, because nothing that is not real can have an influence whatsoever on reality, as clearly stated by Schrödinger himself, in agreement with his objectivism.[95] Therefore, we are led to the conclusion that this formal reality, information, must be somehow real. In this way, we start to answer the question ("how can we ascribe objective reality to quantum systems?") raised in Sect. 2.1.1 as well as the question ("what kind of reality do correlation terms represent?") raised in Sect. 2.1.2. However, if so, we still need to show that all basic requirements that we expect from something being real and having causal influence can be satisfied.

I finally note that the physicist who first took information to be the basic form of quantum reality is John Wheeler, who spoke of *it from bit*. Actually, not by chance, Wheeler seemed to understand this still on the outline of what can be called Bohr's (and in part Wigner's) epistemic phenomenism: as meaning that being any information acquisition by its own nature dual or binary (since it can be expressed with a yes–no question), then the reality of a quantum system is dependent on this binary

[94] As recalled in Auletta (2011c, Sect. 3.2.1).

[95] Quoted in Shimony (1965, p. 317). Shimony defines Schrödinger as a realist.

choice that we experimentally impose on the system (Sects. 2.4.3 and 3.1.1).[96] This means that information, according to Wheeler, depends on information acquisition. At the opposite, I shall show here that information makes perfect sense also independently of its acquisition.

3.2.3 Dealing with Information

Possible Objections

When trying to attribute ontological status to information independently of information acquisition, some possible objections are[97]: Whose information? Information about What? These are obviously very important questions when the issue is the *use* that we humans make of information, and in fact, the two questions are addressed to the possible users and the possible referents of information. However, things can be considered in totally different terms when dealing with an ontological significance of information at a physical level. If quantum systems are considered to essentially or primarily represent information, there could be no subject that makes use of this information as far as its objectivity is concerned and no reference about which this information would be. Thus, I am suggesting that quantum-mechanical systems simply *are* the information that they represent, and therefore such an information is deprived of any functional or intentional meaning but has a pure *combinatorial* character, and, since the possible events are not independent, it consists in the relation among possible events that have some objective probability to occur.

It was Zeilinger who gave an important hint into this direction (a point of view that can appear opposite to his previous standpoint, as reported in Sect. 2.4.2): in the context of QM, he defined information as the connection among possible events (and it is clear that both superposition and entanglement satisfy this criterion).[98] Although there is a clear connection with his previous standpoint (information deals ultimately with detection events), now the stress is not on the events already *happened* (i.e. the information that we in fact acquire through observation) but on the *possible* outcomes (i.e. on the relations among these events before they happen or even independently of their happening). Evidence for this definition of information can be seen in the fact that entropy is defined as a combination of probabilities, and these probabilities may be not independent, as it is the case for quantum-mechanical systems.

Note that the objectivity of information ultimately relies on the objectivity of probability, about which I have already said something in Sect. 2.1.2. This is evident when looking at the form of quantum entropy (3.44), but the same is true for classical

[96] Wheeler (1990).

[97] They are formulated in Bell (1990).

[98] A personal communication reported in Auletta (2011a, p. 38).

entropy. And to a certain extent vice versa: as remarked by Planck, "entropy …is a measure of a physical probability".[99]

Let me stress that we should only avoid the incorrect idea that information is some physical stuff that goes from the system to the observer.[100] As pointed out also by the philosopher of science C. Timpson, a 'received' piece of information is not part of the original message and neither does it contain some transmitted portion of it. No part of the output is systematically related to any part of the input; their various parts are only probabilistically (but not necessarily statistically) related (Sect. 2.1.2). Rather, one ends up with something else entirely. Thus, the output does not count as a piece, portion or part of transmitted information at all. Also in the quantum teleportation protocol (which will be discussed in detail in the next section), the transmission of a piece of information from A to B will consist in the production at B of another token of the type produced at A, where the production at B is consequent on the token's being produced at A.[101]

Freely Evolving Systems

Suppose that a quantum system with which we interacted in the past is subsequently let free (i.e. uncontrolled) for a certain amount of time. What does it happen? The system is no longer confined and will undergo any kind of evolution according to the Schrödinger's and Heisenberg's pictures (Sect. 2.1.1), what means that both the state vector and the expansion axes can rotate in any direction and with whatever angle, since all of these transformations are unitary. In other words, the system could be now in *any* state from which it could evolve through unitary evolution. In fact, the state of a quantum system is described by two parameters: the *amplitude* θ (whose square modulus give us the probability to get one of its components when measuring) and the *relative phase* ϕ among these components (see also Fig. 1.1, Sect. 1.1.2). The phase is an integral part of the description of the state of the system, and therefore, the time evolution operator of the system needs to account for both shifts in amplitude and phase. Let us consider this formally. Since the observable is not predefinite at this stage, let us employ a rather abstract informational basis represented by $|0\rangle$ and $|1\rangle$ (see also Eqs. (1.403)), which could be considered the counterparts of classical bits 0 and 1, as shown in Fig. 3.6, where I make use of a useful geometric representation of the system. Thus, the complete description of a free system expressed in the previous basis is given by

$$|\psi\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\phi}\sin\frac{\theta}{2}|1\rangle\ , \tag{3.80}$$

where the coefficient $\cos(\theta/2)$, $\sin(\theta/2)$ are assumed to vary continuously and the exponential $e^{i\phi}$ expresses the phase difference ϕ between the two components. Note

[99] Planck (1922).

[100] Auletta (2011a, Sect. 2.1).

[101] Timpson (2013, Sects. 2.2.5, 4.4).

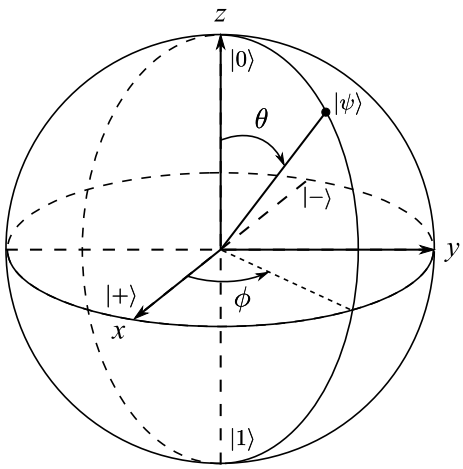


Fig. 3.6 The state vector of a two-state quantum system, for instance $|\psi\rangle$, can be represented as a point on the surface of a unit sphere, called the Poincaré sphere (originally introduced by the French physicist H. Poincaré for representing polarised light; sometimes it is also called *Riemann sphere*, after the name of the German mathematician and physicist Bernhard Riemann (1826–1866)). The parameters ϕ (here represented as angle between the y-axis and the projection of $|\psi\rangle$ on the equatorial plane) and θ (represented as an angle between $|\psi\rangle$ and the z-axis) are sufficient to individuate its location. The states $|+\rangle$ and $|-\rangle$ lie on the equatorial plane and represent two symmetric superpositions of $|0\rangle$ and $|1\rangle$, located at the south and north poles, respectively. (Thus, for the sake of convenience, I have chosen the diameter connecting them to be along the x-axis as well as the diameter connecting $|0\rangle$ and $|1\rangle$ is along the z-axis.) In other words, orthogonal states that were previously represented as orthogonal vectors in the vectorial representation are here represented as the two opposite points of the intersection between a diameter and the surface of the sphere. Adapted from Auletta and Wang (2014, p. 338)

that the factors 1/2 derive from the Poincaré sphere representation in which the angle θ is doubled relative to ordinary rotations in the Hilbert space: in fact, the two orthogonal states $|0\rangle$ and $|1\rangle$ make in the Hilbert space an angle of 90° while in Fig. 3.6 are separated by 180°.

Infinite Information

The two parameters θ and ϕ span the whole surface of the Poincaré sphere, the amplitude being represented by the parallels and the phase difference by the meridians. Then, we see that the system can be actually located at any of the infinite points covering the surface of the sphere (while, as a result of a preparation, we would confine the system at one of the points in that expansion basis). In other, words it is like a binary code covering any possibility. And, for this reason, an infinite or at least unlimited amount of information would be necessary to predict what would be

the outcome,[102] what is another way to say that those outcomes are unpredictable. Note that this statement is true independently of the dimensions of the system and therefore also of the number of the possible outcomes.

For such a reason, any quantum system can be conceived of as a kind of replicator of the fundamental character of the whole "fabric of reality" (Sect. 3.1.2), according to Schrödinger's objectivism[103]: both the part and the whole represent a combinatorics of infinite possibilities. Thus, quantum systems display a certain self-similarity. In other words, quantum systems, if not limited by particular physical conditions, express a superposition of *any* possibility. This can be shown as follows: suppose a quantum trapped in some potential well and assume that we can turn off it in discontinuous way. Then, simultaneously, the quantum has a non-zero probability to be found everywhere in the universe, what means that it can occupy any possible position with some probability >0.

Qubits

The previous discussion helps us to understand better what is a preparation, also called *determinative measurement*. Thus, when we prepare a system, we *select* the state of a system that can be in superposition in a *certain basis*, e.g.

$$|0\rangle := \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad |1\rangle := \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{3.81}$$

Such prepared state, when the superposition is symmetric, in accordance with Eq. (1.403a), can be written as

$$|+\rangle = \frac{1}{\sqrt{2}} \left(|0\rangle + |1\rangle \right), \tag{3.82a}$$

where I have assumed $\phi = 0$. We have here a new information entity represented by a *particular* superposition of 0 and 1: any of such states is called a quantum bit or in short a *qubit* (also quantum of information). I recall that even when we select a particular state like (3.82a) but the system is thereafter free to evolve, it needs to be again represented with the general form (3.80), whose coefficients and relative phase are not specified. Note that, by performing a preparation, we have not yet selected an observable to measure. In fact we could certainly decide to measure the observable whose eigenstates are $|0\rangle$ and $|1\rangle$ but, as the multiple examples previously introduced show, could also decide to measure the observable whose eigenstates are superpositions of this basis: in other words, we could decide to measure the observable whose eigenstates are, e.g. $|+\rangle$, given by Eq. (3.82a), and its orthogonal

[102] As pointed out in Schumacher (1990), Rozema et al. (2014). See also Auletta (2006b, 2011a, Sect. 2.2.2).

[103] See also Deutsch (1997, p. 97).

$$|-\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle - |1\rangle\right), \tag{3.82b}$$

according to Eq. (1.403b) and Fig. 3.6 or whatever other superposition of $|0\rangle$, $|1\rangle$. Two aspects are relevant here. Let us first state them and then explore the consequences. On the one hand, in agreement with Schrödinger's quantum objectivism, the specific physical conditions (for instance, of experimental kind) in which the system is and the processes (like preparation) to which it can be subjected *do not change* the fact that the system is a superposition of the type (3.82) or (3.80). On the other hand, this total amount of information is in large part locally lost in the measurement process (see also Sect. 2.2.2). The crucial issue here is that the measurement process can be considered as a local selection sequence, whose first step after preparation is the choice of a specific *observable* (a problem that we have already considered in Sect. 3.1.2); I recall that this is known as premeasurement. Suppose that we choose to measure, e.g. an observable whose eigenstates are $|0\rangle$ and $|1\rangle$. Then, in a subsequent step, when we will *detect* the system, we shall obtain either $|0\rangle$ or $|1\rangle$ as the measurement outcome. This 'choice' between two possibilities makes precisely one (classical) bit of information, a much smaller amount of information than what is initially contained in the initial superposition state.

Holevo Theorem

The previous result is summarised by the *Holevo theorem*, after the name of the Russian mathematician Alexander Holevo[104]: from a qubit we can acquire at most a bit of classical information. When measuring a parameter X with values $x \in X$ (and having apparatus' outputs given by values of Y), the *Holevo bound* on information accessibility (in the discontinuous case), provided by the mutual information between X and Y, is given by

$$I(X:Y) \leq S(\hat{\tilde{\rho}}) - \sum_{x=1}^{n} \wp_x S(\hat{\rho}_x), \tag{3.83}$$

where the density matrix

$$\hat{\tilde{\rho}} = \sum_{x=1}^{n} \wp_x \hat{\rho}_x \tag{3.84}$$

of the object system is a mixture,

$$S(\hat{\tilde{\rho}}) = S\left(\sum_{x=1}^{n} \wp_x \hat{\rho}_x\right) \tag{3.85}$$

[104] Holevo (1998). See also Nielsen and Chuang (2000, Sect. 12.1.1).

is its von Neumann entropy, X, the variable of the observed system, represents inputs, whereas Y ranges on the possible detection apparatus' outputs. The set of density matrices $\{\hat{\rho}_1, \hat{\rho}_2, \ldots, \hat{\rho}_n\}$ can be also mixtures and stand for possible measurement outcomes, according to von Neumann's formula (3.14). What this inequality tells us is that the apparatus can share information with the object system (and therefore extract information from it) in a measure that is equal to or less than the so-called *Holevo quantity*

$$\chi = S\left(\sum_{x=1}^{n} \wp_x \hat{\rho}_x\right) - \sum_{x=1}^{n} \wp_x S(\hat{\rho}_x). \tag{3.86}$$

Moreover, since $\forall \hat{\rho}_x$ we have

$$S(\hat{\rho}_x) \geq 0, \tag{3.87}$$

with equality when the density matrices are in fact pure, we also have

$$I(X : Y) \leq S(\hat{\bar{\rho}}), \tag{3.88}$$

which tells us the, independently of the kind of measurement that we can perform, the related information transfer from the object system to the detector can never be larger than the entropy of the mixture of the object system that one obtains during the process.

Information Accessibility

Thus, as anticipated, the main difference between a quantum state as described by its wave function before detection and the information that we acquire after detection is due to a problem of information *accessibility*, and therefore it depends on the information that we can factually acquire. This involves three different aspects:

- The reservoir of information that we have when we prepare and premeasure a system must represent *potential information*, that is, information that *could* be acquired in certain physical contexts.
- We must find a reason of why there is an accessibility problem. In other words, why we do not acquire the whole amount of information that is contained in the initial state? We have only partially answered this question.
- We need to understand detection as responsible of a locally irreversible displacement of entropy. This is a point that we have already clarified (in Sect. 3.1.3), but some additional considerations will be necessary.

Encoding Information

Let us consider the first point. The previous examination leads us to the conclusion that the information contained in the superposition of, e.g. two arbitrary orthog-

onal states is independent of any experimental procedure of information acquisition (preparation, premeasurement, detection). As anticipated, we can understand it as linearly *encoded* information. There are several requisites of linear information encoding,[105] and these are satisfied by quantum systems:

(i) There exist (at least) two orthogonal states that represent two mutually exclusive possible measurement outcomes of a quantum system. According to the superposition principle, these two orthogonal states constitute an orthonormal basis that we can take to be the *code alphabet* and thus can be linearly combined into an arbitrary superposition, like in Eq. (3.82a). Note that while not all combination rules are linear, linearity is a necessary requirement for *quantum information* encoding since qubits are represented by vectors in a vector space that can be linearly combined (Sects. 1.1.3 and 1.2.1–1.2.2). The resultant orthonormal basis is referred to as the computational basis, and the two computational basis states are conventionally denoted by $|0\rangle$ and $|1\rangle$, as mentioned.

(ii) By varying the coefficients and the relative phase of the superposition, we can obtain an infinite number of possible superposition states (represented by the different points on the surface of the Poincaré sphere). In other words, it is in principle possible to express *any* given state representing a qubit by means of a superposition of the two computational basis states $|0\rangle$ and $|1\rangle$. Therefore, the value that we assign to the coefficients and the relative phase of the superposition represent the rules according to which we *combine* the two computational basis states $|0\rangle$ and $|1\rangle$, and these combinations can be understood as *codewords* constituting a (binary) code (or code of length 2), as in quantum computation.[106] A code is a set of codewords in a given alphabet.

(iii) It is always possible to choose another set of orthogonal states as the computational basis states, for instance, the states $|+\rangle$ and $|-\rangle$. In other words, different computational bases are different code alphabets used to encode quantum information. For a given state, there are specific rules that allow the translation of expressions from one computational basis to another. These translation rules for quantum information are precisely unitary transformations for the corresponding changes of basis. Indeed, this is a necessary requirement for having information: information cannot be dealt with if there is not a (at least potential) plurality of code alphabets.

The informational dimension of a system is the cardinality of the largest maximal set of perfectly discriminable pure states.[107] Thus, it represents the size of the largest alphabet that can be exactly encoded on a given system. This justifies the multidimensionality of quantum systems in Hilbert spaces. We have remarked (Sect. 1.2.1) that some physicists considered the fact that quantum systems 'evolve' in a Hilbert

[105] A good and basic textbook is Ling and Xing (2004).

[106] On quantum computation a good textbook is Nielsen and Chuang (2000). For a short summary of the subject see Auletta et al. (2009, Sect. 17.7), Auletta and Wang (2014, Sect. 11.4). I shall come back on these problems.

[107] D'Ariano et al. (2017, Sect. 12.3).

space (and not in an ordinary configuration space) as a weird state of affairs. Now, the fact that these dimensions are considered possible outcomes and the number of this outcomes is independent of the basis (observable) that will be chosen is precisely rooted in the multidimensionality of information and the possibility to use different equivalent code alphabets. In other words, information allows us to connect the state of the system and the result of the detection as two different aspects of a single entity.

Potential Information

There is another important point to examine. When we say *potential information* this seems to deprive information of the ontological status that I wish to assign to it. This is a very difficult problem and need a separated investigation. By now, I stress that, when we say *potential* we mean something more than purely possible. Although we have defined quantum information as a correlation among possibilities, when we deal with a real quantum system during, e.g. a measurement procedure, we are working with something that at certain stages represents something more than a pure possibility.[108] In fact, during the steps of preparation and premeasurement we have a specific physical, experimental context that makes the information contained in the system potentially accessible (although not as a whole): during preparation it is an infinite amount of potential (quantum) information (the information that is in a state of a particular selected system), during premeasurement classical bits of information (corresponding to the eigenbasis of a particular selected observable) become *acquirable* in those conditions and not others.[109] It seems to me that Zurek's notion of *actionable* information hints at the same process.[110] These crucial distinctions will be object of further analysis.

Note that, also classically, information that is encoded is not necessarily used or made active, and therefore as such is only potentially available (through certain procedures in determinate contexts). A good example is represented by DNA that, as long as it is not expressed, is dormant and therefore, at least to a certain extent, it is as it did not exist. Nevertheless, it is real, since the organism can make use of it at a later moment, for instance, under a selective pressure or at particular developmental stages.[111]

Information That Is Not Acquired

Now, let us consider the second point: what is that makes this information inaccessible in its totality, i.e. in cumulative terms? We have said the non-commutativity of the observables. And what is the factor determining such a situation? Precisely

[108] On this distinction see Auletta (2011c, Sect. 3.2.5).

[109] Conway and Kochen (2006, 2009).

[110] Zurek (2013).

[111] This is extensively discussed in Auletta (2011a, Chaps. 7–11). See literature quoted there.

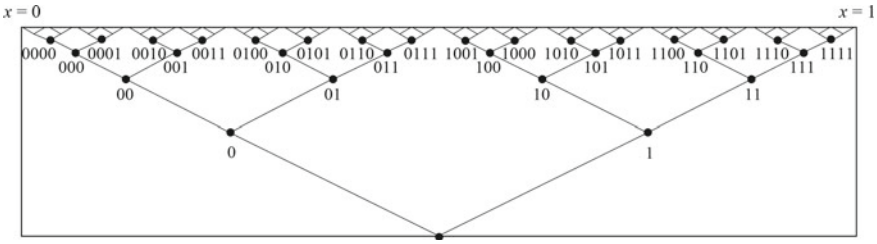


Fig. 3.7 The continuous unit interval [0, 1] can be considered as a limiting case of digitised coding in which the number of digits approaches infinity. Adapted from Auletta (2011a, p. 38)

the interference terms, i.e. the interdependencies among the elements of the code (Sect. 2.1.2). Why? Because they represent information (mutual information, indeed) that *can never be acquired* through a single local measurement: it is locally inaccessible. This is the consequence of a more general epistemological problem: we cannot make experience (and therefore neither directly acquire information) about relations but we always need to consider the things or elements that are *in relation*. It is so because any kind of experience (and information acquisition is one of them) is necessarily *local*. In other words, quantum mechanically we download into the environment those non-local quantum correlations that contribute to the infinite amount of information contained, e.g. in a state of the form (3.80) but also make the information contained in the state inaccessible as long as they are locally dominant. Of course, downloading this kind of information into the environment makes also the information about all components (apart from the one that we actually get through detection) inaccessible.

As a matter of fact, also classically there are accessibility problems, but these are more related with the factual impossibility to reduce to zero any measurement error, as recalled in Sects. 1.1.2 and 2.2.2.[112] Nevertheless, in both (quantum and classical) cases the ultimate reason is the finite resolution of any code that we can use *when measuring a system* (Fig. 3.7). Information and quantisation go together.

Encoding and Accessibility

This allows us to deal with the third problem. At a *formal level* or a priori (on the pure plane of possibility) there are no appreciable differences between an initial state in superposition (whether free or prepared) and a measurement outcome (any of the components of the superposition), provided that the system survives detection. In fact, both $|0\rangle$ and $|1\rangle$ can themselves be expressed as superpositions (for instance, as some combinations of the states $|+\rangle$ and $|-\rangle$ displayed in Fig. 3.6) as follows

[112]The reader may have a look at Auletta (2011a, Sect. 2.2).

$$|0\rangle = \frac{1}{\sqrt{2}}\left(|+\rangle + |-\rangle\right), \qquad |1\rangle = \frac{1}{\sqrt{2}}\left(|+\rangle - |-\rangle\right). \tag{3.89}$$

If quantum systems are self-similar, this circumstance will never changes whatever local operation we can perform. As mentioned, such information is independent from the procedures of preparation, premeasurement and information acquisition, i.e. local detection. Again, on this point the quantum-objectivist point of view is right: quantum-mechanical formalism describes something that is independent of what can happen here or there. Thus, it is quantum objectivism that requires the objectivity of quantum information.

However, this is only one side of the coin. From the point of view of this information acquisition (and even of premeasurement) or *locally* and a posteriori there is a remarkable difference, expressed by the Holevo theorem. Of course, the system that we prepare is a real "object" whose information could be acquired by specifying physical conditions of the extraction and finally performing a real act of detection. We can thus reconcile the a priori (encoding) and the a posteriori (information acquiring) understanding of information in terms of information accessibility and local shift of entropy: every time that we have an *appropriate objective context*, i.e. the appropriate physical conditions for the dynamic interaction between two open systems are satisfied (i.e. anytime after preparation when we have a premeasurement or a premeasurement–like interaction process), we are able to acquire at least a classical bit of information in the form of a measurement result out of an initial and possible higher variety.

For the intelligence of the whole is crucial to understand that there is a fundamental difference between "being in relation" and "being subjective": to assert that things are referred to contexts does not imply that those contexts need to have a subjective component.[113] This relational view of QM has been strongly supported by the Italian physicist Carlo Rovelli.[114] Later on, we shall explore the consequences of this view.

The American information theoretician Charles Bennett defines information as "the notion of distinguishability abstracted away from what we are distinguishing, or from the carrier of information".[115] This fundamental definition needs to be explored step by step. The idea that information can be abstracted away from any physical carrier is very important as far as I have stressed that there is a meaning of information that is independent from any possible physical context, and this constitutes its objectivity. Moreover, it clearly describes acquired information where we have in fact a "choice" and therefore a *difference*. However, at a general level, information only demands *variety* (as expressed, e.g. by the superposition of some components), so that, according to the previous examination, information is what both connects (through mutual information) and distinguishes.

[113] As pointed out in Auletta (2006b).

[114] Rovelli (1996, 2005).

[115] Bennett (1973).

3.2.4 Two Points of View on Entropy and Information

Information and Entropy

I have mentioned the fact that entropy can be understood as the information that would be necessary for describing the system. This is the interpretation of C. Shannon,[116] and has become the canonical one. From Shannon's point of view, seems that information only deals with the knowledge that we can have of a system (and this was in agreement with Wigner's subjectivist interpretation). However, from previous examination, it appears that entropy and information are inversely related. This second point of view has been supported by E. Schrödinger,[117] followed by the French physicist Léon Brillouin (1889–1969).[118] Also the American mathematical physicist Elliott Lieb affirms that information is *negative entropy* or negentropy.[119] Brillouin has stressed that, during the communication along a channel, some information is lost or degraded because of noise (spread of the signal). Therefore, he concludes that, during this process, information is decreasing, while entropy increases. The situation can be compared with the evolution of a thermodynamic or chaotic system: with the time our capacity to describe the initial state from the current state decreases.[120] Using this framework, Brillouin could solve the paradox of Maxwell's demon, after the name of J. Maxwell (a demon controls the door between two chambers filled with gas and sorts the molecules, initially mixed, into fast and slow).[121] Also in QM we expect that any information acquisition, due to the quantum of action, will cost in terms of (local) increase in entropy. Summarising, while for Shannon entropy measures information that needs to be added, for Brillouin it measures *lack* or loss of information.

Noiseless-Channel Coding

Although Brillouin's model is too specific to be generalised to quantum systems as far as it deals with exchanges of signals and therefore does not catch the peculiarity of quantum information when understood in objective terms, there is nonetheless something correct in his understanding of information. The very issue at stake here is the nature of entropy. In the previous subsections, I have introduced the notion of Shannon's entropy without explaining its original derivation in communication theory. Shannon proved that entropy represents the reliable maximal compression of some information (at the source), so that any further compression scheme would

[116] Shannon (1948).
[117] Schrödinger (1992).
[118] Brillouin (1962).
[119] Lieb (1975).
[120] On this issue see Ott (1993), Schuster (1988).
[121] On this point see Cohen-Tannoudji (1991, pp. 52–53, 68–69).

not be reliable[122]: compressibility of data is proportional to the redundancy present in the sequence to be compressed. (We should in fact distinguish between source and channel coding.[123]) Thus, a perfectly compressed message represents maximum entropy. Note that only if the entropy at the source is less than the channel capacity, the data at the source can be reliably transmitted, where with *channel capacity* we understand the maximum of the mutual information between input and output (thus, the Holevo bound (3.83) is a quantum manifestation of this limitation).

The main insight of Shannon was to divide the possible sequences of values $x_1, x_2, \ldots, x_n$ for random variables $X_1, X_2, \ldots X_n$ into two kinds of sequences: typical sequences, which are more likely to occur, and atypical sequences, which occur rarely, according to the source. If they are emitted with a fraction $\wp$ of the source symbols being 0 and a fraction $1 - \wp$ being 1, for n becoming large, we expect also that a fraction $\wp$ of the source outputs will be 0 with high probability, while a fraction $1 - \wp$ will be 1. With the independence assumption for the source, we then have

$$\wp(x_1, x_2, \ldots, x_n) = \wp(x_1)\wp(x_2)\ldots\wp(x_n) \approx \wp^{n\wp}(1-\wp)^{n(1-\wp)} \tag{3.90}$$

for a typical sequence. Taking negative logarithms of both sides, we obtain

$$-\lg \wp(x_1, x_2, \ldots, x_n) \approx -n\wp \lg \wp - n(1-\wp)\lg(1-\wp) = nS(X), \tag{3.91}$$

in agreement with Eq. (3.43). This can be generalised to

$$\left| -\frac{1}{n} \lg \wp(x_1, x_2, \ldots, x_n) - S(X) \right| \leq \epsilon, \tag{3.92}$$

where both sides of the previous equation have been divided by n and ϵ is a small quantity.

Thus, while Shannon (who dealt especially with communication technology) identifies entropy and information and sees maximum entropy as a perfectly compressed message, Brillouin considers it as a perfectly random message (because he sees information as negative entropy). The two perspectives can be brought to agreement if we consider that the objective loss of information by a system during some communication process demands more information to describe it. The reason is that all regularities in nature ground not only order but also our expectations, so that it is violation of expectation and therefore *deviation* from regularity that need to be explained (when possible), and thus objective and subjective perspectives express the same state of affairs. Deviation from expected regularity or uniformity has been always one of the major reason for the search of causal explanations. It may be recalled here that S. Laplace invoked causal explanations for phenomena that he considered highly improbable.[124]

---

[122] Shannon (1948). See also Nielsen and Chuang (2000, Sect. 12.2.1).

[123] On this see Battail (2014, p. 56).

[124] Laplace (1796, pp. 541–44).

Quantum Compression

This intrinsic probabilistic character of quantum information allows for amazing results when considering recent models and experiments allowing new way to compress quantum information (but not as source encoding!).[125] This shows a characteristic difference between classical and quantum information. Indeed, as analysed in Sect. 2.1, the amount of information that can be extracted from a classical system is precisely the same as the amount of information required for a complete description of the system's state (as a consequence this amount of information must be finite), while, as mentioned in the previous subsection, this is not true quantum mechanically; no more than one (classical) bit of information can ever be extracted from a local measurement of a single quantum bit (qubit). It is such fundamental difference between quantum and classical mechanics that opens up the possibility of new kinds of data compression with no classical analogue. In QM an ensemble of identically prepared quantum systems (if pure, with zero Shannon and von Neumann entropy) provides *much more information* than a single copy, while this is clearly not the case classically, where the information encoded in a single system's state can be accessed repeatedly without adding new elements (due to the assumed completeness of information extraction). Although quantum mechanically we cannot compress all of the information contained in an ensemble of systems down to a single quantum copy, an exponential saving can be achieved. The experimental team of the physicists Lee Rozema et al. has shown that an ensemble of n identically prepared qubits (requiring an Hilbert space of 2^n dimensions) can be compressed into a memory of size $\lg(n+1)$ qubits. This can be made because most of the information in a multiqubit state describes permutations (Sect. 1.3.1), which are irrelevant for an ensemble of identical qubits (but one could also envisage to exponentially compress any permutationally invariant pure state).

Suppose that $|\Psi\rangle$ describes a symmetric state of N qubits. In other words, the state vector remains unchanged under application of the unitary permutation operator $\hat{U}_P$. A basis for the set of symmetric states can be chosen in such a way that every basis state has a determinate number of excitations (i.e. qubits in the state $|1\rangle$), so that the respective basis states can be labelled as follows:

$$|N;k\rangle = \binom{N}{k}^{-\frac{1}{2}} \sum_P \hat{U}_P \left(|1\rangle^{\otimes k} \otimes |0\rangle^{\otimes (N-k)} \right), \tag{3.93}$$

where the sum runs through all permutations of the qubit systems, having $\binom{N}{k}$ terms (see Eq. (1.352)). The notation $|1\rangle^{\otimes k}$ means k times the state $|1\rangle$, and similarly for the other cases. In practice, we have partitioned all N qubits in two subsets given by k qubits in the state $|1\rangle$ and $N-k$ qubits in the state $|0\rangle$. The basis states are

[125] The reference papers in this subject are Plesch and Bužek (2010) for the theoretical study and (Rozema et al. 2014) for the performed experiment.

orthonormal. As mentioned, the authors have used a transformation that takes the symmetric states (3.93) into a subset of computational vectors. This subset is formed by (i) the vectors $|0\rangle^{\otimes N}$ and (ii) all vectors having a *single* excitation (a single 1). These are the new codewords. This subset occupies the Hilbert space of the same dimension as symmetric states and defined these two as the new codewords

$$|C\rangle_0 = |0\rangle^{\otimes N}, \tag{3.94a}$$

$$|C\rangle_k = |0\rangle^{\otimes(k-1)} \otimes |1\rangle \otimes |0\rangle^{\otimes(N-k)}, \tag{3.94b}$$

respectively. In practice, we take the original vectors, combine them according to all permutations, that is, each time according to a fixed number of $|0\rangle$'s and $|1\rangle$'s and find for each of these combinations a new codeword that is either $|000\ldots0\rangle$ or $|010\ldots0\rangle$, with *all* its permutations. Then, let use a unitary transformation of this general form

$$\hat{U}\,|N;k\rangle = |C\rangle_k, \tag{3.95}$$

and consider some cases. The one qubit case is trivial since no transformation is needed (with $|0\rangle$ and $|1\rangle$ mapping to themselves). In the two-qubit case, we have the mappings

$$|00\rangle \longmapsto |00\rangle, \tag{3.96a}$$

$$|01\rangle + |10\rangle \longmapsto \sqrt{2}\,|10\rangle, \tag{3.96b}$$

$$|11\rangle \longmapsto |01\rangle, \tag{3.96c}$$

where I have used the synthetic notation $|xy\rangle = |x\rangle_1\,|y\rangle_2$, for $x, y = 0, 1$. Only the singlet state (1.393) or (2.54) is not defined by these transformations and we need to add another codeword:

$$\frac{1}{\sqrt{2}}\left(|01\rangle - |10\rangle\right) \longmapsto e^{i\phi}\,|11\rangle, \tag{3.97}$$

where the singlet state has been reformulated in informational terms on the LHS. However, the performance is increased for increasing number of qubits. Indeed, for three qubits we get four states out of eight independent basis states:

$$|000\rangle \longmapsto |000\rangle, \tag{3.98a}$$

$$\frac{1}{\sqrt{3}}\left(|001\rangle + |010\rangle + |100\rangle\right) \longmapsto |100\rangle, \tag{3.98b}$$

$$\frac{1}{\sqrt{3}}\left(|011\rangle + |101\rangle + |110\rangle\right) \longmapsto |010\rangle, \tag{3.98c}$$

$$|111\rangle \longmapsto |001\rangle. \tag{3.98d}$$

3.3 Hidden Variables and Non-locality

From a general point of view, the EPR argument, presented in Sect. 2.3.1, shows four fundamental aspects:

1. A philosophical part, in particular the notions of reality and completeness,
2. A logical component,
3. A physical–theoretical part, aiming at proving that QM is incomplete, and
4. A physical–experimental part, proposing an ideal experiment for testing the incompleteness of QM.

The first and the latter two issues have been already presented in general terms. In the next subsection, we shall deal with the logical structure of the EPR argument and subsequently analyse the different positions about the other three issues.

3.3.1 Logical Restatement of the EPR Argument

In Sect. 2.3.4 we have seen that it is possible to reject the EPR argument by denying validity to either the assumption of reality or the principle of separability and said that in fact Bohr and Schrödinger have chosen the former and the latter strategy, respectively. In order to understand the root of their objections, we need to consider the abstract logical structure of the EPR argument.[126] Logically speaking, such an argument can be schematically summarised as follows[127]:

(i) The sufficient condition of (classical) reality, and
(ii) The separability condition imply and
(iii) The non-completeness of QM.

In other words, the argument takes the logical form

$$RS \rightarrow C', \tag{3.99}$$

where RS denotes that R and S are connected by the logical AND, $\rightarrow$ is the symbol for logical implication, and the prime denotes negation (which, I recall, inverts the truth value of a statement). Table 3.2 displays the truth table for conjunction. For reasons that will be clear later, I shall use only logical implications, that is, hypothetical statements that are tautologies (can never be false), of the kind $XY \rightarrow Y$. In other words, given the truth of premise, the truth of the consequence necessarily follows.

The argument is developed in three steps. First, suppose that non-commuting observables could have simultaneous reality (definite values), then these values would

[126]This is likely due especially to Podolsky (Home and Whitaker 2007, p. 109), although I think that Einstein considerably contributed.

[127]The reader may also have a look at Auletta et al. (2009, Sect. 16.1), Auletta and Wang (2014, Sect. 10.1).

Table 3.2 Truth table for conjunction

X	0	0	1	1
Y	0	1	0	1
XY	0	0	0	1

Table 3.3 Inclusive disjunction

X	0	0	1	1
Y	0	1	0	1
$X + Y$	0	1	1	1

Table 3.4 Exclusive disjunction

X	0	0	1	1
Y	0	1	0	1
$X \nsim Y$	0	1	1	0

enter in a complete description of the quantum system, according to completeness. But if the wave function of the system provides a complete description of quantum systems and contains such values, they would be predictable but this is not allowed by the uncertainty relations. Thus, it follows that the following two statements are incompatible:

(i) The statement C' that the quantum-mechanical description of reality given by the wave function is not complete.
(ii) The statement R' that when the operators describing two physical quantities do not commute, the two quantities cannot have simultaneous reality.

In formal terms, we have the statement

$$C' \nsim R', \tag{3.100}$$

where the symbol $\nsim$ denotes an exclusive disjunction: either C' or R' are true but cannot be both false, nor both true (in other words, exclusive disjunction is a countervalence or negation of material equivalence). See Tables 3.3 and 3.4, for the truth tables of inclusive (OR) and exclusive disjunction (XOR), respectively. In other words, we have derived the previous statement from (i) completeness, (ii) the hypothetical assumption of the uncertainty relations, and (iii) the fact that, according to the latter, two non-commuting observables cannot simultaneously have definite values. In formal terms,

$$CUR' \rightarrow (C' \nsim R'), \tag{3.101}$$

which is clearly a tautology (always true). In fact, the logical consequence would be falsified only in the case in which the antecedent CUR' is true and the consequent $C' \nsim R'$ us false. However, if the antecedent is true both C and R' are true, what shows that also $C' \nsim R'$ is true, and this represented a contradiction.

Second, if, by further assuming the separability condition, it is possible to show (through some kind of experiment) that two non-commuting observables have in

fact simultaneous reality, we have proved that the criterion of reality (in the sense of *omnimoda determinatio*) *R* holds also for QM. Third, the incompleteness of QM follows from statement (3.100). Then, we have proved the implication (3.99).

In the first stage of his criticism towards QM, Einstein may have tried to show that the uncertainty relations were false and consequently quantum theory inconsistent. Subsequent reflection, at least from 1930 onwards (likely after Bohr had replied to Einstein's clock–shutter device, an issue on which we shall come back), has induced a change of his point of view, since EPR's aim is not to falsify quantum theory but showing its incompleteness.[128] Of course, an important aspect of the whole argument is how we understand incompleteness. In general terms, I have stressed that EPR's aim was to show the *incompleteness* of QM, i.e. its insufficiency in dealing with a complete description of microreality. Some historians have also stressed that such an incompleteness could be understood as a failure of univocity as far as we can simultaneously assign two different wave functions to the state of the two particles.[129] Such an interpretation would certainly have the advantage to disentangle the issue of realism in the sense of the *omnimoda determinatio* (Sects. 1.2.1 and 1.2.4) from the issue of separability. In any case, according to Einstein, the two different descriptions of the same state hinted at properties that are real independently of any measurement or measurement context. Let us consider these problems more carefully.

3.3.2 Bohr's Reply to EPR

EPR's Realism of Observables

Summarising, given the logical structure of the EPR argument, if we desire to reject the conclusion that QM is incomplete, it is necessary to show either the failure of the criterion of physical reality in a quantum framework or the inconsistency of separability with QM. In fact, they are the only non-quantum-mechanical assumptions in the EPR logical structure of the argument and their thought experiment. EPR themselves have anticipated the former objection. In the end of the paper, they replied to a possible critic as follows[130]:

> One could object to this conclusion on the grounds that our criterion of reality is not sufficiently restrictive. Indeed, one would not arrive at our conclusion if one insisted that two

[128] On these problems see Jammer (1974, Chap. 6). See also Harrigan and Spekkens (2010), although I disagree with the authors (and with A. Fine, who first introduced this idea in Fine (1981)) that the EPR paper does not correspond to Einstein's view (the fact that Podolsky may have been the material extensor of the article tells us nothing about the issue at the stake), and consider their judgement about the presumed "opaque" logical structure of the argument a true misunderstanding of what the paper says. The argument is clearly complex but not obscure: it could have been reduced to implication (3.99) and an experimental evidence, but to assume the hypothetical validity of the uncertainty relations was mandatory.

[129] As pointed out in Howard (1992).

[130] Einstein et al. (1935).

> or more physical quantities can be regarded as simultaneous element of reality only when they can be simultaneously measured or predicted. On this point of view, since either one or the other, but not both simultaneously, of the quantities P [momentum] and Q [position] can be predicted, they are not simultaneously real. This makes the reality of P and Q depend upon the process of measurement carried out on the first system, which does not disturb the second system in any way. No reasonable definition of reality could be expected to permit this.

This is a further evidence that EPR had envisaged a possible major conflict on the issue of reality but never occurred to them that separability could be put into discussion. It is worth noticing that the EPR argument deals here with the ontological status of *observables* (like position and momentum). In other words, it is first observables that are considered by them to be elements of reality.[131] This is quite understandable, since their argument is thought to be a kind of limitation of the scope of the uncertainty relations, which in fact deal with the relation between (conjugate) observables (Sect. 1.2.4; see also Sect. 2.4.2). Did EPR also consider the values of those observables (i.e. properties) to be elements of reality? They say: "If, without in any way disturbing a system, we can predict with certainty ... the value of a physical quantity, then there exists an element of physical reality corresponding to this physical quantity". To deal with properties is trickier and does certainly represent a stronger form of realism than the previous one. In the following, I shall try to keep these two different problems apart.

Bohr's Epistemic Restriction on Reality Ascription

In the same year, 1935, in two papers Bohr rejected the sufficient condition of reality precisely along the argumentative lines forecast by EPR.[132] Previously,[133] Bohr had faithfully expressed Einstein's relativistic point of view by saying that, from a classical point of view, observing ultimately rests on the coincidence of two independent events at the same space-time point, and this coincidence cannot be affected by differences among the space-time coordinates of different observers. However, Bohr remarked that this cannot be true for QM, since such a theory requires an interaction with the observing agency that hinders any attribution of independent reality to the system (and, I add following Schrödinger, neither to the agencies of observation). In this first formulation (1927–1928) of the Copenhagen interpretation both Bohr and Heisenberg rely on the notion of the disturbance determined by the observer on the object system during detection (see also Sects. 2.1.2 and 3.1.1). Now, in 1935, Bohr should be aware that this argument is not a valid objection against the EPR model that, as recalled by EPR themselves, supposes a measurement *without disturbance* of (if we like, without interaction with) the second particle, and makes of this a cornerstone of the criterion of reality. From this point of view, the EPR argument can

[131] See also Margenau (1950, pp. 299–300).

[132] Bohr (1935a, b). See also Auletta and Wang (2014, Sect. 10.2).

[133] Bohr (1928).

be considered the first theoretical model of indirect measurement (where we extract information from a system by measuring another one coupled with it)[134] and thus bears a certain resemblance with the interaction-free measurement. For this reason, we may consider the argument as a specific reply of Einstein to the interpretation of QM that Bohr and Heisenberg were developing at that time. Then, Bohr, in his aim at falsifying the EPR criterion of reality, was forced to change the status of the restrictions imposed by QM on ontological ascriptions from a dynamical (operational) into an *epistemic* one.

In fact, in his first paper,[135] Bohr criticised the EPR argument by pointing out that, *even excluding any direct physical interaction* of the system with the measuring apparatus (as it is the case for the second particle), the measurement process "has an essential influence on the conditions on which the very definition of the physical observables in question rests". And, according to Bohr, "these conditions must be considered as an inherent element of any phenomenon to which the term 'physical reality' can be unambiguously applied". Bohr acknowledged that it is possible to determine experimental arrangements such that the measurement of the position or of the momentum of one particle automatically determines the position or the momentum of the other without interfering or interacting with the latter. However, such experimental arrangements for measuring momentum and position are incompatible with each other (an argument that is somehow related not only to complementarily but also to the delayed-choice experiment: see Sect. 2.4.3). Therefore, the central point of Bohr's criticism is that it is not possible to ascribe reality to observables of quantum systems (and therefore certainly not properties to the involved systems) independently of the experimental context in which not only we actually interact with them but in which we *could* interact with them.[136] It is not by chance that the American physicist and philosopher of science Abner Shimony (1928–2015) has spoken of "epistemology without an ontology".[137] It is worth mentioning that this epistemic understanding of the context is precisely what constitutes the received form of the Copenhagen interpretation[138] (Sect. 3.1.4) that has raised a so wide dissatisfaction among physicists.[139]

[134] On this point see Braginsky and Khalili (1992, pp. 40–49). For a short summary see also Auletta et al. (2009, Sect. 9.11.1).

[135] Bohr (1935a).

[136] This way of thinking pertains the family of forms of non-monotonic reasoning that are typical for empirical problems especially when induction is involved: see Pearl (1988, p. 59).

[137] Quoted in Home and Whitaker (2007, p. 64).

[138] For a look at the different forms that the Copenhagen Interpretation has received see Auletta (2000, Chaps. 6–8). See also Shimony (1981) on Bohr's epistemology.

[139] See, e.g. Deutsch (2011, Chap. 12).

Reality Ascription to Observables Depend on Physical Conditions

Now, I wish to point out that the experimental procedures through which observables are determined are *physical operations*. In fact, EPR say that "the elements of physical reality cannot be determined by a priori philosophical considerations, but must be found by an appeal to results of experiments and measurements", where "operations" and "measurements" do not necessarily imply a direct interaction with the system (as it is the case for the second particle).[140] Moreover, from such physical point of view, an *observable* is precisely connected with an operation that could be performed on a system either artificially in our laboratories or that we assume to happen spontaneously in nature when certain kinds of processes occur (Sect. 3.1.3). In this way, an observable describes a range of possible behaviours of a system in certain *physical conditions*. Note in fact, that to measure in QM has not the significance of registering a value that the system already possessed by itself, but rather to check *how a system will react* (during detection) when certain experimental conditions for measuring a certain observable are imposed (Sects. 2.2.1–2.2.2). Therefore, the accent here is not on detection but on those conditions (in general at the premeasurement stage). Most of the properties and parameters that we ascribe even to classical and macroscopic systems show such a dispositional kind of reality.[141] Nevertheless, such dispositions are for us very real, otherwise the reality of most physical systems would dangerously evaporate. This is evident in the case of the effects exerted by a field. Then, nothing prevents us to consider also quantum observables as elements of reality. This statement should be carefully kept *distinct* from the other typical quantum problem, i.e. that observables may not commute. In fact, the complementarity of conjugate observables (and relative uncertainty relations) remains true but becomes not strictly pertinent to the issue of reality ascription to observables. In fact, the EPR argument could work also for a single measurement (of either position or momentum). Also in this case, we are faced with a possible action at a distance (Sect. 2.3.2).[142] The two measurement are only relevant to the issue of showing a failure of the uncertainty relations and, as a consequence, the incompleteness of QM. Thus, as mentioned, we need to disentangle these two aspects.

In this way, as previously anticipated, an important instance of the EPR argument is that to determine the way in which a system can behave in certain conditions *is an ascription of reality*. Thus, we can answer the question about the reality of observables raised in Sect. 2.3.1 in a positive way. However, I am trying to make

[140] Already in 1918 Einstein tells us that "no logical path leads from perceptions to the principles of the theory" (Einstein 1918, p. 109). See also Einstein (1930, p. 114). This conviction accompanies the great physicist through his whole life since still in 1952 he tells us that "there is, of course, no logical way to the establishment of a theory", quoted in Rindler (2001, p. 33). From this correct premise, Einstein infers that "all concepts, even those that are closest to experience, are from the point of view of logic freely chosen conventions" ("Alle Begriffe, auch die erlebnis–nächsten, sind vom logischen Gesichtspunkte aus freie Setzungen") (Einstein 1949a, pp. 12 and 13).

[141] As stressed by authoritative epistemologists (Hempel 1953, Nagel 1961). In Margenau (1950, Sect. 8.2) one speaks of "latency". See also D'Espagnat (1995, pp. 221–22), Bird (2007).

[142] Whitaker (2004), Maudlin (1994).

this ascription compatible with the constraints imposed by QM. The issue is not the maximal amount of realism allowed by QM but the minimal amount of realism that we need to drop according to QM. The definition of realism that catches EPR's main tenet and is sufficiently general to be accepted at this step is: *realism* is the standpoint through which we ascribe reality to certain phenomena when objective physical conditions are satisfied. This realism has to enrich the quantum objectivism that we are supporting here (Sect. 3.2.1).

Epistemic and Ontological Aspects

Nonetheless, by saying this, to a certain extent we are accepting Bohr's instance too, since the experimental arrangement that we could choose would indeed determine our reality attribution, as it is evident for conjugate observables, when the whole EPR thought experiment is considered. The crucial point, however, is that the context in which a certain observable is considered to be real is not just an epistemic one as Bohr assumed in his first 1935 paper, but both an *ontological* and *epistemic* context, and without actually selecting an ontological context (that is, without some fixed objective physical conditions, for instance, of experimental kind) we could certainly make no reality ascription at all or at most only in a conditional sense.[143] As I have stressed in the previous section, to be relative to a certain context does not imply to be dependent on a subjective consideration of the problem, and this is why I call these conditions *ontological*. Note, in fact, that such an operational context comes before detection and therefore also before any role of the observer. If we would limit our considerations to the epistemic (and therefore also formal) issues, we would simply say that the same state can be expanded in different bases, and therefore, we could not be authorised to speak of the reality of any of the observables that we could measure (Sect. 3.1.2). However, if we consider the premeasurement stage as a step, which, although still reversible, sets *additional* operational conditions on the system (making information potentially available), then we are no longer authorised to identify this physical step with the mathematical expansion in different bases, which only expresses a kind of abstract possibility. If we accept this, there is no way to reject EPR conclusion about the reality of observables. In this sense, we are speaking here of a quite different operational context relative to both that of the early Copenhagen interpretation (Sects. 1.2.4–1.3.3) that was centred on detection but also relative to the later epistemic version. Again, phenomenism needs to be corrected by a kind of realism for avoiding its reduction to the pure epistemic standpoint (and consequently to the subjectivist interpretation: Sect. 3.1.1).

[143] It is not by chance that in Bohr (1949, p. 230) Bohr speaks of "the necessity of considering the *whole* experimental arrangement".

Bohr's Second Paper

Actually, Bohr, in his discussion with EPR, considered also the actual experimental context, included detection, but only in the second paper.[144] Here, he came back to the interactionist view that he had previously developed with Heisenberg.[145] In such a case, it makes perfect sense to say with Bohr that we attribute a property *only on the basis of detection*. In other words, a certain observable being an element of reality does not imply that one of its possible properties (that is, one of the eigenvalues of the relative operator) is also a reality. In order to make this inference, we need not only an objective context but an *actual* event, for instance, a detection, as stressed by Bohr. Of course, *occurred* detection can be in turn considered as a physical context for attributing reality to certain observables; if this is the point of EPR is again correct. However, any confusion between dynamical interactions like detection and potential conditions should be avoided. It is important to remove any ambiguity about this point that could be present in their argumentation: in other words, here realism needs to be corrected with a well-understood phenomenism.

Now, what is curious is that, in this second paper, Bohr deals no longer with two systems in the sense of EPR (apart from a footnote) but with a single system under observation, and therefore the argument is impotent to answer the specific aspect of "prediction without disturbance" of EPR. It seems that Bohr did not realise the necessity to put these two aspects together. Summarising, Bohr answered EPR either

(i) Only on an epistemic plane or
(ii) On an operational plane but this time without consideration of the specific EPR argument.

Therefore, Bohr's objection, although important for refining the concept of physical reality, appeared not sufficient to demolish the EPR argument.

3.3.3 Schrödinger's Reply to EPR

The Discovery of Entanglement

In a series of papers,[146] Schrödinger replied to EPR by introducing for the first time the fundamental concept of entanglement in QM. He acknowledged this to be *the* characteristic trait of QM.[147] In the same papers he also proposed for the first time the thought experiment of Schrödinger's cat as a possible consequence of entanglement (see Sects. 2.2.5 and 3.1.4), It is entanglement that (in absence of detection) prevents,

[144] Bohr (1935b).

[145] Although doubts can be cast on whether Bohr had ever consequently supported an interactionist point of view (Stachel 2017).

[146] Schrödinger (1935a, b). See also Auletta and Wang (2014, Sect. 10.2).

[147] Schrödinger (1936).

in general, to attribute properties to a system or its subsystems (since it is a correlation among possible outcomes). When a detection event occurs, it is still entanglement that both

(i) Would allow attribution of properties in a way that is classically unknown and also
(ii) Would forbid other property ascriptions.

As stated by Schrödinger, "the best possible knowledge of a whole does not necessarily include the best possible knowledge of all its parts". For instance, having obtained spin-up on particle 1, we can predict with certainty that particle 2 will be found in spin-down (but not in spin-up). Therefore, the two aspects above cannot be separated, they are like two sides of the same coin: quantum predictions cannot be treated in classical terms but rely on the capability of far away systems to influence each other. But, and this is the core of Schrödinger's contribution, this influence *does not happen through exchange of signals*. In fact, he showed that quantum-mechanical systems can display non-local correlations that make them non-separable also in absence of any dynamical interaction between them, i.e. in full dynamical independence, and thus in disagreement with the separability condition. A word of warning is necessary here: by provisionally excluding that QM violates Einstein's locality (Sect. 2.3.2), from here onwards we can use the word "non-local" for denoting all global quantum phenomena in which there is non-separability when no confusion can arise.

A Model

Schrödinger proved his standpoint through apparently elementary quantum–mechanical considerations.[148] Suppose that we denote by $\hat{x}_1$ and $\hat{p}_1$ the 1D position and momentum of particle 1 into the x-direction, respectively, and by $\hat{x}_2$ and $\hat{p}_2$ the 1D position and momentum observables of particle 2 into the x-direction, respectively. Then, we can always build two other observables like

$$\hat{x} = \hat{x}_1 - \hat{x}_2 \qquad \text{and} \qquad \hat{p}_x = \hat{p}_1 + \hat{p}_2. \tag{3.102}$$

These two observables commute since

$$\begin{aligned} \left[\hat{x}, \hat{p}_x\right] &= (\hat{x}_1 - \hat{x}_2)(\hat{p}_1 + \hat{p}_2) - (\hat{p}_1 + \hat{p}_2)(\hat{x}_1 - \hat{x}_2) \\ &= i\hbar - i\hbar = 0. \end{aligned} \tag{3.103}$$

This means that they are jointly measurable and that we can predict the values of one of them from the values of the other one (since we have, e.g. the corresponding value $x = x_1 - x_2$) without involving any physical interaction at all (what would invalidate also Bohr's epistemic approach). Paraphrasing the words of Schrödinger (already reported in Sect. 2.3.1) it is as if entanglement would provide us of an answer

[148] Schrödinger (1936). For a summary see also Auletta (2000, Sect. 34.1).

to any question (what is another way to say that quantum systems represent infinite reservoirs of information). What is interesting in Schrödinger's model is that we can have violation of separability without any involvement of uncertainty relations, what shows that separability and reality are two completely independent assumptions.

Possible and Actual Detections

There is, in fact, a subtle difference to point out here, which is evident also from the example. Schrödinger seems to have never treated the question in terms of *actual* detection events, but to deal experimentally with observables like those defined in Eqs. (3.102) is far from being trivial. Since he had a continuous, wave-like understanding of quantum physics, he could only attribute properties to the two distant systems in a conditional and *hypothetical* sense, as remarked in Sect. 3.1.2: given the entangled state, any property of one system, supposed that is found to be real, allows a prediction of another property about the distant system. As we know, to him formal aspects appeared to exhaust what reality is. We recover here the different approaches that are behind the Schrödinger and Heisenberg pictures, as pointed out in Sect. 2.1.1. In fact, by synthesising the view expressed in his reply to EPR, the historians of physics J. Mehra and H. Rechenberg say[149]:

> QM replaces the causal relations by a particular statistics: It allows one to compute, from the maximal number of completely determined characteristics, the 'statistical distribution' of every variable at a given instant of time and at any later instants. [...] the new theory thus declared the classical model as being unable to represent the mutual connection of the characteristics (*Bestimmungsstücke*).

However, Schrödinger never says that the whole system (locally) collapses on one or the other property, property attribution remaining only conditional (it is a counterfactual statement), as it was for Bohr in his first paper, although for different reasons. Thus, although the discovery of entanglement represents in fact a rejection of the separability principle, nothing is said about reality: (i) Schrödinger tells nothing about the fundamental issue whether observables can be considered real in certain physical conditions, but (ii) conditional property ascriptions of the kind mentioned are in fact irrelevant to the EPR argument, as already pointed out in the previous subsection. It is here that both the Schrödinger's interpretation and the MWI seem insufficient to deal with the EPR argument: according to the results of the previous subsection, quantum objectivism (Sect. 3.2.1) need to be corrected by a well-understood realism.

Three Forms of Inference

It is interesting to finally consider the different forms through which the actors of this debate developed their argumentation. EPR developed their argument in deductive form. In finding entanglement as a resolution to a possible conflict between a

[149] Mehra and Rechenberg (1982, VI, p. 744).

thought experiment and quantum-mechanical laws, Schrödinger was performing an abduction, while Bohr developed his argument in inductive form.[150] As we shall consider with more details later, abduction deals with the discovery of new properties or characters that allow us to solve a problem in the framework of a previous theory (here, QM), while induction deals with cumulation of experimental evidence for rejecting some theory or general assumption (here, ascription of reality). The final result of abduction is more constructive in the framework of an existing theory, while induction is a kind of no-go theorem: the impossibility to accepts certain general conclusions and the necessity to give rise to a new theoretical framework (which in Bohr's mind was the epistemic Copenhagen interpretation).

In the following, we shall deepen this examination and deal with the main consequences of (i) affirming a rejecting the postulate of classical reality as well as of (ii) affirming and rejecting separability. Let us first discuss reality.

3.3.4 Hidden Variables

Classical Realism

As we know, the aim of EPR was to prove that QM was incomplete: they assumed that quantum systems display real observables and likely real properties (elements of reality), and one may be tempted to say that these are established without any recourse to measurement (given that it is the other system to undergo detection). Now, if the EPR thought experiment could prove such a result, the consequence would have been the search for appropriate variables able to explain in classical deterministic terms what in QM was finally a statistical (and therefore phenomenological) treatment of the micro-realities (Sects. 2.3.1 and 2.4.1). Since it was commonly assumed that these variables underlying quantum phenomena and rounding off the description of quantum theory were not yet detected at that time because of their smallness, the physicists' community baptised them "hidden variables".[151] I define the hidden variables with respect to a theory as that set of variables of a system which are not experimentally detectable by means of the set/array of the conceptual instruments of the theory, although allowing to obtain the correct values of the whole of the observables that describe the pertinent physical systems.[152] A theory of hidden variables (HVs) not only supposes that the system's state and its observables are perfectly determined (this is the *omnimoda determinatio* assumption), but also it is its dynamic evolution (and this is classical determinism). In this way, some of Ein-

[150] As proposed in Auletta (2007).

[151] For historical survey see Jammer (1974, Chap. 7). According to Jammer, hidden variables are one of the most recent attempts at explaining visible things with invisible ones. Einstein himself may have initially contributed (Home and Whitaker 2007, p. 89), although is engagement appears modest and he was never particularly supportive of this research project (Jammer 1974, pp. 254–55).

[152] Auletta (2000, p. 543).

stein's followers, overcoming the incertitudes in the position of the great physicist, adopted a realistic view about properties. This standpoint is *local* or also *classical realist*. Although it may appear reasonable to say that all theories of HVs should agree on the basic assumption that all ontological entities subject to these theories do not depend on measurement, or in general, on interaction,[153] I mention that also theories in which HVs are contextual have appeared.

de Broglie's Empty Wave

In Sect. 2.3.1 I have mentioned that Bohm's idea was to provide a model of the EPR paradox easier to implement. The Anglo-American physicist also set forth a theory of HVs, strongly criticising the interpretation of Copenhagen, according to which the wave function would constitute the best (and complete) description of the quantum entities. Bohm had in mind the building of a classically deterministic theory of the quantum-mechanical world, but also set out to guarantee some objectivity independently of the observer, something that the Copenhagen interpretation, apparently both in its epistemic and interactionist versions, rejected.[154]

Bohm's theory was the follow-up of de Broglie's pilot wave, a development of some insights in the early work of the French physicist, as recalled in Sect. 1.1.3.[155] According to de Broglie, a real particle is embedded in a 'quantum field' that confers to it the specific interference behaviour that we observe. Note that both the particle and the field are purely classical and it is only their combination that gives rise to the observed quantum-mechanical interference phenomena in the particle's trajectory. I remark that there are two noticeable consequences of this interpretation that have been experimentally disproved in the subsequent years[156]:

- The interactions between wave and particle should give rise to *nonlinear effects*.[157]
- There would be some situations in which there is a wave and not a particle, what is called an *empty wave*: for instance, if we assume that a classical particle always takes one of the two paths of an interferometer, we are forced to say that the other path is taken by an empty wave (Fig. 3.8).[158]

[153] Auletta (2000, p. 544).

[154] Bohm (1952).

[155] de Broglie (1956). On this issue see also Auletta (2000, Sect. 28.1), Auletta et al. (2009, Sect. 16.3).

[156] See Auletta (2000, Sects. 28.2–28.3).

[157] As proposed in Bialynicki and Mycielski (1976). Experimental tests disproved this hypothesis (Shull et al. 1980).

[158] The original idea was formulated in Selleri (1969). There is wide literature on that issue (Garuccio et al. 1982; Croca 1987; Hardy 1992; Zou et al. 1992; Auletta and Tarozzi 2004). See also Fano (2014).

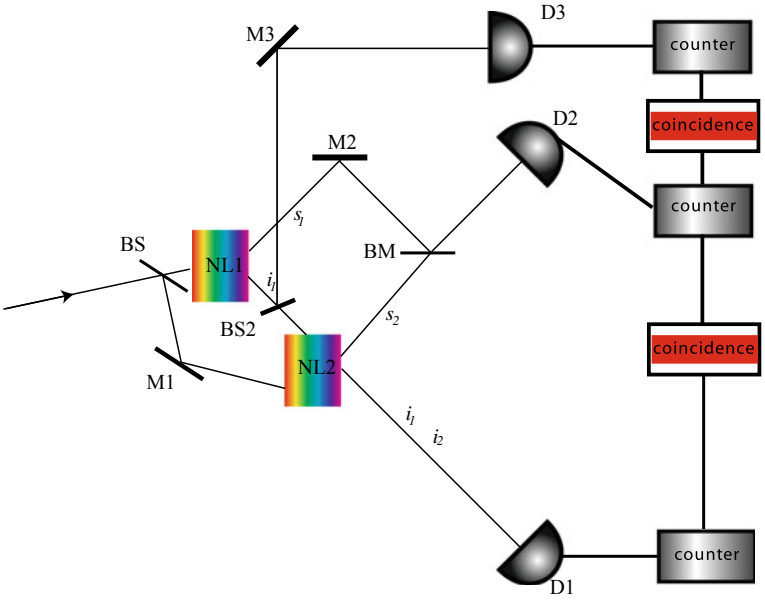


Fig. 3.8 Schematic representation of the experiment on empty waves performed by the team of the German-American physicist Leonard Mandel (1927–2001). An initial photon beam is split by a beam splitter (BS) and the resulting beams, after being deflected by the mirror M1, travel towards two nonlinear crystals (NL1 and NL2). The NLs produce a couple of photons, conventionally called idler and signal. From NL1 the signal photon s_1 and the idler photon i_1 emerge: the i-photon passes through NL2 and will be aligned with the second idler (i_2), which is emitted by NL2 together with the second signal photon (s_2). The two s-photons are combined by the mirror M2 and the beam merger (BM), and the outgoing beam falls on detector D2, whereas the two idler photons fall on detector D1. The second beam splitter (BS2) sends output beams towards the second nonlinear crystal (NL2) and towards the third detector (D3). Note that between detectors there are coincidence counters for determining the time of arrival of photons (in fact the paths are of different length). Now, if detector D3 clicks, from the point of view of the empty wave theory an empty wave (i_1) should fall on D1 and still induce coherence between s_1 and s_2. Experimental results showed no coherence in this case, which supports the interpretation of quantum waves as probability amplitudes and not as real waves. Adapted from Auletta et al. (2009, p. 319)

Bohm's Quantum Potential

Bohm refined this framework and changed the interpretation by putting the stress on HVs.[159] Bohm's theory of the pilot wave consists of a data field that fixes unerringly the trajectory of each particle. This data field depends on the concept of quantum potential developed by Bohm himself. In fact, the quantum potential is a formal consequence of the Schrödinger equation expressed in 2D spherical coordinates or polar coordinates (1.255a) (see also Fig. 1.24, Sect. 1.2.5), when the wave function is separated in its real and imaginary parts.

[159] See Jammer (1974, Sect. 7.5).

In the simplest case, we can represent the wave function through (i) a function $\vartheta(r)$ expressing the amplitude and (ii) an exponential $e^{\frac{i}{\hbar}\phi(r)}$, expressing the relative phase (I have omitted time dependency for the sake of notation). Thus, Bohm started with the wave definition[160]

$$\psi(r) = \langle r \mid \psi \rangle = \vartheta(r)\cos(\phi(r)\hbar^{-1}) + i\vartheta(r)\sin(\phi(r)\hbar^{-1}) = \vartheta(r)e^{\frac{i}{\hbar}\phi(r)} . \tag{3.104}$$

Let us now consider a 3D Schrödinger equation (1.158) with a potential $V(r)$:

$$i\hbar\frac{\partial\psi(r)}{\partial t} = \left(-\frac{\hbar^2}{2m}\nabla^2 + V(r)\right)\psi(r), \tag{3.105}$$

where the nabla (or del) operator is given by Eq. (1.155). Inserting Eq. (3.104) into Eq. (3.105) and multiplying both sides by $e^{-\frac{i}{\hbar}\phi(r)}$ (in order to drop the exponential terms), we get:

$$i\hbar\left[\frac{\partial\vartheta}{\partial t} + \frac{i\vartheta}{\hbar}\frac{\partial\phi}{\partial t}\right] = -\frac{\hbar^2}{2m}\left[\nabla^2\vartheta - \frac{\vartheta}{\hbar^2}(\nabla\phi)^2 + i\left(\frac{2}{\hbar}\nabla\vartheta\cdot\nabla\phi + \frac{\vartheta}{\hbar}\nabla^2\phi\right)\right] + V\vartheta, \tag{3.106}$$

where I have used the product rule (1.206)–(1.207). Now, we separate the real and imaginary parts of the resulting equation (see Eq. (1.9)), where both ϑ and ϕ themselves are assumed to be real. To this purpose, we collect the second term on the LHS and the first, the second and the latter term of the RHS for the real part

$$\frac{\partial\phi}{\partial t} = -\left[\frac{(\nabla\phi)^2}{2m} + V(r) - \frac{\hbar^2}{2m}\frac{\nabla^2\vartheta}{\vartheta}\right] , \tag{3.107a}$$

where I have divided the two sides by ϑ (assuming $\vartheta \neq 0$), and collect the first term on the LHS and the third and the fourth ones on the RHS for the imaginary part

$$\frac{\partial\vartheta}{\partial t} = -\frac{1}{2m}\left[\vartheta\nabla^2\phi + 2\nabla\vartheta\cdot\nabla\phi\right] , \tag{3.107b}$$

where I have divided the two sides by $\hbar$. In the classical limit (where the last term of the RHS of Eq. (3.107a) goes to 0), the phase ϕ is a solution of the classical *Hamilton–Jacobi equation* (after the names of W. Hamilton and C. Jacobi)

$$\frac{\partial\phi}{\partial t} + \frac{(\nabla\phi)^2}{2m} + V(r) = 0 . \tag{3.108}$$

In this case, we can consider the probability density $\wp(r) = \vartheta^2(r)$ as a classical (stochastic) probability density for ensembles of particles (for the notion of proba-

[160] On this subject I recommend the extensive book (Holland 1993). For understanding Bohm's quantum potential theory I recommend (Bricmont 2016).

bility density see Sect. 1.2.3) and ϕ as the action (1.23). By multiplying both sides of Eq. (3.107b) by ϑ, we obtain

$$\frac{\partial}{\partial t}\vartheta^2 + \boldsymbol{\nabla}\left(\vartheta^2 \frac{\boldsymbol{\nabla}\phi}{m}\right) = 0, \tag{3.109}$$

which is called the *continuity equation* and expresses the transport of a conserved quantity (with many applications, both in classical and QM, from fluid dynamics to electromagnetism).[161] Of course, we deal here with a flux of particles.

Instead of neglecting the last term in the RHS of Eq. (3.107a)—as it is the case for the classical approximation—let us reformulate it as a quantum counterpart of the Hamilton–Jacobi equation for an ensemble of particles, the so-called *Bohm–Jacobi equation*

$$\frac{\partial \phi}{\partial t} + \frac{(\boldsymbol{\nabla}\phi)^2}{2m} + V(r) + V_Q(r) = 0\ , \tag{3.110}$$

where

$$V_Q(r) := -\frac{\hbar^2}{4m}\left[\frac{\boldsymbol{\nabla}^2 \wp}{\wp} - \frac{1}{2}\frac{(\boldsymbol{\nabla}\wp)^2}{\wp^2}\right] = -\frac{\hbar^2}{2m}\frac{\boldsymbol{\nabla}^2\vartheta}{\vartheta}. \tag{3.111}$$

After a comparison with Eq. (3.105), this makes evident that the particles are acted upon not only by the 'classical' potential energy $V(r)$, but also by the quantum potential energy $V_Q(r)$. Due to the form of the potential, and in particular to the fact that it remains unchanged when multiplied by an arbitrary constant (since the amplitude appears both in the numerator and denominator of its expression), it was soon clear to Bohm and his followers that it could not represent a form of energy (as we would expect from a true potential). For this reason, he ultimately interpreted it as *active information*, and made it depend only on the shape (the formal structure) of the wave but not on its amplitude.[162] Consequently, the theory is presented as contextual (as far as its deterministic evolution is concerned but not necessarily meaning that also properties are contextual) and holistic. Indeed, the shape of the pilot wave depends on the configuration of the total environment, and thus the motion of each of the particles depends on the changes in the latter. These statements will turn out to be crucial. However, if the piloting model (a remnant of de Broglie's pilot wave) is maintained with an *active* form of guiding, it would require, at a physical level, a kind of aether or a field of connections of trajectories with velocities higher than that of the light, a consequence drawn by Bohm himself (see also Sect. 2.3.2).

Thus, such a theory becomes strongly non-local, given that it supposes that the particles' behaviour depends on all the other particles' dynamics, no matter how far

[161] On this subject see Auletta et al. (2009, Sect. 10.5.3).

[162] Bohm (1980), Bohm and Hiley (1993), Holland (1993), Hiley (1999). See also Auletta (2000, Sect. 32.6). Bohm proposed that the multidimensionality of information should correspond to a certain extension of the Hilbert space's structure.

they could be.[163] A further modification of the theory was aimed at understanding these non-local effects as in fact local effects in a universal system of reference, effects that nonetheless would take place at velocities higher than that of the light.[164] Here, it has also been raised the objection that there could be a difference between the trajectories predicted by the Bohmian theory and the observed ones.[165]

Aharonov–Bohm Effect

Although I do not follow such a strong non-local tenet, this interpretation could be considered as convergent with the point of view supported here.[166] In fact, the so-called quantum potential is a manifestation of the quantum interdependence among systems that I consider crucial for the theory (see Sect. 2.2.2). Of course, we need to renounce the idea that it is a potential and therefore that it is a thread connecting classical (real) trajectories for the reasons already exposed in Sect. 2.1.2: rather it should be a connection among the states (and relative observables) of the quantum systems and therefore among *possible* trajectories. This would be consonant with the interpretation of the 'potential' in terms of information (Sects. 3.2.2–3.2.3), although we should drop the term 'active' since it could result misleading as far as it would reintroduce the idea of some field piloting the particles. In fact, one of Bohm's followers said that it is information about the whole of the context in which the system is, and this can be brought in accordance with our examination.[167] An example could clarify this point: the Aharonov–Bohm (AB) effect, after the names of D. Bohm and the Israeli physicist Yakir Aharonov.[168] This consists in the fact that a beam of particles passing on the left and right side of an isolated magnetic or electric field show phase difference (Fig. 3.9). The two physicists explained this effect in terms of the quantum potential understood as an 'ordinary' physical potential. However, it can be easily accounted for in terms of the so-called geometric (or Berry's) phase, and therefore the geometric properties of the quantum "waves" in relation with a certain environment, without consideration of the nature of the forces, energies or fields involved.[169] I shall come back on the notion of geometric phase.

---

[163] This seems related to the fact, pointed out by the German physicist Walther Bothe (1891–1957), that, even when wave functions of the two EPR particles are factorised (that is, are in a product state of the kind (1.388)), the HVs of the two systems can be still mutually dependent (Home and Whitaker 2007, p. 89).

[164] On this see Auletta (2000, p. 561).

[165] Englert et al. (1994).

[166] One has spoken of exorcised Bohmian theory (Conway and Kochen 2006, p. 1454).

[167] Hiley (1999).

[168] Originally proposed in Aharonov and Bohm (1959). There is a now a wide literature in this subject and a helpful textbook is Peshkin and Tonomura (1989).

[169] Berry (1984, 1987). For a summary see also Auletta et al. (2009, Sect. 13.8).

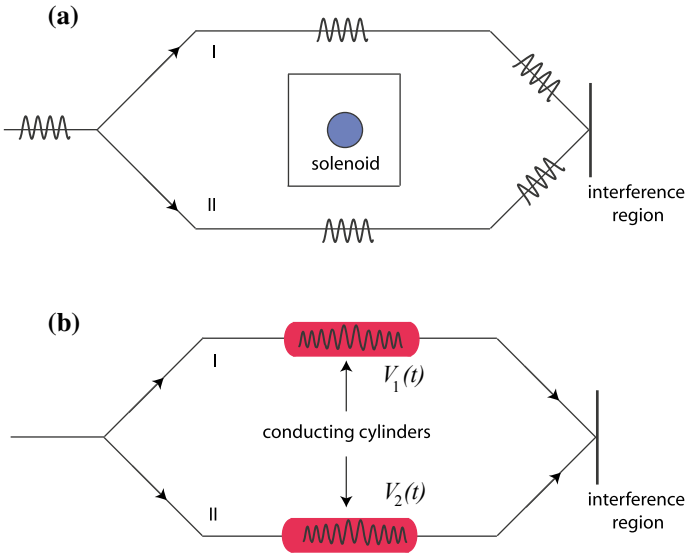


Fig. 3.9 **a** Magnetic AB effect. The axis of the solenoid is perpendicular to the page. The wave function is a split plane wave. **b** Electric AB effect. $V_1 = V_2 = 0$ except when the wave packet is shielded from the electric field. Adapted from Auletta et al. (2009, p. 497)

No Dispersion-Free States

In 1932, before the elaboration of Bohm's proposal (and even of the EPR paper), John von Neumann anticipated such developments and presented a theorem that apparently proved the impossibility of HVs,[170] although, as we shall see, the proof cannot be considered as conclusive. His starting point was that it is experimentally impossible to select states without dispersion corresponding to the values of the two conjugated observables. With *dispersion-free observables* we mean observables with zero variance, where I recall that the variance is the square of the uncertainty (Sect. 1.2.4), while dispersion-free states are those displaying only dispersion-free observables (according to the *omnimoda determinatio* assumption). Such impossibility could be attributed to the disruptive nature of measurements of either variable on the other one (see also Sect. 2.2.2). However, we can figure out that the values of the observables exist before being disturbed (as assumed by EPR), that is, we can assume that sets without dispersion exist independently of any experimental reference. In this way, ideally, we can think of the sets without dispersion ascribing to their elements the precise value determined by one or more HVs. This procedure will only be correct if, according to von Neumann, it is possible to ascertain that any set with dispersion can be represented as a blend of sets without dispersion. The crucial assumption of von Neumann, in the words of his critic J. Bell, can be sum-

[170] Von Neumann (1932, pp. 163–71).

marised as: "Any real linear combination of any two Hermitian operators represents an observable, and the same linear combination of expectation values is the expectation value of the combination."[171] On the basis of such an assumption, the theorem aims at proving that quantum predictions of sets with dispersion cannot be depicted as blends of sets without dispersion, so that, consequently, a pure state cannot be split in subsets that correspond to the states determined by HVs. On this point, Bell pointed out that the mathematical possibility to combine operators does not imply the linear combination of the expectation values and therefore also of eigenvalues (an argument that could to a certain extent also be applied to Schrödinger's reply to EPR (Sect. 3.3.3) when generalised)[172]: the reason is that combining different possible experimental results in specific sets may be not trivial since they require different experimental arrangements, as already Bohr had pointed out (Sect. 3.3.2). This shows that to be a Hermitian operator is only a necessary condition for defining a physical observable, i.e. a quantity that can be experimentally measured. We find again a difference between the formal and experimental aspects.

Bell's Contribution on Dispersion-Free States

Although von Neumann was not able to develop a definitive proof of the impossibility of dispersion-free states able to block the development of HV theories, his contribution stimulated the work of many physicists and mathematicians thereafter. Bell himself, by making use of a theorem due to the American mathematician Andrew Gleason (1921–2008),[173] was able to perform a successful proof that, for a Hilbert space of dimension ≥ 3, there are no dispersion-free states that (as a consequence of normalisation) satisfied the quantum-mechanical requirement of additivity of expectation values of projectors pertaining to the same set[174]:

$$\begin{aligned}\sum_j \left\langle \hat{P}_j \right\rangle_\varphi &= \sum_j \left\langle \varphi \left| \hat{P}_j \right| \varphi \right\rangle \\ &= \sum_j \langle \varphi \mid j \rangle \langle j \mid \varphi \rangle \\ &= \sum_j \wp_j \\ &= 1 \end{aligned} \tag{3.112}$$

for any normalised state $|\varphi\rangle$, in agreement with Eq. (1.101). In fact, he succeeded in proving that, given two arbitrary (pointing in whatever direction and not necessarily normalised) state vectors $|\psi\rangle$, $|\varsigma\rangle$, they cannot be arbitrarily close since we have the

[171] Bell (1966).

[172] Bell (1966, 1971).

[173] Gleason (1957). See also Landsman (2017, Sects. 2.7–2.8 and 4.4), Auletta (2000, Sect. 11.5).

[174] Bell (1966).

following relation between the norms (see Eq. (1.72))

$$\| \, |\varsigma\rangle - |\psi\rangle \, \| > \frac{1}{2} \, \| \, |\psi\rangle \, \| \, , \tag{3.113}$$

which shows that there cannot be dispersion-free states (that is, observables with perfectly determined values) because we have value gaps.[175] As a consequence of Eq. (3.112), note that, for all orthogonal $|\psi\rangle$ and $|\varsigma\rangle$, if

$$\left\langle \hat{P}_{\psi} \right\rangle_{\varphi} = 1, \tag{3.114}$$

then

$$\left\langle \hat{P}_{\varsigma} \right\rangle_{\varphi} = 0. \tag{3.115}$$

Since in the following I shall always consider only mean values taken on the same state $|\varphi\rangle$, I shall drop reference to the latter. Moreover, if for any orthogonal vectors $|\psi_1\rangle, |\psi_2\rangle$, we have

$$\left\langle \hat{P}_{\psi_1} \right\rangle = \left\langle \hat{P}_{\psi_2} \right\rangle = 0, \tag{3.116a}$$

then it follows that

$$\left\langle \hat{P}_{\psi} \right\rangle = 0, \tag{3.116b}$$

with $|\psi\rangle = c_1 \, |\psi_1\rangle + c_2 \, |\psi_2\rangle$, $\forall c_1, c_2 \in \mathbb{C}$. Let us first assume that indeed

$$\left\langle \hat{P}_{\psi} \right\rangle = 1 \text{ and } \left\langle \hat{P}_{\varsigma} \right\rangle = 0. \tag{3.117}$$

If $|\psi\rangle$ and $|\varsigma\rangle$ are orthogonal, Eq. (3.113) is automatically satisfied. Suppose in fact that $\| \, |\varsigma\rangle \, \| = a$ and $\| \, |\psi\rangle \, \| = b$, where $a, b \in \mathbb{R}$. Then, whatever numbers are a, b, we obviously have $a^2 + b^2 > b^2/4$. Let us consider the case in which they are not. Assume in this case that

$$\left|\varsigma'\right\rangle = \left|\psi'\right\rangle + \epsilon \, |\psi_{\perp}\rangle \, , \tag{3.118}$$

where, in agreement with Eq. (1.73)

$$\left|\psi'\right\rangle = \frac{|\psi\rangle}{\| \, |\psi\rangle \, \|} \qquad \text{and} \qquad \left|\varsigma'\right\rangle = \frac{|\varsigma\rangle}{\| \, |\psi\rangle \, \|}, \tag{3.119}$$

$|\psi_{\perp}\rangle$ is some state vector orthogonal to $|\psi\rangle$ and normalised, and ϵ is some quantity that could be very small (since $\left|\varsigma'\right\rangle$ or $|\varsigma\rangle$ are not necessarily normalised, ϵ could be quite arbitrary). Taking into account a state vector $|\psi_{\perp\perp}\rangle$ that is orthogonal to both $|\psi\rangle$ and $|\psi_{\perp}\rangle$, as a consequence of the first of Eqs. (3.117), we have

[175] Auletta (2000, Sect. 32.3), Auletta et al. (2009, Sect. 16.4.1).

$$\left\langle \hat{P}_{\psi_\perp} \right\rangle = \left\langle \hat{P}_{\psi_{\perp\perp}} \right\rangle = 0. \tag{3.120}$$

However, using again Eqs. (3.117), this in turn implies

$$\left\langle \hat{P}_x \right\rangle = \left\langle \hat{P}_{x_\perp} \right\rangle = 0, \tag{3.121}$$

where

$$|\, x\rangle \;=\; \left|\, \varsigma' \right\rangle + a^{-1}\epsilon\, |\, \psi_{\perp\perp}\rangle \qquad \text{and} \qquad |\, x_\perp\rangle \;= -\epsilon\, |\, \psi_\perp\rangle \;+ a\epsilon\, |\, \psi_{\perp\perp}\rangle \tag{3.122}$$

are orthogonal states and $a \in \mathbb{R}$. In fact, $\langle x_\perp \mid x\rangle = \epsilon^2 - \epsilon^2 = 0$. As a consequence of Eqs. (3.121) and taking again into account Eq. (3.116), we have

$$\left\langle \hat{P}_y \right\rangle = 0, \qquad \text{with} \qquad |\, y\rangle \;=\; |\, x\rangle \;+\; |\, x_\perp\rangle\ . \tag{3.123}$$

Now, suppose that $\epsilon < 1/2$. Then, mathematics ensures us that for all such quantities there is a real number a such that $\epsilon(a + a^{-1}) = \pm 1$. Now, since according to Eqs. (3.118) and (3.122), we have

$$|\, y\rangle \;=\; \left|\, \psi' \right\rangle + \epsilon(a + a^{-1})\, |\, \psi_{\perp\perp}\rangle\ . \tag{3.124}$$

this means that we deal in fact with two orthogonal states, $|\, y_+\rangle$ and $|\, y_-\rangle$, according to the sign (direction) of the second component of the superposition:

$$|\, y_\pm\rangle \;=\; \left|\, \psi' \right\rangle \pm\, |\, \psi_{\perp\perp}\rangle\ . \tag{3.125}$$

Clearly, we also have $\left\langle \hat{P}_{y_+} \right\rangle = \left\langle \hat{P}_{y_-} \right\rangle = 0$. However, since

$$\left|\, \psi' \right\rangle \;=\; |\, y_+\rangle \;+\; |\, y_-\rangle\ , \tag{3.126}$$

it follows that

$$\left\langle \hat{P}_{\psi'} \right\rangle = 0, \tag{3.127}$$

in contradiction with assumptions (3.117). This proves that we must have $\epsilon > 1/2$, so that, from Eq. (3.118), it follows that

$$\|\, \left|\, \varsigma' \right\rangle - \left|\, \psi' \right\rangle \,\| \;>\; \frac{1}{2}\, \|\, |\, \psi_\perp\rangle \,\|\ . \tag{3.128}$$

Multiplying both sides by $\|\, |\, \psi\rangle \,\|$ and since by hypothesis the norm of $\|\, |\, \psi_\perp\rangle \,\| = 1$, we have proved Eq. (3.113) also for non-orthogonal states.

Kochen–Specker Theorem

When we try to build a classical reformulation of QM based on HVs, this needs to be classically deterministic, and therefore, as recalled at the beginning of this subsection, needs also to justify ascription of properties according to the *omnimoda determinatio* assumption (Sects. 1.2.1 and 1.2.4). The work of Bell, previously examined, already excluded a classical assignment of properties. Nevertheless, there is a proof on the same outline that is more powerful from a logical point of view although less specific from a physical point of view. The American mathematician S. Kochen and the Swiss mathematician E. Specker developed Bohr's main idea (Sect. 3.3.2) by means of an elegant logic model.[176] With a system of 117 propositions, the two scholars showed that there are statements that are classically false but quantum-mechanically true. Let $\{a_1, a_2, \ldots, a_n\}$ be a set of n atomic statements (an atomic statement is a statement that cannot be broken down into other simpler statements). Atomic statements of this kind express in our framework property ascription (and thus can be considered to be counterparts of projectors). Moreover, b_{ij} and c_{ijk} be compound statements of the form

$$b_{jk} = (a_j a_k)', \quad c_{ijk} = a_i + a_j + a_k, \tag{3.129}$$

where I recall that $+$ denotes the inclusive disjunction (OR) whose truth table is Table 3.4, Sect. 3.3.5. Note that the presence of c statements makes also of this model a 3D one. Let us use De Morgan's law[177] (after the name of the British logician Augustus De Morgan (1806–1871))

$$(XY)' = X' + Y', \quad (X + Y)' = X'Y'. \tag{3.130}$$

Now, we have $b_{jk} = a'_j + a'_k$, which means that classically the disjunction of a'_j and a'_k can be true if at least a_j or a_k is false. At the opposite, in QM the statement b_{ij} can be true even if both a_j and a_k are true. This implies, in QM, that also the statement c_{ijk} can be true even if a_i, a_j, and a_k are all false. Let us see why.

We first consider a statement d composed by the *conjunction* of 18 compound statements involving 10 atomic statements a_j (with $j = 0, 1, \ldots, 9$) as a whole:

$$d = b_{01}b_{02}b_{08}b_{13}b_{15}b_{24}b_{26}b_{35}b_{37}b_{46}b_{47}b_{56}b_{78}b_{79}b_{89}c_{135}c_{246}c_{789}. \tag{3.131}$$

The graphical representation of the statement d is depicted in Fig. 3.10, in which the vertices are the atomic statements. Two vertices a_j and a_k are connected by a straight line iff b_{jk} is in d, while the vertices a_i, a_j, and a_k constitute a triangle iff c_{ijk} is in d. Let the function $f(a)$ denotes the truth value of the statement a. If $f(d) = 1$, then, classically, no pair of atomic statements connected by a straight line can have truth value 1. This is because the statement d is false if one of its constituent statements

[176] Kochen and Specker (1965). See also Pitowsky (1989), Auletta and Wang (2014, Sect. 10.7), Landsman (2017, Sect. 6.1).

[177] De Morgan (1847).

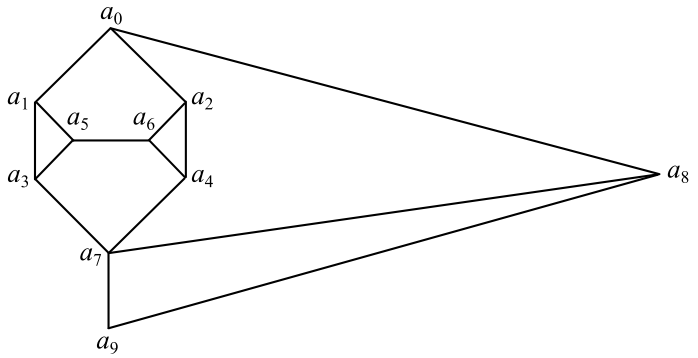


Fig. 3.10 The 10-node Kochen–Specker graph, i.e. the graphical representation of the statement d given by Eq. (3.131). Adapted from Auletta and Wang (2014, p. 326)

b_{jk} is false. Formally, we have the proposition

$$[f(a_j) = f(a_k) = 1 \rightarrow f(b_{jk}) = 0] \rightarrow f(d) = 0. \tag{3.132}$$

Indeed, on the one hand, the disjunction b_{jk} is true only if at least one of the two disjoints a'_j and a'_k is true (or only if a_j or a_k is false). On the other hand, if b_{jk} is false, then the compound statement d is false as well, as the latter is composed of conjunctions between b_{jk}'s and c_{ijk}'s, which implies that neither b_{jk} nor c_{ijk} in d can be false (Table 3.3, Sect. 3.3.5). Now we shall prove the following result:

$$(f(d) = 1)(f(a_0) = 1) \rightarrow (f(a_9) = 1), \tag{3.133}$$

where the function f is again a truth value assignment. In other words, if both d and a_0 are true, also a_9 must hold true. A quick proof by *reductio ad absurdum* can be obtained by inspecting the graph of d depicted in Fig. 3.10. Since a_0 is true by hypothesis, we have that all the statements connected to it through a straight line are classically false, that is, a_1, a_2, and a_8 are all false. Now, if we assume that a_9 is also false, this then implies that a_7 must be true since $c_{789} = a_7 + a_8 + a_9$ must hold true if d is true. However, if a_7 is true, then a_3 and a_4 must be false because both are connected with a_7 by a straight line; and, since both c_{135} and c_{246} must be true, this implies that both a_5 and a_6 are true. But this is a contradiction since the a_5 and a_6 are connected by a straight line. Then, the assumption that a_9 is false is falsified and the previous formula proved.

With the above result, we now proceed to the formulation and proof of the Kochen–Specker theorem. Consider a compound statement e constructed as follows by using 15 copies of the d graph considered above (Fig. 3.11). There are three groups of five interlocking copies of the d graph such that, in the first group, the statement a_9 of one copy of d is identified with the a_0 of the next copy and all copies of a_8 in the same group are identified with each other. Let us now denote by $d_{0,8,9}$ the statement given

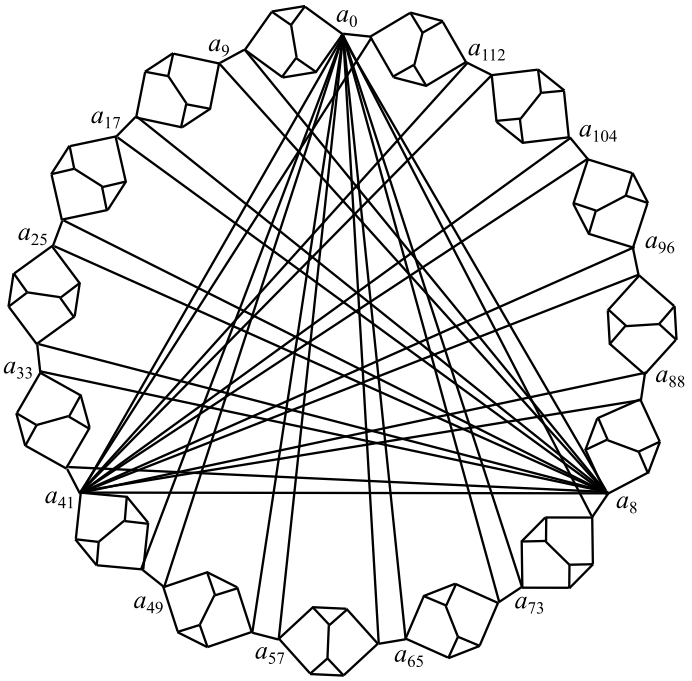


Fig. 3.11 The 117-node Kochen–Specker graph, i.e. the graphical representation of the statement e given by Eq. (3.135), is constructed by using 15 copies of the 10-point Kochen–Specker graph depicted in Fig. 3.10 (consider especially the subgraph containing a_0, a_8 and a_9). The vertices are numbered consecutively in a counterclockwise direction from a_0 to a_{119} except for a_8 and a_{41}. Moreover, vertices a_0 and a_{40} are identified with each other, and so are a_8 and a_{119}, and a_{41} and a_{80}. Adapted from Auletta and Wang (2014, p. 327)

by Eq. (3.131) comprehending all ten atomic statements from a_0 up to a_9, where the subscripts 0, 8 and 9 are the vertices a_0, a_j, and a_k, with j, k to be identified with those of the next copy in the above construction. To avoid confusion, here I have used commas to separate numerical subscripts. Thus, we have the logical form

$$d_{0,8,9}d_{9,8,17}d_{17,8,25}d_{25,8,33}d_{33,8,41}. \tag{3.134}$$

Note that a_0, a_8, and a_{41} constitute a triangle, so that we can add $c_{0,8,41}$. In order that this whole statement is true all the five main components (i.e. the d's) need to be true.

Then, the three groups are cyclically interlocked together by identifying the statement having the function of a_9 of the last copy of d with the statement having the function of a_0 of the first copy of d in the next group. Then, the resulting compound statement e can be expressed as

$$e = (d_{0,8,9}d_{9,8,17}d_{17,8,25}d_{25,8,33}d_{33,8,41})(d_{41,0,49}d_{49,0,57}d_{57,0,65}d_{65,0,73}d_{73,0,8})$$
$$(d_{8,41,88}d_{88,41,96}d_{96,41,104}d_{104,41,112}d_{112,41,0})c_{0,8,41}, \tag{3.135}$$

where I have put the three different groups between parentheses. Assume now that e is true. This implies that the statement $c_{0,8,41}$ must be true, which in turn implies that at least one statement among a_0, a_8, a_{41} must be true. Suppose that $a_0 = 1$. Then, also a_9 must be so, and so on reiteratively along the whole cycle

$$a_0 - a_9 - a_{17} - a_{25} - a_{33} - a_{41} - a_{49} - a_{57} - a_{65} - a_{73} - a_8 - a_{88} - a_{96} - a_{104} - a_{112} - a_0. \tag{3.136}$$

However, a_0 and a_{41} are connected by an edge and cannot be both true. This is a contradiction. Similar contradictions can be derived by assuming that a_8 or a_{41} is true. In fact, in these two cases we cycle again through all involved statements but starting at two different points. This proves the classical logical falsity of statement e.

A Quantum Counterexample

We are now in the position to formulate the Kochen–Specker theorem in terms of the compound statement that we have just constructed: the statement e is classically a logical falsity, but there are cases in which it is true in QM. To prove this, consider all of the 117 atomic statements a_j meaning "The squared spin component of the particle in the direction a_j is zero", where we are considering so-called spin 1 particles whose squared components (along x, y, z) $\hat{S}_x^2$, $\hat{S}_y^2$, and $\hat{S}_z^2$ are all commutable (while $\hat{S}_x$, $\hat{S}_y$, and $\hat{S}_z$ themselves are not). In the common eigenbasis of $\hat{\mathbf{S}}^2$ and $\hat{S}_z$, the operators $\hat{S}_x$, $\hat{S}_y$, and $\hat{S}_z$ are, respectively, represented by the three 3×3 matrices given by Eq. (1.331). Thus, the corresponding squared spin components are given by (in units of $\hbar^2$)

$$\hat{S}_x^2 = \frac{1}{2}\begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \qquad \hat{S}_y^2 = \frac{1}{2}\begin{bmatrix} 1 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & 1 \end{bmatrix}, \qquad \hat{S}_z^2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{3.137}$$

With the above expressions, it is now straightforward to verify that $\hat{S}_x^2$, $\hat{S}_y^2$, and $\hat{S}_z^2$ are mutually commuting, and to see that the following result can be derived from the previous matrices:

$$\hat{\mathbf{S}}^2 = \hat{S}_x^2 + \hat{S}_y^2 + \hat{S}_z^2 = 2\hat{I}. \tag{3.138}$$

Since the eigenvalues of the operators $\hat{S}_x^2$, $\hat{S}_y^2$, $\hat{S}_z^2$ are either 0 or 1, it is necessary that the eigenvalue of one, and only one, of these operators is 0. Now, let us come back to the Kochen–Specker's model. Each of the triangles is a set of three orthogonal directions (like x, y, z), while the other triangles are generated by some kind of rotation of the coordinate axes. Since according to QM only one of these orthogonal

directions is zero, it follows that all of the statements $b_{jk} = (a_j a_k)'$ as well all of the statements of the form $c_{ijk} = a_i + a_j + a_k$ can hold true in QM. This in turn means that any compound statement d that is a conjunction of all of these statements b and c holds true in QM despite the fact that classically it is a logical falsity, as we have previously proved. As a consequence, also the statement e turns out to be true according to QM.

Failure of Classical Realism

Thus, the Kochen–Specker theorem provides a direct and powerful argument against the possibility of integrating QM with HVs capable to reproduce classical property ascription. The fact that in QM there can be value assignments that are not classically meaningful is a consequence of non-commutativity that makes many different experimental arrangements not compatible. A general lesson (that has already been anticipated) is that a realistic tenet like that of EPR is justified when dealing with the local experimental contexts (premeasurements) through which we select a certain observable to be measured. However, as a result of this examination, we cannot identify the eigenvalues of any observable as elements of reality if there is not some actual (detection) event that allows us to make a meaningful property ascription. As mentioned, this also shows that Bohr's position is justified only as a denial of the possibility of any value assignments in absence of detection (Sect. 3.3.2).[178] On this point there could be even some agreement with EPR since their argument is focused on predictions that are formulated *after* detections, although their argument deals with indirect measurement.

Thus, the problem arises only when we (implicitly) assume that the properties that we predict and assign in this way could be *identified* with elements of reality. In such a case, detection events would appear only as an a posteriori verification of the properties and the value assignment itself becomes only an a priori prediction of them. I have called this stronger version of realism *classical* or *local* realism. This is the basis of the belief that nature is essentially constituted of objects or bodies having classical properties. However, since the EPR's ontology consists essentially in dynamical and local interactions among events and objects that are taken to be real together with the physical observables that define them, we could better call this view *dynamical realism* as a further specification of Einstein's realistic tenet. It can be considered as a character of the original realism of EPR.

The identification of properties with elements of reality without the need to specify physical conditions is a general classical belief, but, as the following examination will show, it also represents an epistemological mistake. The British mathematician, logician and philosopher Alfred N. Whitehead (1861–1947) called this the "Fallacy of misplaced concreteness", namely the "error of mistaking the abstract for the concrete".[179] Moreover, this examination shows that we cannot consider quantum

[178] See also Bub (1989).

[179] Whitehead (1925).

systems as objects in the classical sense of the word either. This means that the question ("in which sense does the quantum-mechanical formalism describe kinds of objects?") raised at the beginning of Sect. 2.1.1, was still influenced by a classical view of the physical world. In the following we shall look for a new formulation of that question.

3.3.5 Bell Inequalities for Testing EPR

Bell's Inequality

Having shown the impossibility of the classical-realistic assumption, we focus now on the separability assumption. Apart from his reformulation of von Neumann's theorem, Bell's historical merit was to discover a precise mathematical formulation for expressing the conflict between the predictions of theories based on HVs and those of QM.[180] This opened new paths for dealing with the separability problem. Following EPR (Sect. 3.3.1), Bell based his theorem on two hypotheses: one about reality and one about separability. The former states that the properties that belong to physical states are represented by HVs, which, as we know, would mean that properties are independent of the experimental context for their verification (they are elements of reality in the sense of local realism). For this reason, these forms of HV theories are not contextual.[181] As previously recalled, non-contextual solutions have been also proposed: if we admit (or we are forced to admit) that the 'responses' given by a quantum system are not determined by its past history (since they depend on the current context only), we cannot assume that HVs determine these results in every context.[182] Nevertheless, I maintain that only non-contextual HV theories express the essence of the classical realism of the EPR argument against QM, as far as it presupposes a kind of objectivity that cannot depend on the operations at hand. The second assumption holds that events do not bear instant influence at a distance, and in this case they are separable. Since the condition of separability bans any interdependence between particles in absence of any dynamic interaction (Sects. 2.3.1–2.3.2), Bell introduced *local* hidden parameters on which the spin detection of a particle in a certain direction exclusively depended, so that the probability distributions of the measurement outcomes for the two particles would turn out to be independent of each other. Mathematically, this is formulated by means of the so-called Bell's inequality: a deterministic theory of non-contextual HVs, based on the condition of separability, must satisfy this inequality, while the quantum theory would violate it.[183]

---

[180] Bell (1964).

[181] Auletta (2000, pp. 549–50).

[182] Conway and Kochen (2006).

[183] I have, also in collaboration with other scholars, dealt several times with these problems (Auletta 2000, pp. 589–91; Auletta et al. 2009, Sect. 16.4.2; Auletta and Wang 2014, Sect. 10.4).

Let us now consider some details. The *Gedankenexperiment* proposed by Bell was a further refinement of the EPR-Bohm proposed experiment (Sect. 2.3.1). As said, Bell assumed the existence of a hidden parameter λ_{HV} such that it provides a full deterministic description of the measurement results obtained in the EPR-Bohm experiment. In particular, let us now imagine to perform a joint measurement in which the Cartesian 3D spin $\hat{\boldsymbol{\sigma}}_1$ of particle 1 along a chosen direction **a**, i.e. the observable mathematically expressed by the scalar product

$$\hat{\boldsymbol{\sigma}}_1 \cdot \mathbf{a} = a_x \hat{\sigma}_{1x} + a_y \hat{\sigma}_{1y} + a_z \hat{\sigma}_{1z} \tag{3.139a}$$

is measured with apparatus A and the Cartesian 3D spin $\hat{\boldsymbol{\sigma}}_2$ of particle 2 along another chosen direction **b**, i.e. the observable

$$\hat{\boldsymbol{\sigma}}_2 \cdot \mathbf{b} = b_x \hat{\sigma}_{2x} + b_y \hat{\sigma}_{2y} + b_z \hat{\sigma}_{2z} \tag{3.139b}$$

is measured with apparatus B, where the Euclidean vectors **a** and **b** are taken to be unit vectors as they represent here spatial directions and the Pauli matrices for both systems are given by Eq. (1.324). In other words, we are working in units of $\hbar/2$.

Up to now we have expressed the requirement of classical realism. Given the hidden parameter λ_{HV}, the results of this joint measurement can be uniquely described by the deterministic function $M_{\mathbf{a},\mathbf{b}}(\lambda_{\mathrm{HV}})$. Since the joint measurement is carried out after the time the two particles no longer interact, we satisfy the separability condition and can express such a condition in a mathematical rigorous way as

$$M_{\mathbf{a},\mathbf{b}}(\lambda_{\mathrm{HV}}) = A_{\mathbf{a}}(\lambda_{\mathrm{HV}}) \cdot B_{\mathbf{b}}(\lambda_{\mathrm{HV}}) \,, \tag{3.140}$$

which can be interpreted as a factorisation rule: the function $A_{\mathbf{a}}(\lambda_{\mathrm{HV}})$ describing the results obtained by measuring the spin of particle 1 in the direction **a** with the apparatus A, regardless of the measurement results of particle 2, and the function $B_{\mathbf{b}}(\lambda_{\mathrm{HV}})$ describing the results of measuring the spin of particle 2 in the direction **b** with an apparatus B, regardless of the measurement results of particle 1. I consider here coupling apparatus and detector as a single device for simplicity. Moreover, in order to satisfy classical or local realism too, also the functions $A_{\mathbf{a}}(\lambda_{\mathrm{HV}})$ and $B_{\mathbf{b}}(\lambda_{\mathrm{HV}})$ themselves, like $M_{\mathbf{a},\mathbf{b}}(\lambda_{\mathrm{HV}})$, need to be deterministic functions of the HVs. In sum, the separability condition above expresses the fact that the measurement results for the two 'separated' particles are mutually independent but only depend on local HVs.

Let us stress the novelties introduced by Bell relative to the discussion between EPR, Bohr, and Schrödinger. First, the issue of completeness does no longer play a central role. Second, we deal here with measurements, and this could be seen as a follow-up of Bohr's contribution, as far as he criticised EPR's reality criterion by pointing out that only detection can allow reality ascription to properties (Sect. 3.3.2). Third, differently from Schrödinger (and von Neumann) (Sects. 3.3.3–3.3.4), he only employed 'local' observables.

Since $A_{\mathbf{a}}(\lambda_{\mathrm{HV}})$ and $B_{\mathbf{b}}(\lambda_{\mathrm{HV}})$ are functions describing the possible measurement results of spin components in certain directions, which can be either $+1$ (representing

spin-up) or -1 (representing spin-down), then we have

$$A_{\mathbf{a}}(\lambda_{\mathrm{HV}}) = \pm 1, \quad B_{\mathbf{b}}(\lambda_{\mathrm{HV}}) = \pm 1. \tag{3.141}$$

This means that, for any orientation $\mathbf{a}$, we have

$$|A_{\mathbf{a}}(\lambda_{\mathrm{HV}})|^2 = 1. \tag{3.142}$$

If $\wp(\lambda_{\mathrm{HV}})$ denotes the probability distribution of λ_{HV}, then the expectation value of the product of the two components $\hat{\boldsymbol{\sigma}}_1 \cdot \mathbf{a}$ and $\hat{\boldsymbol{\sigma}}_2 \cdot \mathbf{b}$ is (see Eq. (1.102) or (1.126))

$$\langle \mathbf{a}, \mathbf{b} \rangle = \int_{\Lambda_{\mathrm{HV}}} d\lambda_{\mathrm{HV}} \wp(\lambda_{\mathrm{HV}}) A_{\mathbf{a}}(\lambda_{\mathrm{HV}}) B_{\mathbf{b}}(\lambda_{\mathrm{HV}}) \, , \tag{3.143}$$

where Λ_{HV} represents the set of all possible values of λ_{HV} and I have used the simplified notation

$$\langle \mathbf{a}, \mathbf{b} \rangle := \left\langle \left(\hat{\boldsymbol{\sigma}}_1 \cdot \mathbf{a} \right) \left(\hat{\boldsymbol{\sigma}}_2 \cdot \mathbf{b} \right) \right\rangle \tag{3.144}$$

for the joint mean value. Note that when the directions $\mathbf{a}$ and $\mathbf{b}$ are parallel, being the system in the singlet state (2.54) by hypothesis (I recall that EPR used the basic equations of QM), quantum physics requires the anticorrelation between the results of the two measurements:

$$\langle \mathbf{a}, \mathbf{a} \rangle_{\Psi_0} = -1. \tag{3.145}$$

However, this condition holds iff we also have

$$A_{\mathbf{a}}(\lambda_{\mathrm{HV}}) = -B_{\mathbf{a}}(\lambda_{\mathrm{HV}}). \tag{3.146}$$

In this case, $\langle \mathbf{a}, \mathbf{b} \rangle$ reach its minimum value. This allows us to rewrite Eq. (3.143) using the apparatus A only

$$\langle \mathbf{a}, \mathbf{b} \rangle = -\int_{\Lambda_{\mathrm{HV}}} d\lambda_{\mathrm{HV}} \wp(\lambda_{\mathrm{HV}}) A_{\mathbf{a}}(\lambda_{\mathrm{HV}}) A_{\mathbf{b}}(\lambda_{\mathrm{HV}}) \, , \tag{3.147}$$

Then, by assuming a third measurement direction $\mathbf{c}$ alternative to $\mathbf{b}$ (for particle 2) and thanks to Eq. (3.142), we can write

$$\begin{aligned} \langle \mathbf{a}, \mathbf{b} \rangle - \langle \mathbf{a}, \mathbf{c} \rangle &= -\int d\lambda \wp(\lambda) [A_{\mathbf{a}}(\lambda) A_{\mathbf{b}}(\lambda) - A_{\mathbf{a}}(\lambda) A_{\mathbf{c}}(\lambda)] \\ &= \int d\lambda \wp(\lambda) A_{\mathbf{a}}(\lambda) A_{\mathbf{b}}(\lambda) [A_{\mathbf{b}}(\lambda) A_{\mathbf{c}}(\lambda) - 1] \, , \end{aligned} \tag{3.148}$$

where I have dropped the subscript HV for simplicity. Since for any integrable function $f(x)$, we have the evident mathematical inequality

$$\left| \int dx f x \right| \leq \int dx |f(x)|, \tag{3.149}$$

then, thanks to the fact that, given the value assignments (3.141), we have

$$|A_{\mathbf{b}}(\lambda) A_{\mathbf{c}}(\lambda) - 1| = 1 - A_{\mathbf{b}}(\lambda) A_{\mathbf{c}}(\lambda), \tag{3.150}$$

we can deduce

$$| \langle \mathbf{a}, \mathbf{b} \rangle - \langle \mathbf{a}, \mathbf{c} \rangle | \leq \int d\lambda \wp(\lambda) [1 - A_{\mathbf{b}}(\lambda) A_{\mathbf{c}}(\lambda)] \, . \tag{3.151}$$

On this basis, we have derived the (first) *Bell inequality*:

$$| \langle \mathbf{a}, \mathbf{b} \rangle - \langle \mathbf{a}, \mathbf{c} \rangle | \ \leq 1 + \langle \mathbf{b}, \mathbf{c} \rangle \ , \tag{3.152}$$

with all mean values having the general form of Eq. (3.147). The inequality tells us that these three mean values (the sum of two of them in absolute value) cannot exceed 1 if there is dependence on a local hidden parameter.

CHSH Inequality and Bell Operator

A more used expression is due to the physicists J. Clauser, M. Horne, A. Shimony, and M. Holt (known as CHSH, as already mentioned in Sect. 2.3.2)[184]:

$$| \langle \mathbf{a}, \mathbf{b} \rangle - \langle \mathbf{a}, \mathbf{b}' \rangle + \langle \mathbf{a}', \mathbf{b} \rangle + \langle \mathbf{a}', \mathbf{b}' \rangle | \leq 2 \, , \tag{3.153}$$

where $\mathbf{a}$ and $\mathbf{a}'$ as well as $\mathbf{b}$ and $\mathbf{b}'$ are alternative settings for particles 1 and 2, respectively. It is evident that settings $\mathbf{a}$ and $\mathbf{b}'$ are antiparallel. Let us expand here the expectation values occurring in inequality (3.197) as

$$\langle \mathbf{a}, \mathbf{b} \rangle = \wp(1, 1|\mathbf{a}, \mathbf{b}) + \wp(-1, -1|\mathbf{a}, \mathbf{b}) - \wp(1, -1|\mathbf{a}, \mathbf{b}) - \wp(-1, 1|\mathbf{a}, \mathbf{b}), \tag{3.154}$$

according to statistics (photon count), and similarly for the other terms, where the probability of the experimental outcomes $(+1, -1)$ is conditional on the apparatus orientation ($\mathbf{a}$, $\mathbf{a}'$ and $\mathbf{b}$, $\mathbf{b}'$). I recall that the conditional probability $\wp(B|A)$ of the occurring of an event B when the event A is already occurred is given by

$$\wp(B|A) = \frac{\wp(A, B)}{\wp(A)}. \tag{3.155}$$

[184]Clauser et al. (1969). The authors used for the first time a stochastic HV theory and not a deterministic one as in the case of Bell's previous paper. The formulation that I have reported here is slightly different relative to that of their original paper and is due to a later paper of Bell (1971).

Note that the expression in the modulus is the expectation value of the so-called Bell operator[185] given by

$$\begin{aligned}\hat{\mathrm{B}} &= (\hat{\boldsymbol{\sigma}}_1 \cdot \mathbf{a})(\hat{\boldsymbol{\sigma}}_2 \cdot \mathbf{b} - \hat{\boldsymbol{\sigma}}_2 \cdot \mathbf{b}') + (\hat{\boldsymbol{\sigma}}_1 \cdot \mathbf{a}')(\hat{\boldsymbol{\sigma}}_2 \cdot \mathbf{b} + \hat{\boldsymbol{\sigma}}_2 \cdot \mathbf{b}') \\ &= (\hat{\boldsymbol{\sigma}}_1 \cdot \mathbf{a})(\hat{\boldsymbol{\sigma}}_2 \cdot \mathbf{b}) - (\hat{\boldsymbol{\sigma}}_1 \cdot \mathbf{a})(\hat{\boldsymbol{\sigma}}_2 \cdot \mathbf{b}') + (\hat{\boldsymbol{\sigma}}_1 \cdot \mathbf{a}')(\hat{\boldsymbol{\sigma}}_2 \cdot \mathbf{b}) + (\hat{\boldsymbol{\sigma}}_1 \cdot \mathbf{a}')(\hat{\boldsymbol{\sigma}}_2 \cdot \mathbf{b}'),\end{aligned} \tag{3.156}$$

so that its expectation value is

$$\left\langle \hat{\mathrm{B}} \right\rangle = \langle \mathbf{a}, \mathbf{b} \rangle - \left\langle \mathbf{a}, \mathbf{b}' \right\rangle + \left\langle \mathbf{a}', \mathbf{b} \right\rangle + \left\langle \mathbf{a}', \mathbf{b}' \right\rangle. \tag{3.157}$$

In other words, when testing the CHSH inequality, we are in fact measuring the Bell operator. Note that it is mostly the CHSH inequality that has ben experimentally tested in the following years. Although relatively complicated, the Bell operator is not in principle different from any quantum observable, and thus, the expansion (3.156) is analogical with respect to the basic expression (1.96) for spectral decomposition of any quantum observable (here for a composite system). This result is outmost important as far as we have found a non 'fictional' observable for two entangled subsystems.

Experiments

Let us first consider an ideal experiment and have a look at Fig. 3.12. Let us take $\mathbf{a}$, $\mathbf{a}'$, $\mathbf{b}$ and $\mathbf{b}'$ to be coplanar and the angle between $\mathbf{a}'$ and $\mathbf{b}$, between $\mathbf{b}$ and $\mathbf{a}$, and between $\mathbf{a}$ and $\mathbf{b}'$ to be $\theta = \pi/4$. Moreover, I recall that all vector norms (magnitudes) are $\| \mathbf{a} \| = \| \mathbf{b} \| = \| \mathbf{a}' \| = \| \mathbf{b}' \| = 1$. First, note that, for the singlet state (2.54), we have the expectation

$$\langle \mathbf{a}, \mathbf{b} \rangle = -\mathbf{a} \cdot \mathbf{b}, \tag{3.158}$$

for all directions $\mathbf{a}$, $\mathbf{b}$. In fact, using Eqs. (3.139), we have

$$\left\langle \Psi_0 \left| a_x b_x \hat{\sigma}_{1x} \hat{\sigma}_{2x} \right| \Psi_0 \right\rangle = -a_x b_x, \tag{3.159a}$$

$$\left\langle \Psi_0 \left| a_y b_y \hat{\sigma}_{1y} \hat{\sigma}_{2y} \right| \Psi_0 \right\rangle = -a_y b_y, \tag{3.159b}$$

$$\left\langle \Psi_0 \left| a_z b_z \hat{\sigma}_{1z} \hat{\sigma}_{2z} \right| \Psi_0 \right\rangle = -a_z b_z, \tag{3.159c}$$

while the remaining six cross-terms (i.e. terms resulting from the product between different components) are all zero. Since the scalar product is given by formula (1.60), we have $\langle \mathbf{a}, \mathbf{b} \rangle = -\cos\theta$. In these conditions, all expectation values are $-1/\sqrt{2}$ apart from $\left\langle \mathbf{a}', \mathbf{b}' \right\rangle$ that is equal to $1/\sqrt{2}$. Thus, we have for the quantum expectation value of the Bell operator.

[185] The reference paper is Braunstein et al. (1992).

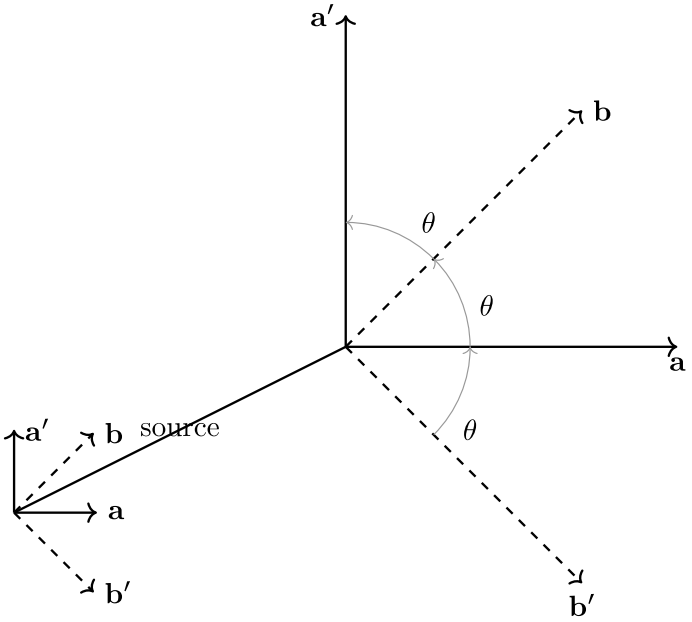


Fig. 3.12 Optimal orientations $\mathbf{a}$, $\mathbf{a}'$, $\mathbf{b}$, and $\mathbf{b}'$ for testing inequality (3.153). Adapted from Auletta et al. (2009, p. 592)

$$\left|\langle \mathbf{a}, \mathbf{b}\rangle - \langle \mathbf{a}, \mathbf{b}'\rangle + \langle \mathbf{a}', \mathbf{b}\rangle + \langle \mathbf{a}', \mathbf{b}'\rangle\right|_{\mathrm{QM}} = \left|-\frac{1}{\sqrt{2}} + \left(-\frac{1}{\sqrt{2}}\right) - \frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}}\right| = 2\sqrt{2}\,, \tag{3.160}$$

which is clearly ≥ 2.

Experiments for testing the CHSH inequality already started in the mid of 1970s and proved that the inequality is violated showing that QM's predictions are correct.[186] The early performed experiments made use of radioactive decay (Fig. 3.13a) but were exposed to two main objections. First, it was well possible that some hidden and deterministic (causal) but non-local exchanges of signal between the polarisers and the source could explain the violation of the CHSH inequality. To face this problem, the French physicist Alain Aspect performed in 1982 another kind of experiment (Fig. 3.13b).[187] He interposed in the photons' trajectory extremely fast and random change-of-direction devices, so that, once conveniently separated the source of photons from the measuring instrument, the time needed for a photon to reach the measuring instrument was twice as long as the time needed for the change-of-

[186] I invite the reader who likes to know more about these experiments to consult (Auletta 2000, Chap. 35) and literature quoted there. With the language of category theory, we have now an equaliser of quantum theory and experiment (Spivak 2013, Sect. 2.5.3).

[187] Aspect et al. (1982).

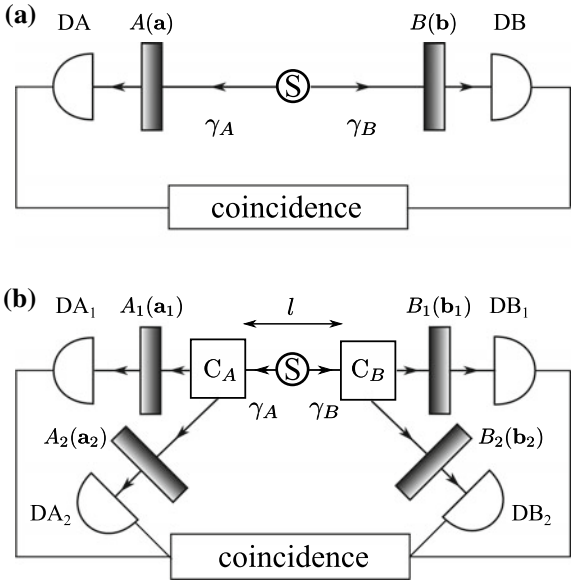


Fig. 3.13 First tests of CHSH inequality. Photons became the principal quantum object system. The correlated photons γ_A, γ_B coming from the source S impinge upon the linear polarisers I, II oriented in directions **a**, **b**, respectively. The rate of joint detection by the photomultipliers is monitored for various combinations of orientations (**a**). Experiment proposed by Aspect: The optical commutator C_A directs the photon γ_A either towards polariser I_1 with orientation $\mathbf{a}_1$ or to polariser I_2 with orientation $\mathbf{a}_2$. Similarly for C_B. The two commutators work independently (the time intervals between two commutations are taken to be stochastic). The four joint detection rates are monitored and the orientations $\mathbf{a}_1, \mathbf{a}_2, \mathbf{b}_1, \mathbf{b}_2$ are not changed for the whole experiment. l is the separation between the switches (**b**). Adapted from Auletta et al. (2009, p. 600)

direction device to switch directions. No difference was found for different direction changes. Thus, any possible influence of the measuring instrument on the source was avoided: it turned out that Bell's inequality was still violated, but now it could not be attributed to the influence of the measuring instrument on the source. Thus, the detection of the polarisation on one of the photons ensured that the other one be in the opposite polarisation direction, at the same time and without interaction.

Another objection was the possibility that there could be an angular correlation between the orientations of the particles during the radioactive decay.[188] To face this problem, spontaneous parametric down-conversion (SPDC) experiments were performed instead of radioactive decay ones.[189] In such experiments, the photons are pumped by a laser and go through a nonlinear crystal (Fig. 3.14). In particular, SPDC consists in pumping a nonlinear crystal with ultraviolet (high-energy) light. An ultraviolet photon may eventually decay into a couple (of less energetic) photons,

[188] Santos (1991).

[189] Among the first reported experiments I quote (Shih and Alley 1988; Ou and Mandel 1988).

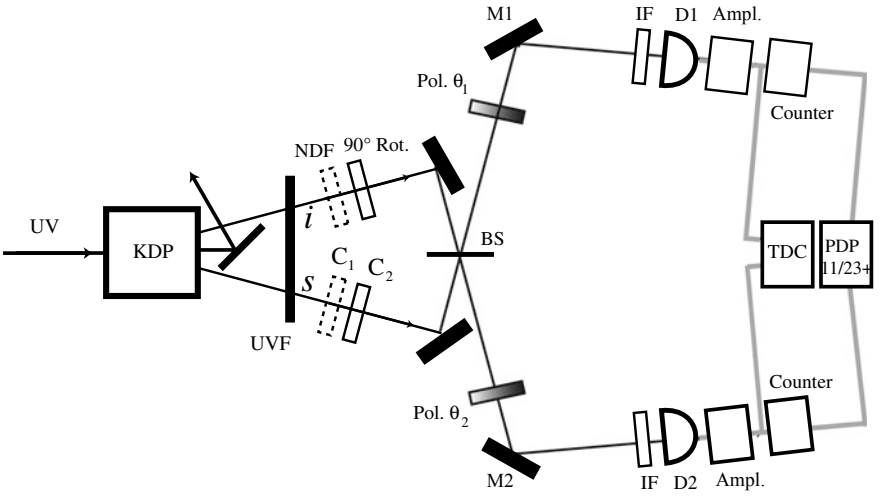


Fig. 3.14 Outline of the Alley–Shih and Ou–Mandel's experiment (have also a look at Fig. 3.8, Sect. 3.3.5). Light from the 351.1-nm line of an argon-ion laser falls on a nonlinear crystal of potassium dihydrogen phosphate (KDP), where down-converted photons of wavelength of about 702 nm are produced. When the condition for degenerate phase matching is satisfied, down-converted, linearly polarised signal and idler photons emerge at angles of about $\pm 2°$ relative to the ultraviolet (UV) pump beam with the electric vector in the plane of the diagram. The *idler* (i) photons pass through a 90° polarisation rotator, while the *signal* (s) photons traverse a compensating glass plate C_1 producing an equal time delay. s-photons and i-photons are then directed from opposite sides towards a beam splitter (BS). The input to the BS consists of an x-polarised s-photon and of a rotated y-polarised i-photon. The light beams emerging from BS, consisting of a mixing of i-photons and s-photons, pass through linear polarisers set at adjustable angles θ_1 and θ_2, through similar interference filters (IF) and finally fall on two photodetectors D_1 and D_2. The photoelectric pulses from D_1 and D_2 are amplified and shaped and fed to the start and stop inputs of a time-to-digital converter (TDC) under computer control which functions as a coincidence counter. Adapted from Auletta et al. (2009, p. 602)

obeying the laws of conservation of energy (and momentum). In other words, we have

$$\nu_p = \nu_i + \nu_s \qquad \text{and} \qquad \hat{\mathbf{k}}_p = \hat{\mathbf{k}}_i + \hat{\mathbf{k}}_s, \tag{3.161}$$

where the indexes p, i, s stand for pump, idler and signal photons, respectively, and $\hat{\mathbf{k}}$ is the propagation vector (1.166). Moreover, we distinguish between type I of SPDC if the outgoing photons have the same polarisation direction and type II if they have orthogonal directions. Thus, the crucial point is that these devices produce photon pairs with good angular correlation and they do not need to be collinear, as in the cascade experiments.

Also other objections could be raised and in fact were raised.[190] However, these objections instead of invalidating the experimental results (and the predictions of QM) in fact pushed towards experimental solutions that were more and more refined and that confirmed always the quantum-mechanical predictions. Therefore, towards the end of last century, the scientific community became growingly convinced of the correctness of the latter also about the issue of non-locality. Nevertheless, it should be remarked that without those objections the theory and the experiments would not have advanced so far. This shows the relevance of discussions about interpretative issues even for the fine-tuning of experimental testing. In general, the whole discussion of the EPR argument up to the experimental tests is one of the best examples of the whole history of science, a true case study, for understanding how theories are tested and interpretations eventually corrected.

3.3.6 Entanglement Swapping, Teleportation, Quantum Cryptography

As a conclusion of the previous examination, we can take for granted that both the CHSH inequality is violated and experimental tests on quantum systems contradict separability. The question now is: is there any direct evidence for entanglement being a non-local correlation without exchange of signals? I recall that this was the essence of Schrödinger's objection to EPR, mentioned in Sect. 3.3.3 and is the object of the following examination.

The Entanglement Swapping Experiment

The question can be specifically reformulated as follows: is it possible for quanta that, in a controlled context, have never directly interacted before to get entangled? In fact, if we are able to entangle systems without direct interaction, we shall have found experimental evidence for entanglement being a correlation without exchange of signals. The milestone step towards the generalisation of the entanglement concept was the *entanglement swapping* experiment performed by Zeilinger and collaborators in 1993.[191] Consider two pairs of entangled photons emitted by two independent sources as shown in Fig. 3.15. The states of the two photon pairs are given by

[190] Marshall et al. (1983), Garuccio and Selleri (1984), Marshall and Santos (1985), Ferrero et al. (1990).

[191] Żukowski et al. (1993). See also Auletta et al. (2009, Sect. 16.6).

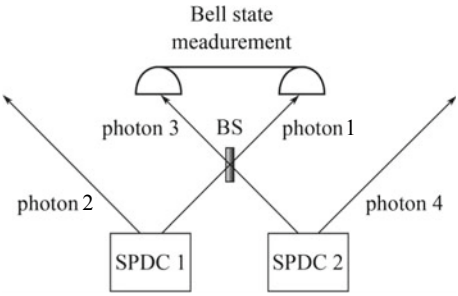


Fig. 3.15 Schematic set-up of entanglement swapping. Two pumped spontaneous parametric down-conversion sources SPDC 1 and SPDC 2 each emit a photon pair (1–2 and 3–4, respectively). Note that photons 2 and 3 merge through a beam splitter BS before being measured by a Bell state measurement device. If the orientation of the nonlinear optical crystal is chosen appropriately, two such possible decay processes become indistinguishable and lead to quantum correlations between the daughter photons generated by those crystals. Adapted from Auletta et al. (2009, p. 611)

$$|\Psi\rangle_{12} = \frac{1}{\sqrt{2}}\big(|0\rangle_1 |1\rangle_2 - |1\rangle_1 |0\rangle_2\big), \tag{3.162a}$$

$$|\Psi\rangle_{34} = \frac{1}{\sqrt{2}}\big(|0\rangle_3 |1\rangle_4 - |1\rangle_3 |0\rangle_4\big), \tag{3.162b}$$

where the subscripts 12 and 34 denote the photon pairs composed of photons 1 and 2, and photons 3 and 4, respectively. Note that these states are similar to the singlet state (2.54), but here they are reformulated in a rather abstract way, since $|0\rangle$ may stand for horizontal polarisation or for spin-up whereas $|1\rangle$ for vertical polarisation or spin-down, but also further possibilities are allowed. Evidently, the state of the composite four-photon system is factorised, i.e. we have

$$\begin{aligned} |\Psi\rangle_{1234} &= |\Psi\rangle_{12} \otimes |\Psi\rangle_{34} \\ &= \frac{1}{2}\left(|0101\rangle - |0110\rangle - |1001\rangle + |1010\rangle\right)_{1234}. \end{aligned} \tag{3.163}$$

In other words, while the two photons in each photon pair are entangled, there is *no* entanglement between either of the two photons with either of the two photons taken from different photon pairs.

If we now perform a Bell-like joint measurement on photons 1 and 3, we are able to project photons 2 and 4 onto an entangled state that depends on the result of the measurement on photons 2 and 3. To be specific, let us use the so-called *Bell states* for photons 1 and 3, which in the basis $\{|0\rangle, |1\rangle\}$ are given by[192]

[192] Braunstein et al. (1992).

$$\left|\Psi^{+}\right\rangle_{13} := \frac{1}{\sqrt{2}}(|0\rangle_1|1\rangle_3 + |1\rangle_1|0\rangle_3), \tag{3.164a}$$

$$\left|\Psi^{-}\right\rangle_{13} := \frac{1}{\sqrt{2}}(|0\rangle_1|1\rangle_3 - |1\rangle_1|0\rangle_3), \tag{3.164b}$$

$$\left|\Phi^{+}\right\rangle_{13} := \frac{1}{\sqrt{2}}(|0\rangle_1|0\rangle_3 + |1\rangle_1|1\rangle_3), \tag{3.164c}$$

$$\left|\Phi^{-}\right\rangle_{13} := \frac{1}{\sqrt{2}}(|0\rangle_1|0\rangle_3 - |1\rangle_1|1\rangle_3), \tag{3.164d}$$

where $\left|\Psi^{-}\right\rangle_{13}$ is again generalisation of the singlet state (2.54). With a suitable reformulation it may be shown that the Bell states are eigenstates of the Bell operator $\hat{\mathrm{B}}$ given by Eq. (3.156). The Bell states are four maximally entangled states of photons 1 and 3 and also constitute a complete orthonormal basis of the Hilbert space of photons 1 and 3. A measurement of the Bell operator is called *Bell measurement*.

In order to find the effect of the joint measurement that projects the compound state (3.163) onto one of the Bell states given by Eqs. (3.164), we express the states for both pairs of photons 2, 4 and 1, 3 in terms of their own Bell states rewriting the state $|\Psi\rangle_{1234}$ as

$$|\Psi\rangle = \frac{1}{2}\left(\left|\Psi^{+}\right\rangle_{24}\left|\Psi^{+}\right\rangle_{13} - \left|\Psi^{-}\right\rangle_{24}\left|\Psi^{-}\right\rangle_{13} - \left|\Phi^{+}\right\rangle_{24}\left|\Phi^{+}\right\rangle_{13} + \left|\Phi^{-}\right\rangle_{24}\left|\Phi^{-}\right\rangle_{13}\right), \tag{3.165}$$

where I have dropped the subscript 1234 for the sake of notation. Note also that in the above expansion the components for photons 2 and 4 are paired with components of the same type for photons 1 and 3. It follows from Eq. (3.165) that the projection of the state $|\Psi\rangle$ onto a Bell state of photons 1 and 3, also projects photons 2 and 4 onto an entangled state precisely of the same form (and vice versa). In other words, by performing a joint measurement on photons 1 and 3 that entangles these two initially unentangled photons, we are able to entangle also the two initially unentangled photons 2 and 4. Indeed, suppose that we get one of such components, say $\left|\Phi^{+}\right\rangle_{24}\left|\Phi^{+}\right\rangle_{13}$. Now, it is evident that we can write

$$\begin{aligned}\left|\Phi^{+}\right\rangle_{24}\left|\Phi^{+}\right\rangle_{13} &= \frac{1}{2}\left(|0000\rangle + |0011\rangle + |1100\rangle + |1111\rangle\right)_{1234} \\ &= \frac{1}{2}\left(|00\rangle_{12}|00\rangle_{34} + |00\rangle_{12}|11\rangle_{34} + |11\rangle_{12}|00\rangle_{34} + |11\rangle_{12}|11\rangle_{34}\right) \\ &= \frac{1}{2}\sum_{j,k=0,1,\,j\neq k}\left(|jj\rangle_{12}\otimes|jj\rangle_{34} + |jj\rangle_{12}\otimes|kk\rangle_{34}\right),\end{aligned} \tag{3.166}$$

what shows that, while the couples 13 and 24 are entangled, the couples 12 (as well as 34) are no longer entangled (at least locally and in the ideal controlled case) since they are in a product state (Fig. 3.16). Thus, entanglement swapping can be considered as an *indirect* entangling.

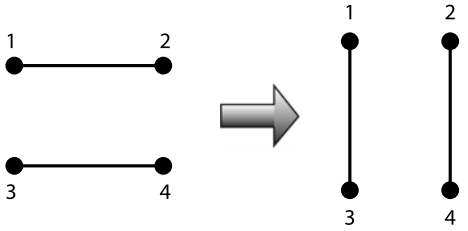


Fig. 3.16 Scheme of entanglement swapping. The edges represent entanglement

The Causal Influence of Entanglement

Note that photons 2 and 4 likely never locally interacted before and therefore are not involved at all in this dynamical interaction. Therefore, we are finally justified in saying that entanglement represents a form of interdependence in which no dynamic interaction is necessary. For this character of quantum correlations, the American physicist and philosopher Abner Shimony has aptly spoken of *passion at a distance* instead of action at a distance[193] (and this is the main difference with Bohm's active piloting: Sect. 3.3.4). This is what justifies the fact that all quantum systems of our universe are plus or minus entangled: at the global scale no interaction (although some local interaction is needed to manifest that influence), and therefore no propagation of signals is necessary but it suffices that some systems which are (even indirectly) correlated with other systems interact locally to produce detectable consequences on the latter ones, in agreement with the analysis of Sect. 3.2.2.[194] In other words, entanglement can show *causal influence without mechanical or efficient causation* and therefore without exchange of signals, what makes us appreciate its expression in terms of mutual information. However, also the reciprocal is true: entanglement can locally 'disappear' as the result of interactions between other couples of particles, in agreement with what decoherence tells us (Sect. 3.1.3). It is worth recalling that the first scholar to have considered this kind of possibility was W. Leibniz, who had first envisaged correlations without efficient causation but had also pointed out that monads are shielded against the exterior and therefore evolve not thanks to external signals but "out of themselves" (Sect. 2.3.3).

The Quantum Teleportation Protocol

Another striking aspect of QM that shows the connection between entanglement and information is *quantum teleportation*.[195] By this term, it is meant a procedure that is able to 'transfer' with certainty the state of an input quantum system (understood

[193] Shimony (1983).

[194] This is why to say that entanglement is an exclusive and discriminating connection among particles, as unfortunately is written in Maudlin (1994, p. 23), is a mistake.

[195] First proposed in Bennett and Wiesner (1992), Bennett et al. (1993). See also Auletta and Wang (2014, Sect. 11.5).

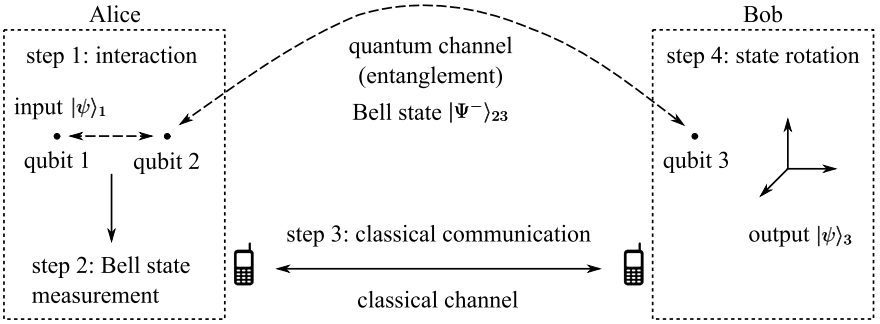


Fig. 3.17 Scheme of quantum teleportation. In step 1, Alice let qubit 1 (the qubit to be teleported) interact with qubit 2 which is entangled with qubit 3, previously given to Bob, in the Bell state $\left|\Psi^-\right\rangle_{23}$. In step 2 she reads the outcome of the Bell measurement performed on qubits 1 and 2. In step 3, she classically communicates this result to Bob. In function of this classical piece of information, in step 4 Bob rotates the state of qubit 3 and obtains as the output state $|\psi\rangle_3$ which has the same form of the state $|\psi\rangle_1$ of qubit 1. Adapted from Auletta and Wang (2014, p. 364)

as codified information) onto the state of a distant output system of the same type. As we know, an instantaneous transfer of information is not possible: due to the intrinsic randomness of QM we cannot manipulate at a distance the outcomes of a system, neither, as we shall see, their probability distributions (or experimental context), although that system may be entangled with another one that we are able to manipulate locally (obviously, we can do this by choosing certain measurement procedures and not by controlling the possible outcomes). However, it is possible to 'teleport' some information by exploiting entanglement and without violating the Einstein's locality (that is, without overcoming the limitations on the speed of the spread of signals set by the relativity theory: see Sect. 2.3.2).

Suppose that a sender, conventionally called Alice, wishes to send to a receiver, conventionally called Bob, the information about a quantum system, say a qubit labelled 1, prepared in a particular state $|\psi\rangle_1$ but *unknown* to her as well as to Bob (Alice may represent some dealing agency). To this purpose, Alice allows qubit 1, called the *ancilla*, to interact with qubit 2 that is entangled with qubit 3, previously given to Bob (Fig. 3.17). Now, Alice performs a Bell measurement on her composite two-qubit system 1–2, so that, by classically communicating to Bob the result of her measurement, Bob can obtain the initial state of qubit 1 onto qubit 3. In order to discuss this procedure specifically, let the entangled qubits 2 and 3 be in the EPR state

$$\left|\Psi^-\right\rangle_{23} = \frac{1}{\sqrt{2}}(|0\rangle_2|1\rangle_3 - |1\rangle_2|0\rangle_3), \tag{3.167}$$

which is again a generalisation of the singlet state (2.54). If qubit 1 is in the state

$$|\psi\rangle_1 = c_0|0\rangle_1 + c_1|1\rangle_1, \tag{3.168}$$

where c_0 and c_1 are complex probability amplitudes unknown to Alice (and Bob), then before the measurement the state of the whole composite three-qubit system 123 is given by the product state

$$\begin{aligned} |\Psi\rangle_{123} &= |\psi\rangle_1 \left|\Psi^-\right\rangle_{23} \\ &= \frac{c_0}{\sqrt{2}}(|0\rangle_1 |01\rangle_{23} - |0\rangle_1 |10\rangle_{23}) + \frac{c_1}{\sqrt{2}}(|1\rangle_1 |01\rangle_{23} - |1\rangle_1 |10\rangle_{23}). \end{aligned} \tag{3.169}$$

Alice could choose to measure qubits 1 and 2 in the Bell basis (3.164) for qubits 1–2

$$\left|\Psi^+\right\rangle_{12} = \frac{1}{\sqrt{2}}(|0\rangle_1 |1\rangle_2 + |1\rangle_1 |0\rangle_2), \tag{3.170a}$$

$$\left|\Psi^-\right\rangle_{12} = \frac{1}{\sqrt{2}}(|0\rangle_1 |1\rangle_2 - |1\rangle_1 |0\rangle_2), \tag{3.170b}$$

$$\left|\Phi^+\right\rangle_{12} = \frac{1}{\sqrt{2}}(|0\rangle_1 |0\rangle_2 + |1\rangle_1 |1\rangle_2), \tag{3.170c}$$

$$\left|\Phi^-\right\rangle_{12} = \frac{1}{\sqrt{2}}(|0\rangle_1 |0\rangle_2 - |1\rangle_1 |1\rangle_2). \tag{3.170d}$$

Given the local operation performed by Alice on systems 1–2, the state $|\Psi\rangle_{123}$ in terms of the above Bell states is

$$\begin{aligned} |\Psi\rangle_{123} = -\frac{1}{2}\big[&\left|\Psi^-\right\rangle_{12}(c_0|0\rangle_3 + c_1|1\rangle_3) + \left|\Psi^+\right\rangle_{12}(c_0|0\rangle_3 - c_1|1\rangle_3) \\ &- \left|\Phi^-\right\rangle_{12}(c_0|1\rangle_3 + c_1|0\rangle_3) - \left|\Phi^+\right\rangle_{12}(c_0|1\rangle_3 - c_1|0\rangle_3)\big]. \end{aligned} \tag{3.171}$$

This expression can be further simplified to

$$|\Psi\rangle_{123} = -\frac{1}{2}\big[\left|\Psi^-\right\rangle_{12}\hat{I}|\psi\rangle_3 + \left|\Psi^+\right\rangle_{12}\mathrm{Z}|\psi\rangle_3 - \left|\Phi^-\right\rangle_{12}\mathrm{X}|\psi\rangle_3 + \left|\Phi^+\right\rangle_{12} i\mathrm{Y}|\psi\rangle_3\big], \tag{3.172}$$

where X, Y, and Z are the usual Pauli matrices (1.324): since I have used the informational basis $|0\rangle, |1\rangle$, which, as remarked, can stand for different physical observables, the Pauli matrices, being logical gates acting on qubits, are conventionally denoted X, Y, Z. This allows us to make use of Table 3.5, in analogy with Table 1.2, Sect. 1.3.2. Moreover, note in such a context that the states (1.328) can be rewritten in informational terms as $|+\rangle, |-\rangle$, given by Eqs. (3.82), for expressing the eigenstates of X, as displayed in Fig. 3.6, Sect. 3.2.4, as well as

$$|y_+\rangle := \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle), \qquad |y_-\rangle := \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle) \tag{3.173}$$

Table 3.5 The actions of the Pauli matrices on the computational basis

$\mathrm{X}\,\vert 0\rangle = \vert 1\rangle$	$\mathrm{X}\,\vert 1\rangle = \vert 0\rangle$
$\mathrm{Y}\,\vert 0\rangle = i\,\vert 1\rangle$	$\mathrm{Y}\,\vert 1\rangle = -i\,\vert 0\rangle$
$\mathrm{Z}\,\vert 0\rangle = \vert 0\rangle$	$\mathrm{Z}\,\vert 1\rangle = -\vert 1\rangle$

for expressing the eigenstates of Y. Note that in fact, for spin directions, the latter two states are on the y-axis (as the eigenstates of the x component are on the x-axis). This clearly shows that a change of phase and not only a change of amplitudes can induce a change of (the Cartesian component of an) observable. It is only when we are interested in the probabilities of measurement outcomes that the phase becomes irrelevant, as it is evident in Eqs. (2.33). This is why, in general, time unitary transformations need to take into account both possible changes.

After Alice's measurement, the three-qubit system is projected into one of the four states superposed in $\vert\Psi\rangle_{123}$, depending on the measurement outcome. According to Eq. (3.172), this means that given Alice's measurement outcome, the state of Bob's qubit is given by

$$\left\vert\Psi^{-}\right\rangle_{12} \Rightarrow \hat{I}\left(\hat{I}\,\vert\psi\rangle_3\right), \tag{3.174a}$$

$$\left\vert\Psi^{+}\right\rangle_{12} \Rightarrow \mathrm{Z}\left(\mathrm{Z}\,\vert\psi\rangle_3\right), \tag{3.174b}$$

$$\left\vert\Phi^{-}\right\rangle_{12} \Rightarrow \mathrm{X}\left(\mathrm{X}\,\vert\psi\rangle_3\right), \tag{3.174c}$$

$$\left\vert\Phi^{+}\right\rangle_{12} \Rightarrow -i\mathrm{Y}\left(i\mathrm{Y}\,\vert\psi\rangle_3\right). \tag{3.174d}$$

Note that three of the four possible states of qubit 3 (Bob's one) are simply related to the state $\vert\psi\rangle_1$ of qubit 1 (the one Alice wished to teleport) by a unitary transformation, namely, identity and the Pauli unitary matrices X, Y, or Z (this is schematically resumed in Fig. 3.18). Note that, being the Pauli matrices the inverses of themselves, an application of an operation by Bob on a state already determined by the same operation, according to Eqs. (1.325a) and (3.172), ensures that the final state of the system 3 is identical to $\vert\psi\rangle$ of system 1. Note, finally, that the operator $i\mathrm{Y}$ is still unitary but not self-adjoint since $(i\mathrm{Y})^{\dagger} = -i\mathrm{Y}$, what implies

$$(i\mathrm{Y})\hat{I}(-i\mathrm{Y}) = (-i\mathrm{Y})\hat{I}(i\mathrm{Y}) = \mathrm{Y}^2 = \hat{I}. \tag{3.175}$$

This makes the choice of the sign in the product of this operator with $\hat{I}$ quite arbitrary considering the inverse operation performed by Bob.

Resuming, in the case of the first measurement outcome $\left\vert\Psi^{-}\right\rangle_{12}$, the state of qubit 3 is the same as that of qubit 1 except for an irrelevant phase factor, so that Bob needs no further operation to recover the state of qubit 1. In the three other cases, in order to convert the state of qubit 3 to the state of qubit 1, Bob must apply to qubit 3 one of the unitary operators X, Y, and Z, which, according to Eq. (1.338), represents rotations around the x-, y- or z-axis by π radians, respectively. The possible operations involved in the teleportation procedure are summarised in Table 3.6. What Bob has to do, obviously depends on the classical communication

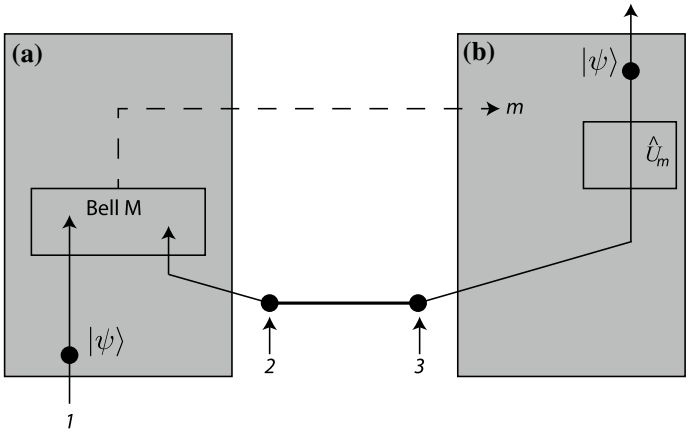


Fig. 3.18 How teleportation works. The local operations performed by Alice and Bob are framed in the left and right grey rectangles, respectively. The unitary transformation $\hat{U}_m$ performed by Bob depends on the result m of Alice's measurement classically communicated to Bob. Adapted from Abramsky and Coecke (2009)

Table 3.6 Possible events in a teleportation process. The fact that each of Alice's measurement result is uniquely mapped (through the entangled state $|\Psi^-\rangle_{23}$) to the input state $|\psi\rangle_1$ of qubit 1 (which could represent a code) allows that Alice classical instructs Bob about the kind of operation to be performed so as to recover the state $|\varphi\rangle_3$ of qubit 3

Alice's measurement	State of Bob's qubit	Bob's operation
$\|\Psi^-\rangle_{12}$	$\hat{I}\,\|\psi\rangle_3$	$\hat{I}$
$\|\Psi^+\rangle_{12}$	$\mathrm{Z}\,\|\psi\rangle_3$	Z
$\|\Phi^-\rangle_{12}$	$\mathrm{X}\,\|\psi\rangle_3$	X
$\|\Phi^+\rangle_{12}$	$i\mathrm{Y}\,\|\psi\rangle_3$	$-i\mathrm{Y}$

of Alice's measurement result, say via phone call or email. In other words, once that Alice obtains a certain measurement result on qubits 1–2, she sends to Bob a simple instruction. For instance, the instruction may reads "Do nothing," "Perform the rotation about the z axis," "Perform the rotation about the x-axis," or "Perform the rotation about the y-axis." For this reason, both Alice and Bob could be even computers with no understanding whatsoever of QM and therefore no knowledge of the initial conditions and the equations describing the state of the three qubits. It suffices that two bits of classical information is sent by Alice (according to a previously established code, like the binary numbers 00, 01, 10, and 11) so that Bob may perform the required operation as per prearranged rules. Thanks to the shared information, i.e. the entanglement between qubits 2 and 3, Bob is able to recover the encoded quantum information. In fact, as we shall see later, this can be made with both classical and quantum means. Finally, we note that quantum teleportation has

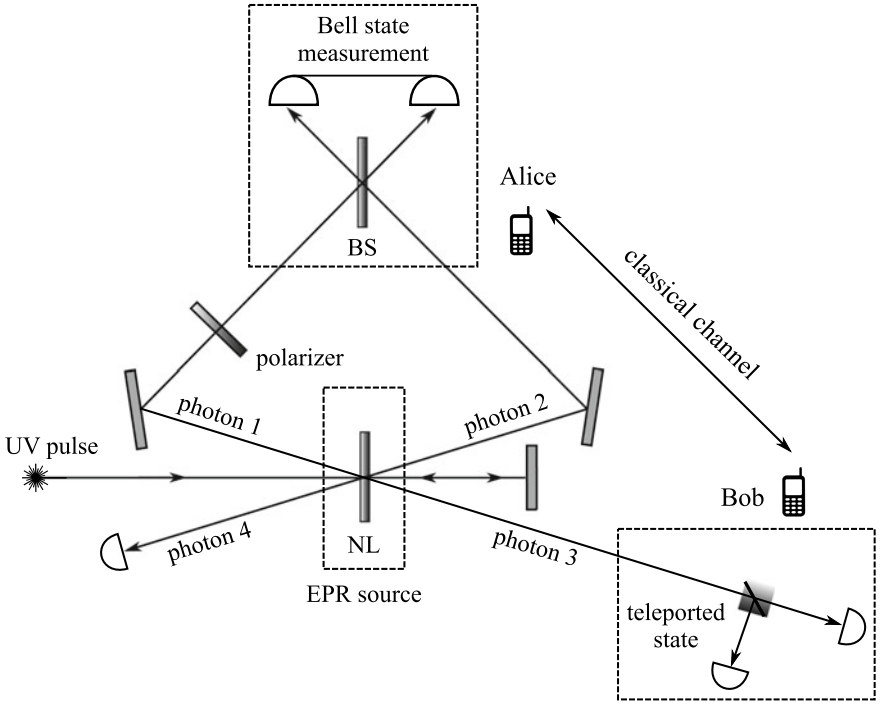


Fig. 3.19 Experimental realisation of quantum teleportation by Zeilinger and colleagues. A pulse of ultraviolet radiation passing through a nonlinear crystal NL produces the EPR photon pair 2–3. After retroflection during its second passage through NL, the radiation creates another pair, 1–4, of which the photon 1 is that to be teleported and photon 4 is a trigger indicating that the other photon is underway. Alice looks for coincidence counts of photons 1 and 2, after the BS. Finally, Bob, after receiving the classical bit of information may retrieve the input state of photon 1 through appropriate detection. Adapted from Auletta and Wang (2014, p. 368)

been experimentally realised, first with photons.[196] The set-up of a typical quantum teleportation experiment is schematically shown in Fig. 3.19.

Classical and Quantum Channels

In conclusion, quantum teleportation is based on two channels:

- A *quantum channel* (the entangled bit or *ebit* representing the correlation between qubits 2 and 3), by which Alice teleport to Bob the state $|\psi\rangle_1$ of qubit 1, and
- A *classical channel*, which we ideally assume to be noiseless (Sect. 3.2.4) and by which Alice communicates to Bob the result of her measurement on systems 1–2.

[196] Bouwmeester et al. (1997), Furusawa et al. (1998).

Thus, an *ebit* is the amount of entanglement between a maximally entangled pair of qubits (for example, two spin $\frac{1}{2}$ particles in a singlet state). After having received the classical information about Alice's measurement result, and hence without violating Einstein's locality requirement, Bob uses the entanglement between qubits 2 and 3 to recover exactly the state $|\psi\rangle_1$ on qubit 3. However, the original state $|\psi\rangle_1$ of qubit 1 is destroyed as a result of the entanglement between qubits 1 and 2 created by Alice's measurement. We note that this is in accordance with the no-cloning theorem, which dictates the impossibility to clone arbitrary unknown quantum states.

Then, by teleportation we transmit a qubit by means of (i) a shared ebit and (ii) two bits of classical information. Note that an ebit is a weaker resource than a qubit. In fact, the 'transmission' of a qubit can always be used to create one ebit (as it is evident with premeasurement, like for mapping (2.43)), while the sharing of an ebit (or many ebits) does not suffice to 'transmit' a qubit as we also need e.g. classical information to make an appropriate state rotation.

Teleportation shows how helpful is to interpret quantum information in specific experimental contexts as potential information. In fact, the entanglement between particles 2 and 3 (the quantum channel) can be considered as reservoir of *additional* information that can be used at any later time for transferring some qubit.[197] In fact, for quantum states, the conditional entropy can be negative,[198] an issue that demands to be interpreted in quantum-mechanical terms later on. Now, when the conditional entropy appears to be negative in a teleportation protocol, Alice and Bob share pure maximally entangled states which can be used to teleport quantum states between the two parties using only classical communication. The conditional entropy plays the same role in quantum information theory as it does in the classical theory, except that here, the quantum conditional entropy can be negative in an operationally meaningful way. In fact, negative conditional entropy gives Bob the potential to receive future quantum information for free, so that we can say that negative conditional entropy is just the potential future communication gained.

Quantum Cryptography

A related development dealing with information is represented by *quantum cryptography* that makes use of entanglement for establishing a common key between two distant partners without making necessary to send it physically, what is always exposed to the risk of eavesdropping. Quantum cryptography was initiated by a pioneering study of the Israeli physicist Stephen Wiesner, proposed in the 1970s but published only in 1983.[199] The computer scientists C. Bennett and G. Brassard published in 1984 a seminal paper in which they proposed the first protocol for quantum key distribution.[200] In Bennett and Brassard's protocol, now know universally as the

[197] Horodecki et al. (2005), Cavalcanti et al. (2011).

[198] Cerf and Adami (1997).

[199] Wiesner (1983).

[200] Bennett and Brassard (1984).

BB84 protocol, Alice and Bob are connected by a quantum communication channel which allows the 'transmission' of quantum states. In addition, they may communicate via a public classical channel. Both these channels may possibly be not secure. In particular, if the quantum channel is represented by the transmission of photons (e.g. in an optical fibre or in free space), they can make use of two possible bases of polarisation as alphabets:

$$+ = \{|\Leftrightarrow\rangle\ , |\Updownarrow\rangle\} \quad \text{and} \quad \times = \{\left|\nearrow\right\rangle\ , \left|\nwarrow\right\rangle\}, \tag{3.176}$$

which represent, respectively, the vertical–horizontal polarisation basis and the 45°–135° polarisation basis. First, Alice and Bob establish a one-to-one correspondence between each basis state of the two bases and the classical bits 0 and 1 that they desires to communicate with each other. They may agree to use, for instance, the following codification

$$0 \longleftrightarrow |\Leftrightarrow\rangle\ , \ \left|\nearrow\right\rangle\ , \quad \text{and} \quad 1 \longleftrightarrow |\Updownarrow\rangle\ , \ \left|\nwarrow\right\rangle\ . \tag{3.177}$$

Then, Alice sends to Bob a sequence of classical bits, encoded in photon polarisation states using one or the other polarisation basis chosen at random. Bob decodes the encoded bits by measuring the polarisation state of the transmitted photons, using the basis randomly chosen from one of the two bases. After the measurement, they publicly inform each other via a classical channel which basis they have chosen to use for each transmitted bit. They can immediately discard the bits for which the two chosen bases do not match (on average 50% of the transmitted bits). The bits for which Alice and Bob used the same basis constitute a shared secret key. An example of the transmission sequence is schematically shown in Table 3.7. Although eavesdropping by a third agency (conventionally called Eve) is still possible also with a quantum cryptography protocol, the disturbance that it will induce on quantum systems makes it relative easy to detect it, as displayed in Table 3.8. Another issue is whether Alice herself decides to cheat, a problem that will be discussed later. Anyway, also cryptography is clear evidence of how relevant are the effects of shared information although it does not represent a form of mechanical causation.

Information as Abstract Entity

Some scholars do worry about the abstract nature of information. The previous examination shows indeed that it is a *type* and not a token: tokens are often taken to be physical individual systems (provided that they can be discriminated) or also properties, while types are *equivalence classes of tokens* or other types. In fact, information is invariant under (see Sect. 3.2.3): (i) the physical nature of the medium which bears it (we can in fact perform essentially the same kind of experiments and with the same results using different kinds of quantum systems, as it is evident for, e.g. teleportation), (ii) the alphabet size (and this is why everything can be finally codified with a binary alphabet), (iii) translation from one alphabet to another (which, for quantum

Table 3.7 An example of the transmission sequence in the BB84 protocol in the absence of eavesdroppers. At the end of the protocol, Alice and Bob have the shared key 011010

Alice's bits	0	1	1	1	0	0	1	0	0	1
Alice's basis	+	+	×	×	×	+	+	×	+	+
Alice's states	$\|\leftrightarrow\rangle$	$\|\updownarrow\rangle$	$\|\nwarrow\rangle$	$\|\nwarrow\rangle$	$\|\nearrow\rangle$	$\|\leftrightarrow\rangle$	$\|\updownarrow\rangle$	$\|\nearrow\rangle$	$\|\leftrightarrow\rangle$	$\|\updownarrow\rangle$
Bob's basis	+	×	×	×	×	×	+	+	+	×
Bob's states	$\|\leftrightarrow\rangle$	$\|\nearrow\rangle$	$\|\nwarrow\rangle$	$\|\nwarrow\rangle$	$\|\nearrow\rangle$	$\|\nwarrow\rangle$	$\|\updownarrow\rangle$	$\|\leftrightarrow\rangle$	$\|\leftrightarrow\rangle$	$\|\nearrow\rangle$
Shared key	0	–	1	1	0	–	1	–	0	–

Table 3.8 An example of the transmission sequence in the BB84 protocol in the presence of an eavesdropper. At the end of the protocol, Alice and Bob are left with the partially correlated bit strings 011010 and 010010, respectively. If Alice and Bob will compare their sequences, they will discover that the third bit of these two sequences (which is 1 for Alice but 0 for Bob) presents a mismatch and this will reveal the presence of the eavesdropper

Alice's bits	0	1	1	1	0	0	1	0	0	1
Alice's basis	+	+	×	×	×	+	+	×	+	+
Alice's states	$\|\leftrightarrow\rangle$	$\|\updownarrow\rangle$	$\|\nwarrow\rangle$	$\|\nwarrow\rangle$	$\|\nearrow\rangle$	$\|\leftrightarrow\rangle$	$\|\updownarrow\rangle$	$\|\nearrow\rangle$	$\|\leftrightarrow\rangle$	$\|\updownarrow\rangle$
Eve's basis	×	×	+	+	×	+	+	+	×	+
Eve's states	$\|\nearrow\rangle$	$\|\nearrow\rangle$	$\|\updownarrow\rangle$	$\|\leftrightarrow\rangle$	$\|\nearrow\rangle$	$\|\leftrightarrow\rangle$	$\|\updownarrow\rangle$	$\|\leftrightarrow\rangle$	$\|\nwarrow\rangle$	$\|\updownarrow\rangle$
Bob's basis	+	×	×	×	×	×	+	+	+	×
Bob's states	$\|\leftrightarrow\rangle$	$\|\nearrow\rangle$	$\|\nwarrow\rangle$	$\|\nearrow\rangle$	$\|\nearrow\rangle$	$\|\nwarrow\rangle$	$\|\updownarrow\rangle$	$\|\leftrightarrow\rangle$	$\|\leftrightarrow\rangle$	$\|\nearrow\rangle$
Check	0	–	1	0	0	–	1	–	0	–

information, coincides with unitary transformations for one code to another).[201] In such a way, information is best defined as an *equivalence class* of sequences with respect to the previous three transformations, and the most convenient representative of information is the shortest binary sequence which belongs to such a class. This is called the *information message*. In other words, (quantum) information needs to

[201] Battail (2014, p. 15).

be treated as a type theory.[202] This is the point, since, on the one hand all physical phenomena by definition involve tokens, but, on the other hand, as we have seen, information makes the specific physical details irrelevant relative to the procedures of information codification, sharing and communication *without interfering* in any way with those physical details. This enables us to justify C. Bennett's dictum, reported in Sect. 3.2.3, according to which information is the "distinguishability abstracted away from […] the carrier of information".

It is obviously true that information, in all the mentioned experiments, can be effective only through ordinary physical means and processes, so that "one should not be seeking in an information-theoretic protocol—quantum or otherwise—for some particular, denoted by 'the information', whose path one is to follow, but rather concentrating on the physical processes by which the information is transmitted, that is, by which the end result of the protocol is brought about".[203] This is certainly right and it is in fact what happens in the real world. However, there are some subtle differences that concern the notion of causality that force us to focus on the way in which information is dealt with, and this will be object of further analysis in the next section.

3.3.7 What Is Non-separability

Eberhard's Theorem

In 1978 by the French physicist Philippe Eberhard proved an important theorem.[204] Let us consider the problem in all its generality, i.e. in the following we do not confine our examination to the specific spin model previously introduced. To this purpose, we consider a composite system $\mathcal{S}$ consists of two subsystems $\mathcal{S}_1$ and $\mathcal{S}_2$, which are set to be measured with apparata A and B, respectively. Let $\hat{O}_1$ and $\hat{O}_2$ be the respective observables of systems $\mathcal{S}_1$ and $\mathcal{S}_2$. Given that the respective settings of apparata A and B are $\mathbf{a}$ and $\mathbf{b}$, suppose that the probability that a joint measurement of $\hat{O}_1$ on $\mathcal{S}_1$ and $\hat{O}_2$ on $\mathcal{S}_2$ yield the respective results o_a and o_b is given by the conditional probability $\wp(o_a, o_b|\mathbf{a}, \mathbf{b})$. Let $\wp(o_a|\mathbf{a})$ denote the probability that a measurement of $\hat{O}_1$ on $\mathcal{S}_1$ yield the outcome o_a given the setting of apparatus A is $\mathbf{a}$, regardless of the setting of apparatus B and the measurement outcome of $\hat{O}_2$ on $\mathcal{S}_2$. From the addition rule

$$\wp(A|B) = \sum_j \wp(A_j|B), \tag{3.178}$$

where A, B are event sets and $\sum_j A_j = A$, the conditional probability $\wp(o_a|\mathbf{a})$ is obtained by summing the joint probabilities $\wp(o_a, o_b|\mathbf{a}, \mathbf{b})$ over all the possible

[202] See Abramsky and Coecke (2009).

[203] As pointed out in Timpson (2013, Sects. 3.7 and 4.1).

[204] Eberhard (1978). See also Auletta (2000, Sect. 36.5), Auletta and Wang (2014, Sect. 10.6).

outcomes o_b of $\hat{O}_2$ on $\mathcal{S}_2$ (I recall that it is called marginal probability), i.e.

$$\wp(o_a|\mathbf{a}) = \sum_{o_b} \wp(o_a, o_b|\mathbf{a}, \mathbf{b}). \tag{3.179a}$$

A similar probability $\wp(o_b|\mathbf{b})$ can be defined for the measurement outcome o_b of $\hat{O}_2$ on $\mathcal{S}_2$ when the setting of apparatus B is $\mathbf{b}$, regardless of the setting of apparatus A and the measurement outcome of $\hat{O}_1$ on S_1:

$$\wp(o_b|\mathbf{b}) = \sum_{o_a} \wp(o_a, o_b|\mathbf{a}, \mathbf{b}). \tag{3.179b}$$

According to Eberhard, if the above requirements were violated we would have a *non-local causal interdependence* between the two subsystems. This is because, by changing the setting of one apparatus, we would be able to have influence on the measurement outcomes of the other apparatus, and hence, if we performed experiments on subsystems that are spacelike separated, i.e. not connected by a light signal (Sect. 2.3.2), we would be able to transmit a message at superluminal or even infinite speed. Actually, the violation of the above requirement does not necessarily imply a non-local causal interconnection because there could still be some form of interdependence between the possible *settings* of the apparata. As I have mentioned in Sect. 2.3.2 and we shall see in detail in what follows, to exclude also the latter possibility sets a stronger bound than to exclude a violation of Einstein's locality. Eberhard himself interpreted the theorem in the weaker form (no causal interconnection among settings) instead of the stronger formulation (no violation of the separability among settings), where separability is sufficient condition of the lack of causal interconnections between the settings. As a matter of fact, in order to recover the weaker form, the theorem should tell that, e.g. the probability (3.179a) holds for *any* setting $\mathbf{b}$. In other words, we should have summed not only on the possible outcomes but also on the possible settings.

In order to prove the theorem, we consider the situation in which we first perform a measurement of $\hat{O}_1$ on $\mathcal{S}_1$, followed by a subsequent measurement of $\hat{O}_2$ on $\mathcal{S}_2$. Let us denote by $|o_a, \mathbf{a}\rangle$ the state of subsystem $\mathcal{S}_1$ when the setting of apparatus A is $\mathbf{a}$ and the measurement outcome of $\hat{O}_1$ is o_a, and by $|o_b, \mathbf{b}\rangle$ the state of subsystem $\mathcal{S}_2$ when the setting of apparatus B is $\mathbf{b}$ and the measurement outcome of $\hat{O}_2$ is o_b. Then the conditional probability (3.179a) of obtaining the outcome o_a when the setting of apparatus A is $\mathbf{a}$, can be reformulated, in agreement with Eq. (1.379), in terms of the density matrix of the bipartite system $\hat{\rho}$ as

$$\wp(o_a|\mathbf{a}) = \mathrm{Tr}\left(\hat{P}_{o_a,\mathbf{a}}\,\hat{\rho}\right), \tag{3.180}$$

where $\hat{P}_{o_a,\mathbf{a}}$ is the projector on the state $|o_a, \mathbf{a}\rangle$ of $\mathcal{S}_1$, i.e.

$$\hat{P}_{o_a,\mathbf{a}} = |o_a, \mathbf{a}\rangle\langle o_a, \mathbf{a}|. \tag{3.181}$$

After the measurement of $\hat{O}_1$ on S_1 when the setting of apparatus A is $\mathbf{a}$ and the outcome is o_a, the density matrix $\hat{\rho}$ reduces to (see Eq. (3.14))

$$\hat{\rho}' = \frac{\hat{P}_{o_a,\mathbf{a}}\,\hat{\rho}\,\hat{P}_{o_a,\mathbf{a}}}{\wp(o_a|\mathbf{a})}. \tag{3.182}$$

If we perform a subsequent measurement of $\hat{O}_2$ on $\mathcal{S}_2$, then the conditional probability of obtaining the outcome o_b when the setting of apparatus B is $\mathbf{b}$, is given by

$$\begin{aligned}\wp(o_b|o_a,\mathbf{a},\mathbf{b}) &= \mathrm{Tr}\left(\hat{P}_{o_b,\mathbf{b}}\,\hat{\rho}'\right)\\ &= \frac{\mathrm{Tr}\left(\hat{P}_{o_b,\mathbf{b}}\hat{P}_{o_a,\mathbf{a}}\,\hat{\rho}\,\hat{P}_{o_a,\mathbf{a}}\right)}{\wp(o_a|\mathbf{a})},\end{aligned} \tag{3.183}$$

where $\hat{P}_{o_b,\mathbf{b}}$ is the projector on the state $|\,o_b,\mathbf{b}\rangle$ of S_2, i.e.

$$\hat{P}_{o_b,\mathbf{b}} = |\,o_b,\mathbf{b}\rangle\,\langle o_b,\mathbf{b}\,|\,. \tag{3.184}$$

Therefore, the joint probability of obtaining the outcomes o_a and o_b, given the respective settings $\mathbf{a}$ and $\mathbf{b}$ of apparatus A and B, thanks to the general formula for probabilities of events A, B

$$\wp(A,B) = \wp(A)\wp(B|A), \tag{3.185}$$

is given by

$$\begin{aligned}\wp(o_a,o_b|\mathbf{a},\mathbf{b}) &= \wp(o_a|\mathbf{a})\,\wp(o_b|o_a,\mathbf{a},\mathbf{b})\\ &= \wp(o_a|\mathbf{a})\,\frac{\mathrm{Tr}\left(\hat{P}_{o_b,\mathbf{b}}\hat{P}_{o_a,\mathbf{a}}\hat{\rho}\hat{P}_{o_a,\mathbf{a}}\right)}{\wp(o_a|\mathbf{a})}\\ &= \mathrm{Tr}\left(\hat{P}_{o_b,\mathbf{b}}\hat{P}_{o_a,\mathbf{a}}\hat{\rho}\hat{P}_{o_a,\mathbf{a}}\right).\end{aligned} \tag{3.186}$$

By using the cyclic property of the trace (1.371c), the fact that the projectors $\hat{P}_{o_a,\mathbf{a}}$ and $\hat{P}_{o_b,\mathbf{b}}$ commute because they pertain to different subsystems, and the general property $\hat{P}^2_{o_a,\mathbf{a}} = \hat{P}_{o_a,\mathbf{a}}$ (see Eq. (1.90)), we may further simplify the above expression to

$$\begin{aligned}\wp(o_a,o_b|\mathbf{a},\mathbf{b}) &= \mathrm{Tr}\left(\hat{P}_{o_a,\mathbf{a}}\hat{P}_{o_b,\mathbf{b}}\hat{P}_{o_a,\mathbf{a}}\,\hat{\rho}\right)\\ &= \mathrm{Tr}\left(\hat{P}_{o_b,\mathbf{b}}\hat{P}_{o_a,\mathbf{a}}\hat{P}_{o_a,\mathbf{a}}\,\hat{\rho}\right)\\ &= \mathrm{Tr}\left(\hat{P}_{o_b,\mathbf{b}}\hat{P}_{o_a,\mathbf{a}}\,\hat{\rho}\right).\end{aligned} \tag{3.187}$$

Now, we sum the above result over all possible outcomes of o_a in Eq. (3.187) in order to write the probability $\wp(o_b|\mathbf{b})$ as

$$\begin{aligned}\wp(o_b|\mathbf{b}) &= \sum_{o_a} \wp(o_a, o_b|\mathbf{a}, \mathbf{b}) \\ &= \sum_{o_a} \mathrm{Tr}\left(\hat{P}_{o_b,\mathbf{b}} \hat{P}_{o_a,\mathbf{a}}\, \hat{\rho}\right) \\ &= \mathrm{Tr}\left[\hat{P}_{o_b,\mathbf{b}} \left(\sum_{o_a} \hat{P}_{o_a,\mathbf{a}}\right) \hat{\rho}\right] \\ &= \mathrm{Tr}\left(\hat{P}_{o_b,\mathbf{b}}\, \hat{\rho}\right), \end{aligned} \tag{3.188}$$

where use has been made of the property that $\sum_{o_a} \hat{P}_{o_a,\mathbf{a}} = \hat{I}$ for any complete set of orthogonal projectors. We note that the result for $\wp(o_b|\mathbf{b})$ given by the last equality in Eq. (3.188) is indeed independent of the setting $\mathbf{a}$ of apparatus A. The same conclusion applies to $\wp(o_a|\mathbf{a})$ if we start the proof by considering a measurement of $\hat{O}_2$ on $\mathcal{S}_2$, followed by a measurement of $\hat{O}_1$ on $\mathcal{S}_1$. In fact, we have:

$$\wp(o_a|o_b, \mathbf{a}, \mathbf{b}) = \mathrm{Tr}\left(\hat{P}_{o_a,\mathbf{a}}\, \hat{\rho}'\right) \tag{3.189}$$

$$= \frac{\mathrm{Tr}\left(\hat{P}_{o_a,\mathbf{a}} \hat{P}_{o_b,\mathbf{b}}\, \hat{\rho}\, \hat{P}_{o_b,\mathbf{b}}\right)}{\wp(o_b|\mathbf{b})}, \tag{3.190}$$

where

$$\hat{\rho}' = \frac{\hat{P}_{o_b,\mathbf{b}}\, \hat{\rho}\, \hat{P}_{o_b}}{\wp(o_b|\mathbf{b})}. \tag{3.191}$$

Then, it follows:

$$\wp(o_a, o_b|\mathbf{a}, \mathbf{b}) = \mathrm{Tr}\left(\hat{P}_{o_a,\mathbf{a}} \hat{P}_{o_b,\mathbf{b}}\, \hat{\rho}\, \hat{P}_{o_b,\mathbf{b}}\right) \tag{3.192}$$

$$= \mathrm{Tr}\left(\hat{P}_{o_a,\mathbf{a}} \hat{P}_{o_b,\mathbf{b}}\, \hat{\rho}\right), \tag{3.193}$$

that allows us to finally obtain:

$$\wp(o_a|\mathbf{a}) = \sum_{o_b} \wp(o_a, o_b|\mathbf{a}, \mathbf{b}) \tag{3.194}$$

$$= \sum_{o_b} \mathrm{Tr}\left(\hat{P}_{o_a,\mathbf{a}} \hat{P}_{o_b,\mathbf{b}}\, \hat{\rho}\right) \tag{3.195}$$

$$= \mathrm{Tr}\left[\hat{P}_{o_a,\mathbf{a}} \left(\sum_{o_b} \hat{P}_{o_b,\mathbf{b}}\right) \hat{\rho}\right]$$

$$= \mathrm{Tr}\left(\hat{P}_{o_a,\mathbf{a}}\,\hat{\rho}\right), \tag{3.196}$$

where again I have used the property that $\sum_{o_b} \hat{P}_{o_b,\mathbf{b}} = \hat{I}$ for any complete set of orthogonal projectors. In conclusion, settings satisfy in QM the separability principle. This demands a further deepening of our analysis.

Popescu and Rohrlich's Locality Requirement

In Sect. 2.3.2 I have already mentioned that the physicists S. Popescu and D. Rohrlich established that systems satisfying the locality requirement show a bound of the CHSH inequality (3.153) that is much higher than the separability requirement[205]:

$$\left|\langle \mathbf{a}, \mathbf{b}\rangle + \langle \mathbf{a}, \mathbf{b}'\rangle + \langle \mathbf{a}', \mathbf{b}\rangle - \langle \mathbf{a}', \mathbf{b}'\rangle\right| = |\left\langle \hat{\mathrm{B}}\right\rangle| \leq 4, \tag{3.197}$$

where $\hat{\mathrm{B}}$ is the Bell operator (3.156) and I have slightly changed the expression. This can be proved as follows. The only requirement of relativistic (Einstein's) locality is that the operations one perform locally here are not influenced by *signals* due to the operations someone performs far away elsewhere. This implies in particular that the probability that one obtains a certain outcome (say 1) when choosing the direction $\mathbf{a}$ is independent from the outcomes (either $+1$ or -1) when someone elsewhere chooses a direction $\mathbf{b}$ or $\mathbf{b}'$, that is, we have the conditional probabilities

$$\wp(1, 1|\mathbf{a}, \mathbf{b}) + \wp(1, -1|\mathbf{a}, \mathbf{b}) = \wp(1, 1|\mathbf{a}, \mathbf{b}') + \wp(1, -1|\mathbf{a}, \mathbf{b}'). \tag{3.198}$$

Similar considerations hold for arbitrary directions. If we consider only this requirement, we are allowed to build the set of probabilities

$$\left.\begin{array}{l} \wp(1, 1|\mathbf{a}, \mathbf{b}) = \wp(-1, -1|\mathbf{a}, \mathbf{b}) \\ \wp(1, 1|\mathbf{a}, \mathbf{b}') = \wp(-1, -1|\mathbf{a}, \mathbf{b}') \\ \wp(1, 1|\mathbf{a}', \mathbf{b}) = \wp(-1, -1|\mathbf{a}', \mathbf{b}) \\ \wp(1, -1|\mathbf{a}', \mathbf{b}') = \wp(-1, 1|\mathbf{a}', \mathbf{b}') \end{array}\right\} = \frac{1}{2} \tag{3.199}$$

while assuming that all of the other probabilities are zero. I Therefore, the expectation value (3.154) reduces to

$$\langle \mathbf{a}, \mathbf{b}\rangle = \wp(1, 1|\mathbf{a}, \mathbf{b}) + \wp(-1, -1|\mathbf{a}, \mathbf{b}). \tag{3.200a}$$

Similarly, for the other three expectation values we have

[205] The original paper is Popescu and Rohrlich (1994). A summary of this subject can be found in Auletta and Wang (2014, Sect. 12.2).

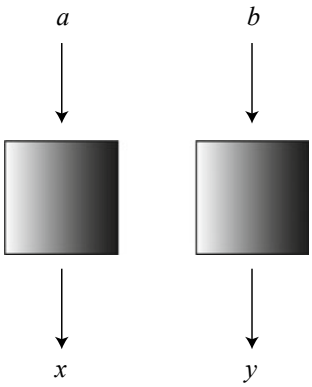


Fig. 3.20 PR boxes: they are devices that fulfil the locality (non-signalling) but not the separability condition. Alice takes inputs a's and get output x's while Bob takes inputs b's and get output y's. Since the only requirement of PR boxes is no-signalling, between the two boxes there could be (classical or quantum) correlations

$$\langle \mathbf{a}, \mathbf{b}' \rangle = \wp(1, 1|\mathbf{a}, \mathbf{b}') + \wp(-1, -1|\mathbf{a}, \mathbf{b}'), \tag{3.200b}$$
$$\langle \mathbf{a}', \mathbf{b} \rangle = \wp(1, 1|\mathbf{a}', \mathbf{b}) + \wp(-1, -1|\mathbf{a}', \mathbf{b}), \tag{3.200c}$$
$$\langle \mathbf{a}', \mathbf{b}' \rangle = -\wp(1, -1|\mathbf{a}', \mathbf{b}') - \wp(-1, 1|\mathbf{a}', \mathbf{b}'). \tag{3.200d}$$

In this way, as a consequence of probabilities (3.199), we obtain the upper bound 4 of the inequality (3.197).

PR Boxes and Information

The ideal devices used for dealing with this kind of problems are called *PR boxes* (Fig. 3.20). This allows us to rewrite the set of probabilities (3.199) as $\wp(x, y|a, b)$. It is helpful to use binary codification for all variables here involved[206]: instead, of having results $+1, -1$, we write $x, y = 0, 1$, while we set $\mathbf{a}, \mathbf{b}$ as $a, b = 0$ and $\mathbf{a}', \mathbf{b}'$ as $a, b = 1$. This allows us to rewrite the 4 terms of the CHSH inequality in terms of correlations C_{xy} as follows

$$\mathrm{C}_{00} = \wp(11|00) + \wp(00|00), \tag{3.201a}$$
$$\mathrm{C}_{01} = \wp(11|01) + \wp(00|01), \tag{3.201b}$$
$$\mathrm{C}_{10} = \wp(11|10) + \wp(00|10), \tag{3.201c}$$
$$\mathrm{C}_{11} = -\wp(10|11) - \wp(01|11), \tag{3.201d}$$

and the whole CHSH expression, being the expectation value of the Bell operator, as

$$\mathrm{B} = \mathrm{C}_{00} + \mathrm{C}_{01} + \mathrm{C}_{10} - \mathrm{C}_{11}. \tag{3.202}$$

By making use of the modulo–2 addition $\oplus$, whose rules are given by

[206] Auletta (2011b).

$$0 \oplus 0 = 0, \quad 0 \oplus 1 = 1 \oplus 0 = 1, \quad 1 \oplus 1 = 0, \tag{3.203}$$

we can rewrite the set of probabilities (3.199) as

$$\wp_{PR}(x, y|a, b) = \frac{1}{2}\delta_{x\oplus y=ab}, \tag{3.204}$$

where the subscript PR denotes PR boxes and δ is the Kronecker delta (1.66).[207] In fact, according to the rules of modulo–2 addition, when $x \oplus y = 0$, the outputs x and y need to be equal (according to the first three probabilities (3.199)), what means that at least one of the couple a, b is zero (that is, at least one of the settings is unprimed), and therefore their product is 0. However, when $x \oplus y = 1$, this means that the outputs x and y need to be different (according to the fourth probability) and both a and b are 1, that is, both settings are primed, and their product is 1.

Tsirelson's Theorem

Now, as shown in Sect. 3.3.5, QM violates the bound 2. Therefore, the natural question arises of what is the exact bound for quantum systems. Is it 4? The Russian-Israeli mathematician Boris Tsirelson proved that is lower and corresponds to (see also Eq. (3.160))[208]:

$$\left|\left\langle \hat{\mathcal{B}} \right\rangle\right| \leq 2\sqrt{2}, \tag{3.205}$$

where $\left\langle \hat{\mathcal{B}} \right\rangle$ is the mean value (3.157) of a generalisation of the Bell operator. This inequality can be proved in the following way. Let $\hat{O}_a$, $\hat{O}_{a'}$, $\hat{O}_b$, $\hat{O}_{b'}$ be arbitrary Hermitian operators on a 2D Hilbert space (in this general proof we no longer need to deal with the spin observable), each having eigenvalues 1 and -1 and satisfying the conditions $[\hat{O}_a, \hat{O}_b] = 0$, and so on, for the other pairs (a, b'), (a', b), and (a', b'). Then, we can prove that the following inequality holds in QM:

$$\left|\langle \hat{O}_a \hat{O}_b \rangle + \langle \hat{O}_{a'} \hat{O}_b \rangle + \langle \hat{O}_a \hat{O}_{b'} \rangle - \langle \hat{O}_{a'} \hat{O}_{b'} \rangle\right| \leq 2\sqrt{2}. \tag{3.206}$$

In such a context, the Bell operator (3.156) is generalised as

$$\hat{\mathcal{B}} = \hat{O}_a \hat{O}_b + \hat{O}_{a'} \hat{O}_b + \hat{O}_a \hat{O}_{b'} - \hat{O}_{a'} \hat{O}_{b'}. \tag{3.207}$$

This allows us to define a new Hermitian operator $\hat{\mathcal{A}}$ by

[207] Pawłowski and Scarani (2016).

[208] Tsirelson (1980).

$$\begin{aligned}\hat{\mathcal{A}} &= 2\sqrt{2}\hat{I} - \hat{\mathcal{B}} \\ &= \frac{1}{\sqrt{2}}\left(\hat{O}_a^2 + \hat{O}_{a'}^2 + \hat{O}_b^2 + \hat{O}_{b'}^2\right) - \hat{\mathcal{B}},\end{aligned} \tag{3.208}$$

since, being by definition the square of each of this observables = 1, we have

$$\hat{O}_a^2 + \hat{O}_{a'}^2 + \hat{O}_b^2 + \hat{O}_{b'}^2 = 4\hat{I}. \tag{3.209}$$

Equation (3.208) can be rewritten as

$$\hat{\mathcal{A}} = \frac{1}{\sqrt{2}}\left[\left(\hat{O}_a - \frac{\hat{O}_b + \hat{O}_{b'}}{\sqrt{2}}\right)^2 + \left(\hat{O}_{a'} - \frac{\hat{O}_b - \hat{O}_{b'}}{\sqrt{2}}\right)^2\right], \tag{3.210}$$

what shows that $\hat{\mathcal{A}}$ consists in a sum of squares of Hermitian operators, and it is evidently positive semidefinite and has a non-negative expectation value (see Eqs. (1.88b)), that is,

$$\left\langle\hat{\mathcal{A}}\right\rangle = 2\sqrt{2} - \left\langle\hat{\mathcal{B}}\right\rangle \geq 0, \tag{3.211}$$

which leads to

$$\left\langle\hat{\mathcal{B}}\right\rangle \leq 2\sqrt{2}. \tag{3.212a}$$

Thanks to a similar argument with the Hermitian operator $\hat{\mathcal{A}}' = 2\sqrt{2}\hat{I} + \hat{\mathcal{B}}$ we are able to derive

$$\left\langle\hat{\mathcal{B}}\right\rangle \geq -2\sqrt{2}. \tag{3.212b}$$

Therefore, we conclude that

$$\left|\left\langle\hat{\mathcal{B}}\right\rangle\right| \leq 2\sqrt{2}, \tag{3.213}$$

which proves the theorem. This is known as *Tsirelson's bound*.

Three Cases

The expression (3.202) for the CHSH inequality allows us to make use of a parameter D for writing[209]

$$\begin{aligned}\mathrm{D} &= \frac{\mathrm{B}}{2} - 1 \\ &= \frac{1}{2}(\mathrm{C}_{00} + \mathrm{C}_{01} + \mathrm{C}_{10} - \mathrm{C}_{11}) - 1.\end{aligned} \tag{3.214}$$

[209] Masanes et al. (2006).

Let us now consider the simplest case in which

$$C_{00} = C_{01} = C_{10} = -C_{11} = C > 0, \tag{3.215}$$

from which we deduce $B = 4C$. Then, we can rewrite the expression (3.214) as

$$D(C) = 2C - 1. \tag{3.216}$$

It is easy to distinguish, in terms of the value of C, the following three cases:

(i) Classical case: for $0 < C \leq \frac{1}{2}$, we have $-1 < D \leq 0$ and $0 < B \leq 2$. Indeed, in this case we have classical separability.
(ii) Quantum case: for $\frac{1}{2} < C \leq \frac{1}{\sqrt{2}}$ we have $0 < D \leq \sqrt{2} - 1$ and $2 < B \leq 2\sqrt{2}$. This is the case in which correlations are quantum mechanical and we have quantum non-separability.
(iii) Hyper-correlation case: for $\frac{1}{\sqrt{2}} < C \leq 1$ we have $\sqrt{2} - 1 < D \leq 1$ and $2\sqrt{2} < B \leq 4$. This is the case in which the causal no-signalling requirement is respected, but there exist correlations stronger than the strongest quantum correlations.

Principle of Information Causality

Thus, quantum systems are somehow intermediate between (i) classical systems and (ii) hyper-correlated systems that do not violate Einstein's locality (represented by the bound 4). We may ask why quantum systems satisfy the above bound. The reason is to find in the principle of information causality[210] that tells us that the information gain (for instance, during teleportation) that, by using all his local resources (which may be still correlated with Alice's resources) and n classical bits communicated to him by Alice, Bob can get about a data set of Alice that was previously unknown to him is at most n bits: this is in fact related to the Holevo bound (3.83).[211]

Suppose that Alice receives a random list $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ while in a separate location Bob receives a random variable $b \in (1, 2, \ldots, n)$. Then, Alice sends m classical bits to Bob and the latter is asked to guess the value of the kth bit of Alice's list: a_k. Then, if Bob's output is denoted by y_k, the maximal amount of information gain is given by the mutual information $I(a_k : y_k)$. This quantity, in order to satisfy the information causality principle need to be such that

$$\sum_{j=1}^{n} I(a_j : y_j) \leq m. \tag{3.217}$$

[210] Formulated in Pawłowski et al. (2009). See also Pawłowski and Scarani (2016).

[211] Another way to consider the problem is that hyper-correlations would violate the uncertainty relations (Oppenheim and Wehner 2010).

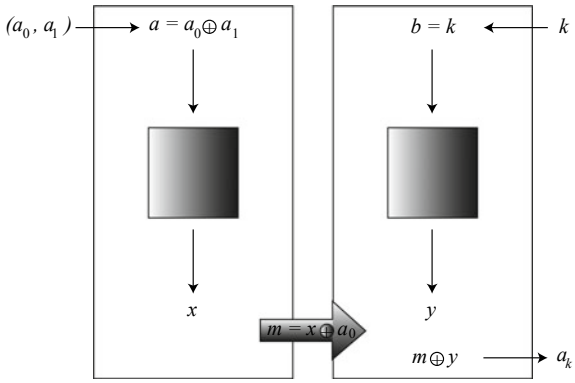


Fig. 3.21 Alice receives input $a = a_0 \oplus a_1$ and gets output x. Then, she communicates the classical one-bit information $m = x \oplus a_0$ to Bob

This result could seem obvious. To understand why it is important, let us assume that such a principle (and also Tsireslon's bound (3.213), which can be derived from the principle of information causality) were violated. This would imply that, when we received some (classical) information from a partner, thanks to correlations we could be able to get not only this piece of information but something additional. How and what? This would be possible in the case in which entanglement would provide us of immediate access to the *whole code alphabet* used by our partner (that allows us to know the whole of the a_k's). Now, this additional information we are looking for cannot be represented by another message that was never sent by hypothesis. In other words, we would know the whole code alphabet with a single communication. In such a case, without necessarily violating the bound of 4 (since we are hypothesising a correlation between codes without exchange of superluminal signals), we would certainly violate the Tsirelson bound. As we have seen, all observables involved in Tsirelson's theorem mutually commute, but the transmission of information over a distance could only take place, if we can identify operators at a certain (space-time) location and operators at another (space-time) location that do not commute. And this is not the case.[212]

This conclusion can be nicely shown with the following example.[213] We make use again of PR boxes (Fig. 3.21). Note that the example can be both classical or quantum mechanical. What is crucial is the classical one-bit communication from Alice to Bob. Now, with this single bit of classical information Bob was able to guess what is Alice's input, whether a_0 or a_1. On his side, Bob inputs $b = k$, reads y and computes $C = m \oplus y = a_0 \oplus x \oplus y$. This protocol is called Random Access Code (RAC). Now, if $k = 0$, then, according to Eq. (3.204), we know that $y = x$, and therefore $C = a_0$. On the other hand, if $k = 1$, $y = x \oplus a$ (clearly, the value of y depends on whether we have a_0 or a_1). However, in agreement with rules (3.203), we have

[212] 'T Hooft (2016, p. 32).

[213] Pawłowski and Scarani (2016).

$$\begin{aligned} C &= a_0 \oplus x \oplus x \oplus a \\ &= a_0 \oplus a \\ &= a_0 \oplus a_0 \oplus a_1 \\ &= a_1. \end{aligned} \tag{3.218}$$

Clearly, $k = a_k$ and thus Bob has access to the whole of Alice's input code alphabet (a_0, a_1). But Alice has transmitted only a single bit (a) and the PR boxes are supposed to be no-signalling so they cannot be used to transmit the other bit. Thus, it is only correlation between code alphabets that could explain this result when the sole requirement of locality is satisfied.

No Entanglement Between Code Alphabets

Therefore, what this discussion shows is that QM allows for entanglement *between possible outcomes* (outputs or events or also messages) but *not between code alphabets* (or codes in general).[214] Now, it is crucial to understand that to use a specific code alphabet is equivalent to premeasure, that is, to select a specific experimental set-up. Indeed, by choosing an alphabet for expressing a quantum state, we are in fact experimentally selecting the observable whose eigenstates are precisely those elements of the alphabet (represented by the axes in the reference frame of that observable): see Fig. 2.1, Sect. 2.1.2, and Fig. 3.6, Sect. 3.2.4, and relative comments. Thus, codification remains a *local* procedure (and a local choice) that cannot be shared through quantum channels only, but to this purpose always some additional form of classical communication is needed, since a quantum channel only represents a form of mutual information among possible *outcomes* (Sect. 3.2.2). The theorem of triorthogonal decomposition ensures us of the effectiveness of such a local choice (Sect. 3.1.3). This state of affairs is clearly shown by quantum cryptography where a common codification can be shared only through classical communication and implementation of different experimental settings (see the previous subsection). In conclusion, we can say that both measuring and codifying are local procedure. Ultimately, they depend on local choices, whatever the reason of these choices can be.[215]

Coupling Code Alphabets

A remark seems appropriate here. In Sect. 3.2.3, I have spoken of linear transfer of information during the premeasurement stage but also clarified that this cannot consist in pouring something from one system to another. This should be clear now because we have definitively excluded that entanglement consisted of some signal sending. From what we have said so far there is also an important postil: to entangle

[214] As pointed out in Auletta (2011b).

[215] Wiesner (1983).

systems means ultimately that these systems come to *couple* (but not to share) code alphabets, as it is evident from Eq. (2.43). This is connected with the fact that the information of quantum systems means nothing and refers to nothing. The same equation also shows that, for this premeasurement step, the codification of the information that can be potentially present in the initial state of the system does not affect the code of the apparatus, since, for acquiring information, what counts is the latter and the related physical context that we impose by choosing a certain observable of the object system. In this consists the linear transfer of information: it is an *arbitrary* connection between code alphabets but through a *physical operation*; if quantum systems would share code alphabets, the premeasurement step would be irrelevant, but this is in contradiction with our experience. To say "physical operation" allows us to recover the connection between the notion of observable and the notion of codification.

To show this, let us come back to the model of coupling introduced in Sect. 3.1.1, but inverting somehow the terms of the problem. In particular, I choose the operator $\hat{\sigma}_z^{\mathcal{S}}$ for the observable of $\mathcal{S}$ and $\hat{\sigma}_z^{\mathcal{M}}$ for the observable of the metre (see the matrices (1.324)). This is permitted by the circumstance that a state is entangled in different bases and therefore for different observables. The system $\mathcal{S}$ is initially prepared in a superposition of the two eigenstates of $\hat{\sigma}_z^{\mathcal{S}}$, i.e. $|\uparrow\rangle_{\mathcal{S}}, |\downarrow\rangle_{\mathcal{S}}$, given by Eq. (1.322). The metre is assumed to be initially in the x spin-down state, $|\downarrow_x\rangle_{\mathcal{M}}$. The interaction Hamiltonian can then be written as

$$\hat{H}_{\mathcal{SM}} = \varepsilon_{\mathcal{SM}} \left(1 + \hat{\sigma}_z^{\mathcal{S}}\right) \hat{\sigma}_z^{\mathcal{M}} \,, \tag{3.219}$$

where, again, $\varepsilon_{\mathcal{SM}}$ is some coupling function between the metre and the object system. Now, we need to write its initial state in terms of the eigenkets of the two systems

$$\begin{aligned} |\Psi(\tau)\rangle_{\mathcal{SM}} &= \hat{U}_{\tau}^{\mathcal{SM}} |\Psi(0)\rangle_{\mathcal{SM}} \\ &= e^{-\frac{i}{\hbar}\tau\varepsilon_{\mathcal{SM}}(1+\hat{\sigma}_z^{\mathcal{S}})\hat{\sigma}_z^{\mathcal{M}}} \left[\left(c_\uparrow |\uparrow\rangle_{\mathcal{S}} + c_\downarrow |\downarrow\rangle_{\mathcal{S}}\right) |\downarrow_x\rangle_{\mathcal{M}}\right] \,, \end{aligned} \tag{3.220}$$

where τ is the time interval of the interaction between $\mathcal{S}$ and $\mathcal{M}$. Then, we have

$$\begin{aligned} |\Psi(\tau)\rangle_{\mathcal{SM}} = \frac{1}{\sqrt{2}} \Big(& c_\uparrow e^{-\frac{2i}{\hbar}\tau\varepsilon_{\mathcal{SM}}} |\uparrow\rangle_{\mathcal{S}} |\uparrow\rangle_{\mathcal{M}} + c_\downarrow |\downarrow\rangle_{\mathcal{S}} |\uparrow\rangle_{\mathcal{M}} \\ & - c_\uparrow e^{+\frac{2i}{\hbar}\tau\varepsilon_{\mathcal{SM}}} |\uparrow\rangle_{\mathcal{S}} |\downarrow\rangle_{\mathcal{M}} - c_\downarrow |\downarrow\rangle_{\mathcal{S}} |\downarrow\rangle_{\mathcal{M}} \Big) \,. \end{aligned} \tag{3.221}$$

We finally obtain

$$\begin{aligned} |\Psi(\tau)\rangle_{\mathcal{SM}} = \frac{1}{\sqrt{2}} \Bigg(& -c_\uparrow \frac{2i}{\sqrt{2}} \sin\frac{2\tau\varepsilon_{\mathcal{SM}}}{\hbar} |\uparrow\rangle_{\mathcal{S}} |\uparrow_x\rangle_{\mathcal{M}} + \sqrt{2} c_\downarrow |\downarrow\rangle_{\mathcal{S}} |\downarrow_x\rangle_{\mathcal{M}} \\ & + c_\uparrow \frac{2i}{\sqrt{2}} \cos\frac{2\tau\varepsilon_{\mathcal{SM}}}{\hbar} |\uparrow\rangle_{\mathcal{S}} |\downarrow_x\rangle_{\mathcal{M}} \Bigg) \,. \end{aligned} \tag{3.222}$$

Choosing again

$$\frac{2\tau\varepsilon_{\mathcal{SM}}}{\hbar} = \frac{\pi}{2}\,, \quad \text{or} \quad \tau = \frac{\pi\hbar}{4\varepsilon_{\mathcal{SM}}}\,, \tag{3.223}$$

we kill the last term and get

$$|\,\Psi(\tau)\rangle_{\mathcal{SM}} = c_{\downarrow}\,|\downarrow\rangle_{\mathcal{S}}\,|\downarrow_x\rangle_{\mathcal{M}} - ic_{\uparrow}\,|\uparrow\rangle_{\mathcal{S}}\,|\uparrow_x\rangle_{\mathcal{M}}\,, \tag{3.224}$$

which works perfectly as in the original model but this time it is the x up state of $\mathcal{M}$ to be correlated with the z up state of $\mathcal{S}$ and similarly for the down states.

Summarising, coupling code alphabets is necessary for *acquiring information*. We are able to explain the linear transfer of information precisely by assuming that, by singling out an observable, *an objective code alphabet* is present in the object system (with the related orthogonality requirement) and the apparatus can couple its own states with such a codification, independently of the meaning or the physical nature of such states making such an acquiring possible. This eliminates any element of magic that could be presupposed to work in the linear transfer of information between object system and apparatus. Note that, in order to avoid the basis degeneracy problem, supporters of the MWI could in fact postulate that there is universal alphabet–sharing. In such a case, in fact, many problems previously discussed would clearly disappear. But such an assumption contradicts the results proved in this subsection. If, on the other hand, there were no alphabet coupling, no transfer of information would be possible. Thus, both ingredients are necessary.

LOCC Protocol

There is a subtle point here: although the initial code alphabet of the input system appears irrelevant to the measurement process and what counts is a coupling of code alphabets between it and the apparatus in a certain physical context (we are here in a controlled situation), it is not irrelevant in general, since it is what guarantees that quantum systems can exchange and acquire information also without the presence of an observing agency. Of course, the problem is trickier when we deal with systems whose entanglement we do not control. In such a case, alphabets are still coupled, but, as displayed in the case of quantum cryptography (discussed in the previous subsection), often we do not know which element with which element (e.g. whether horizontal polarisation in a photon is coupled with vertical or horizontal polarisation in the other one), if we do not exchange additional classical information with a partner or compare the detections statistics of many entangled systems in the same state. This is evident in the so-called Local Operations and Classical Communication(LOCC)

protocol.[216] In this case to Alice and Bob are given two entangled qubits and their task is to know whether the global state is, e.g. $\left|\Psi^{+}\right\rangle$ or $\left|\Phi^{+}\right\rangle$, given by Eqs. (3.170a) and (3.170c). Now, suppose that Alice measure her qubit and gets $|0\rangle$. She can classical communicate such a result to Bob. Now, the latter can discriminate among the two EPR states by measuring his own qubit. If he get $|1\rangle$, the global state is $\left|\Psi^{+}\right\rangle$, if $|0\rangle$ it is $\left|\Phi^{+}\right\rangle$. In other words, only the comparison between these two results allows to know which is in fact the entangled state. Thus, coupling code alphabets is very different from sharing the same alphabet, an operation that we have shown to be impossible in absence of exchange of classical information.

Conclusion

Because QM satisfies and saturates the bound imposed by the principle of information causality, and because in so doing it also sets specific constraints on the possible correlations in our universe, the fact that non-local encoding of information is forbidden by quantum physics justifies quantum information as a general theory of information. However, in order to get a consistent theory of causality that go further than the no-go theorem or the principle of information causality, we need to deepen our analysis of the causal effect of correlations.

3.4 Back to the Problem of Causality

3.4.1 Form, Structure and Information

Information is a formal entity but according to the interpretation that we are developing here is also able to have real influence (Sect. 3.3.6). However, we have agreed that this influence cannot be understood in terms of mechanical causation (Sect. 3.3.7). To assume the opposite could not only lead to strong non-local consequences (like the later interpretation formulated by Bohm), but seems also conceptually untenable. If information is a formal entity, it cannot act mechanically for a basic reason: it cannot be a dynamical cause and therefore a fortiori it cannot be mechanical.[217] Nevertheless, apart from the processes involved in local information acquisition, information can be shared as for entangled systems. In that case, the result can be a network of correlated systems. Then, depending on the context, we can have the arising of structures.[218] In other words, I am trying to split here the issue of information as such

[216] See the summary in Nielsen and Chuang (2000, Sect. 12.5.1).

[217] As pointed out in Auletta (2011a, Chap. 2).

[218] The connection between network of entangled systems and non-locality has been explored in Cavalcanti et al. (2011).

and the problem of signal exchanging (that is fundamental for mechanical causality: Sect. 2.3.2) together with information acquisition. While the latter must rely on local processes and involves mechanical causes, for the former it is not necessarily so.

GHSZ States

To understand this formal–structural character of quantum systems, let us consider a simple example: a three-particle state. The physicists D. Greenberger, M. Horne, A. Shimony, and A. Zeilinger (GHSZ) have been able to prove that QM violates separability in the spirit of Bell proofs but without using inequalities at all. The authors studied a particle with zero mean momentum that (radioactively) decays into three corpuscles.[219] If all three daughter particles have the same energy, by momentum conservation they must be emitted 120° apart from each other. The central source is surrounded by an array of six apertures: a, b and c at 120° separation, and a', b' and c' also at 120° separation with respect to each other (Fig. 3.22). Because of the placement of apertures, the three particles 1, 2 and 3 must emerge either through a, b and c or through a', b' and c', respectively. Thus, the state of the three particles beyond the apertures will be given by the superposition of these two possibilities, called a *GHSZ state*,

$$|\Psi_{GHSZ}\rangle = \frac{1}{\sqrt{2}}\left(|a\rangle_1|b\rangle_2|c\rangle_3 + |a'\rangle_1|b'\rangle_2|c'\rangle_3\right) . \tag{3.225}$$

Beyond the apertures, beams $|a\rangle_1$ and $|a'\rangle_1$ are totally reflected so as to overlap at a $50-50$ BS, and the two outgoing beams are monitored by detectors D_1 and D'_1. Suppose also that the component $|a'\rangle_1$ passes through a phase plate which causes a phase shift ϕ_1. A similar arrangement is at place for the other beams. Consequently, in agreement with Eqs. (2.28) and (2.30), we have the unitary transformations

$$|a\rangle_1 \longmapsto \frac{1}{\sqrt{2}}(|\mathrm{D}_1\rangle + |\mathrm{D}'_1\rangle) , \tag{3.226a}$$

$$|a'\rangle_1 \longmapsto \frac{e^{i\phi_1}}{\sqrt{2}}(|\mathrm{D}_1\rangle - |\mathrm{D}'_1\rangle) , \tag{3.226b}$$

where $|\mathrm{D}_1\rangle$ and $|\mathrm{D}'_1\rangle$ denote the states of the particle emerging towards detectors D_1 and D'_1, respectively. Particles 2 and 3 are subjected to similar treatment with detectors D_2 and D'_2 for particle 2 and detectors D_3 and D'_3 for particle 3.

Thus, the initial state $|\Psi_{GHSZ}\rangle$ of the three particles evolves at a later time t into

[219] Greenberger et al. (1989, 1990). See also Auletta et al. (2009, Sect. 16.7.2).

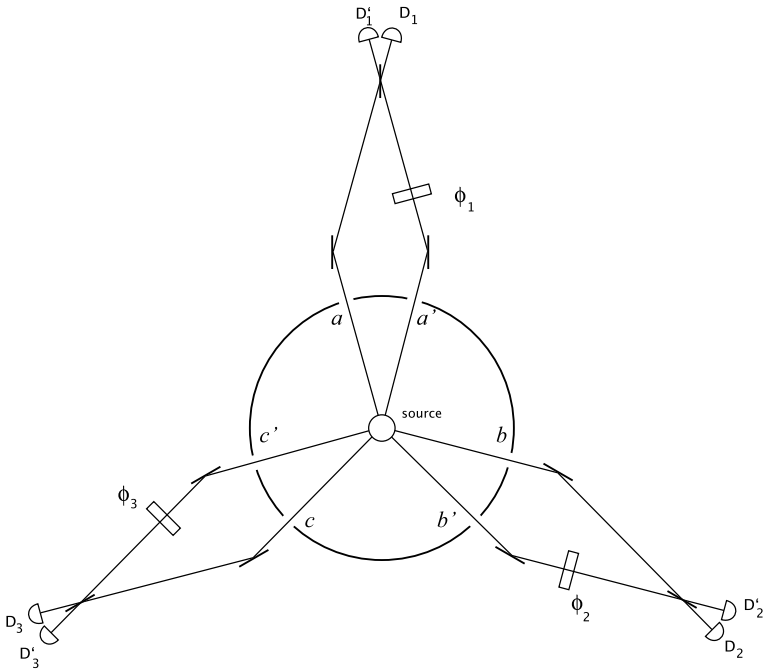


Fig. 3.22 The GHSZ basic experiment with a three-particle interferometer. The source emits a triple of particles, 1, 2 and 3, in six beams, with the state given by Eq. (3.225). A phase shift ϕ_1 is imparted to beam a' of particle 1, and beams a, a' are brought together on a BS before illuminating detectors $\mathrm{D}_1, \mathrm{D}_1'$. Likewise for particles 2 with beams b, b' and for particle 3 with beams c, c'. Adapted from Auletta et al. (2009, p. 616)

$$|\Psi_{GHSZ}(t)\rangle = \frac{1}{4}\Big[\Big(1 + e^{i(\phi_1+\phi_2+\phi_3)}\Big)\big(|\mathrm{D}_1\rangle|\mathrm{D}_2\rangle|\mathrm{D}_3\rangle + |\mathrm{D}_1\rangle|\mathrm{D}_2'\rangle|\mathrm{D}_3'\rangle + |\mathrm{D}_1'\rangle|\mathrm{D}_2\rangle|\mathrm{D}_3'\rangle + |\mathrm{D}_1'\rangle|\mathrm{D}_2'\rangle|\mathrm{D}_3\rangle\big)\Big] + \frac{1}{4}\Big[\Big(1 - e^{i(\phi_1+\phi_2+\phi_3)}\Big)\big(|\mathrm{D}_1\rangle|\mathrm{D}_2\rangle|\mathrm{D}_3'\rangle + |\mathrm{D}_1\rangle|\mathrm{D}_2'\rangle|\mathrm{D}_3\rangle + |\mathrm{D}_1'\rangle|\mathrm{D}_2\rangle|\mathrm{D}_3\rangle + |\mathrm{D}_1'\rangle|\mathrm{D}_2'\rangle|\mathrm{D}_3'\rangle\big)\Big] , \tag{3.227}$$

which results out of a combination of all transformations of the kind (3.226). The probability for detection of the three particles by the respective detectors $\mathrm{D}_1, \mathrm{D}_2$ and D_3 is (see also Eq. (2.33a))

$$\wp_{\mathrm{D}_1\mathrm{D}_2\mathrm{D}_3}(\phi_1, \phi_2, \phi_3) = \frac{1}{16}\left|1 + e^{i(\phi_1+\phi_2+\phi_3)}\right|^2 = \frac{1}{8}[1 + \cos(\phi_1 + \phi_2 + \phi_3)] , \tag{3.228a}$$

where I have made use of Eqs. (2.34). Likewise (see also Eq. (2.33b))

$$\wp_{\mathrm{D}_1'\mathrm{D}_2\mathrm{D}_3}(\phi_1, \phi_2, \phi_3) = \frac{1}{8}[1 - \cos(\phi_1 + \phi_2 + \phi_3)] , \tag{3.228b}$$

and so on for the remaining six possible outcomes. The sum of the probabilities for all eight possible outcomes is of course 1.

Now, we calculate the expectation value on the state $|\Psi_{GHSZ}(t)\rangle$ of the product of the three measurement outcomes at the three arms of the interferometer, given that the relative phases are ϕ_1, ϕ_2, ϕ_3,

$$\begin{aligned}\langle\phi_1, \phi_2, \phi_3\rangle_{\Psi_{GHSZ(t)}} &= \wp_{D_1D_2D_3} + \wp_{D_1D'_2D'_3} + \wp_{D'_1D_2D'_3} + \wp_{D'_1D'_2D_3} \\ &\quad - \wp_{D'_1D'_2D'_3} - \wp_{D'_1D_2D_3} - \wp_{D_1D'_2D_3} - \wp_{D_1D_2D'_3} \\ &= \cos(\phi_1 + \phi_2 + \phi_3) \ , \end{aligned} \tag{3.229}$$

where I have dropped the dependence of probabilities on ϕ_1, ϕ_2, ϕ_3 for simplicity of notation. The above expression represents correlations among probabilities (Sect. 2.1.2), which, as we shall see, need to have the general form of a difference among the probabilities of the "paths".

Given an HV parameter λ that determines the state of the whole multiparticle system, we may define three functions $\alpha_\lambda(\phi_1)$, $\beta_\lambda(\phi_2)$, and $\gamma_\lambda(\phi_3)$ which represent the measurement result at the detector pairs 1, 2 and 3, respectively. There is an implicit assumption (the EPR-like separability requirement) in the introduction of these three functions: $\alpha_\lambda(\phi_1)$ does not depend on ϕ_2 and ϕ_3, and so on, but only on the local (hidden) parameter λ. For the sake of simplicity (and following what Bell did), we assign to each of these functions the value $+1$ when the particle enters an unprimed detector, and -1 when it enters a primed one (Sect. 3.3.5).

Let us show that in fact a contradiction arises between Eq. (3.229) and the EPR-like requirement that three measurement functions depend on a parameter λ and possess definite values once λ, ϕ_1, ϕ_2, and ϕ_3 are specified. For $\phi_1 + \phi_2 + \phi_3 = 0$ the previous expectation is $+1$ while for $\phi_1 + \phi_2 + \phi_3 = \pi$ we get -1. Stated in terms of the three functions $\alpha_\lambda(\phi_1)$, $\beta_\lambda(\phi_2)$, and $\gamma_\lambda(\phi_3)$, we have that

$$\alpha_\lambda(\phi_1)\beta_\lambda(\phi_2)\gamma_\lambda(\phi_3) = \begin{cases} +1 \text{ if } \phi_1 + \phi_2 + \phi_3 = 0 \\ \\ -1 \text{ if } \phi_1 + \phi_2 + \phi_3 = \pi \end{cases} . \tag{3.230}$$

We have a single choice $(0, 0, 0)$ that satisfies the former requirement as well as three different choices of the phase angles that satisfy the latter assignment, i.e. $(\pi/2, \pi/2, 0)$, $(0, \pi/2, \pi/2)$, $(\pi/2, 0, \pi/2)$. For the -1 case, we may write the product of the outcomes as

$$\alpha_\lambda(\pi/2)\beta_\lambda(\pi/2)\gamma_\lambda(0) = -1 \ , \tag{3.231a}$$

$$\alpha_\lambda(0)\beta_\lambda(\pi/2)\gamma_\lambda(\pi/2) = -1 \ , \tag{3.231b}$$

$$\alpha_\lambda(\pi/2)\beta_\lambda(0)\gamma_\lambda(\pi/2) = -1 \ . \tag{3.231c}$$

Multiplying the three Eqs. (3.231) we have

$$\alpha_\lambda(0)\beta_\lambda(0)\gamma_\lambda(0) = -1 \ , \tag{3.232}$$

because the other factors are $\alpha_\lambda^2(\phi) = \beta_\lambda^2(\phi) = \gamma_\lambda^2(\phi) = 1$ for any ϕ. This result clearly contradicts that, for $\phi_1 + \phi_2 + \phi_3 = 0$, we get $+1$ for the expectation (3.229).

Note that the observed count rates for coincidences (for instance, among detectors D_1, D_2, D_3), will depend on the phases. That is, if $\phi_1 + \phi_2 + \phi_3$ is varied linearly in time, then the three-particle coincidence rate (the three-particle interference) will vary sinusoidally (see Eq. (3.228a)). However, there will be no two-particle interference fringes.

Borromean and Hopf Rings

Let us now consider an interesting consequence of the GHSZ state: entanglement, in a given compound system, is strongly affected by measurements performed on one of its constituents.[220] For instance, consider the case of three spin–$\frac{1}{2}$ particles (an instance of the GHSZ state), defined by

$$|\Psi_{GHSZ}\rangle_S = \frac{1}{\sqrt{2}}\left(|\uparrow\rangle_1|\uparrow\rangle_2|\uparrow\rangle_3 + |\downarrow\rangle_1|\downarrow\rangle_2|\downarrow\rangle_3\right) \ , \tag{3.233}$$

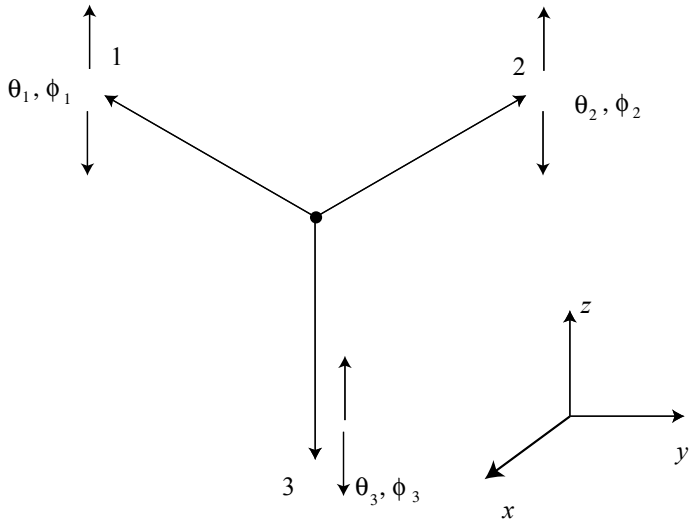


Fig. 3.23 Conceptual scheme allowing investigation of the conditional entanglement arising between two out of three particles in a GHSZ state. Adapted from Auletta et al. (2009, p. 619)

[220] As shown in Krenn and Zeilinger (1996).

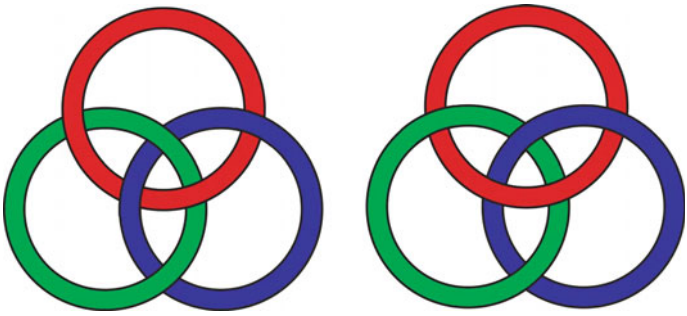


Fig. 3.24 Borromean rings on the left and Hopf rings on the right. Note that the latter are entangled pair by pair while in the former case only one of the rings (e.g. the red one) entangles the other two

where $|\uparrow\rangle_j$ and $|\downarrow\rangle_j$ represent the particle j's up and down states of the spin component along the z-direction, respectively. Now, we look for two-particle entanglement between particles 1 and 2 when the spin of particle 3 is measured along the component defined by the spherical angles (θ_3, ϕ_3) (Fig. 3.23; see also Fig. 1.24, Sect. 1.2.5). If one performs spin measurements within the x–y plane on particles 1 and 2 ($\theta_1 = \theta_2 = \frac{\pi}{2}$), two subensembles are generated depending on the result (up or down) of the measurement performed on particle 3. Then, the two-particle correlation functions for these subensembles are given by

$$\mathrm{Exp}_{12}^{\pm} = \pm \sin(\theta_3) \cos(\phi_1 + \phi_2 + \phi_3). \tag{3.234}$$

Equation (3.234) tells us that, after having measured the third particle, the other two still remain entangled, unless the third is measured along the z-direction ($\theta_3 = 0$), because in this case we obviously obtain: $\mathrm{Exp}_{12}^{+} = \mathrm{Exp}_{12}^{-} = 0$.

Now, if the spin along the z-direction is measured on one particle and the other two are found in a factorised state, this situation can be represented by Borromean rings; if, however, the spin along the z-direction is measured on a particle and the other two remain entangled, this structure that can be represented by Hopf (torus) rings (Fig. 3.24).[221] Of course, we deal here with a geometric representation and we cannot immediately attribute these structures to quantum systems. However, the structural difference is a fact. There is also another aspect to be considered: quantum systems are sensitive to the geometry of their environment. I have already given some hints (Sects. 2.1.2 and 3.3.4), but more will be said in the following. Thus, if we do not like to speak here of structures, the more general notion of *form* (which we have already introduced in Sect. 3.2) makes sense. The term "form" is used not by chance, because I have stressed that information is a pure formal entity (Sects. 3.2.2–3.2.3 and 3.3.7). Since information is finally a correlation among possible outcomes, I shall focus now on the connection between information and quantum correlations.

[221] As proposed in Aravind (1997).

3.4.2 *Correlations and Causal Constraints*

A Causal Power

The main conundrum here is that for most scholars if something does not act mechanically, is not physical. Thus, the realisation that quantum information is not a concrete particular (token), nor some kind of physical stuff, and is not part of the spatial–temporal web of the world (a problem to which I shall come back), leads many scholars to think that it does not exist at all or at least is totally irrelevant to any physical theory. In fact, the crucial point is that no form or structure can activate itself for producing an effect. And this is what makes Timpson's statements reported in Sect. 3.3.7 correct. Nonetheless, although not efficient, I have shown that information be provided with a certain causal power, clearly not of mechanical kind.[222] This is evident in the case of entanglement (a form of mutual information), since it canalises possible events into certain directions and not others (it restricts the space of possible joint events) and in this way allows for operations that do not have a classical analogue (like teleportation or quantum cryptography), as already mentioned. The crucial question is: when and how does it have such a power? Only when there is some kind of measurement or classical communication (what often goes together with some detection), as it is namely the case for teleportation or quantum cryptography. In other words, it is a local *action* (having causal efficiency by itself) that, together with this formal (and global) causal power, determines certain surprising kinds of effects. Otherwise, those resources of information are doomed to remain at most *potential* (Sect. 3.3.6). It might be recalled that the ancient Greek philosopher Aristotle said that something potential can become actual only when there is already something active.[223] In other words, types (which in the traditional philosophical language are called "universals") can have some influence on tokens only thanks to physical processes involving the dynamical contribution of *other tokens*.

This is so because quantum systems, as far as are prepared in a specific state (or undergo spontaneously similar processes), are types–tokens, while free-evolving systems (as representing pure possibles) are types.[224] In fact, as noted, in the latter case the quantum is in a superposition of any possibility, while in the former, without loosing this character, it shows limitations that depend on the physical context (e.g. preparation). The notion of type–tokens is due to the fact that a quantum system in a particular physical context has token-like connotations (e.g. it goes through, or is 'trapped' in, some apparatus) but nonetheless displays also type-like connotations as far as it shows some kind of interference behaviour. Let us consider these concepts a

[222] The reference paper is Auletta et al. (2008).

[223] Aristotle Phys. (1950, I, 7–9).

[224] I think that the notion of *Quasi-gegenstand* (almost object) introduced by Carnap aims at both types and type–tokens, since it concerns general objects but also some individual ones (Carnap 1928, Sect. 27). Armstrong has supported a theory of universals (Armstrong 1978). In Armstrong (1983, pp. 100–101) he introduces the notion of *quasi-universal* that perhaps expresses the same concept of type–token here. See also Margenau (1950, Sect. 15.4).

little further. If a system is prepared in a certain state, it is in that state with probability equal to 1 (while the other elements pertaining to the same orthonormal basis of that state have probability equal to zero). This has led the Dutch physicist G. 'T Hooft to speak of *ontological state* for distinguishing it from an ordinary superposition.[225] However, the Dutch physicist also affirms that "an ontological state evolves into an ontological state; superpositions evolve into superpositions".[226] Apparently, the first part of the statement holds if the system is not disturbed. It is true that if, after a preparation, a system is let free, it may be in any state into which it can evolve according to the Schrödinger or Heisenberg equation, but that statement remains true for stationary states, as explained in Sect. 1.2.3. Of course, often we cannot be certain about the state in which the system is if not submitting it to a further control or preparation procedure.[227] These issues clearly deserve later discussion, but by now we can say that what is a token here is not the state, which always preserves its quantum nature, as remarked in Sect. 3.2.3, and therefore, it is still a type (a superposition of possibilities) but the *event* that has generated a system in a certain state, and therefore *the system being in that state* must be a type–token. In order to avoid any confusion, I shall take here and in the following only *occurring events* as tokens. Note that an idealised classical system considered as a singular system is assumed to be a token, although CM is accustomed to deal with typical models like an elastic collision between two bodies, so that it seems that also this discipline deals with type–tokens. Recapitulating, we have the classification: (i) types (information and correlations among possibilities), (ii) type–tokens (physical systems, i.e. systems that interact and are correlated, through entanglement, with other systems and physical contexts), (iii) tokens (quantum events). Then, the distinction between possibility, potentiality, and actuality can be expresses as the question of whether we deal with types, type–tokens or tokens.

Then, we can explain in a relatively easy way how, in EPR-like experiments, there can be *reciprocal* causal influence of each particle on the other one without any exchange of signals. Such a reciprocal influence is precisely the mark of global constraints (the types) while the efficient-causation side is represented by local operations (dealing with type–tokens), and, thanks to the tokens represented by measurement outcomes (detection of particle 1), correlations as types are able to influence other tokens or type–tokens ((outcome on) particle 2) in the form of causal constraints. I remark that there are very few philosophers that have taken into consideration the effects of types on tokens. I recall here an exception, represented by the British philosopher Harry Collins, who, introducing the terms strings (types) and entities (tokens), asks about the causal influence of the former on the latter.[228]

[225] 'T Hooft (2016, pp. 8–9). Superposition states are called *templates*, which corresponds somehow to the notion of type.

[226] 'T Hooft (2016, p. 44).

[227] The author seems aware of this problem ('T Hooft 2016, pp. 51–54).

[228] Collins (2010).

Classical Examples

Is this completely new and specific to QM? Apparently not: there are many classical examples that have essentially the same character (mesoscopic and macroscopic systems are even richer in structural characters than quantum systems). Let us mention two of them. A hydraulic network has a structure and is inert as such. However, when there is a dynamical (efficient) agency like the water pumped in, this structure is crucial for the further dynamical development (water is canalised along specific directions), and thus comes to exert causal constraints on the process.[229] So, we cannot say that geometry (or structural factors) is (are) irrelevant. The quick motion of gas molecules if confined in the walls of a piston (the form or structure, here) can generate momentum along a certain direction (and therefore also exert pressure). Unfortunately, some followers of the great Austrian physicist and philosopher L. Boltzmann, the father of statistical mechanics, were not completely aware that without such walls no macroscopic effect whatsoever could arise.[230] With the authoritative words of Max Born, "thermodynamic is definitively connected with walls and enclosures".[231] What do these examples tell us? When there is a dynamical factor, this can make operative or *activate* a given structure, according to Aristotle's point of view.

A remark seems appropriate here. The reader may be surprised by the use of classical and even everyday examples for discussing quantum-mechanical issues. In fact, as I shall show, many characters of quanta are absolutely general and apply to many situations, so that similar examples, when taken in the restricted sense in which they are used, can be very helpful to solve such conceptual riddles.

A Historical Remark

The problem raised has been analysed and discussed in philosophy since a long time. Aristotle called this kind of causal factors or powers *formal causes*,[232] although I like to call them (for remaining more faithful to current scientific terminology) causal *constraints*.[233] Thus, in order to avoid any confusion on this point, I use the term *cause* for indicating mechanical or efficient causation, as it is current use, and *constraint* (or at most causal factor or power) when speaking of what Aristotle denoted with the term "formal causes".

---

[229] Auletta et al. (2008).

[230] Boltzmann (1896, 1898). On this subject see Auletta (2011a, Chap. 25).

[231] Born (1949, p. 44).

[232] Aristotle Phys. (1950, II, 3). For deepening these issues see Anagnostopoulos (2009).

[233] It could be said that this insight was originally due to the Italian physicist Franco Selleri (1936–2013) when he affirmed that quantum waves can have physical effects although deprived of momentum and energy, i.e. of dynamical characters, what hints implicitly to a kind of causal constraint (Selleri 1969). Unfortunately, in the following Selleri was not always consequent with this position and searched for an empty wave as a kind of localised object (Sect. 3.3.4), loosing in this way the notion of causal constraint that was implicit in his former point of view.

I have insisted on the distinction between the notion of "cause" and "constraint" not by chance. In fact, in the philosophical schools subsequent to Aristotle, especially late Middle-Age Neoplatonism, formal causes were understood in dynamical terms and considered efficient by themselves.[234] This way to consider the problem fully discredited Aristotle's doctrine and led to the rejection of formal causes in Modern ages. There was likely a second source of trouble due to the identification between formal and final causes. The reason is that Aristotle's philosophy was essentially focused on the biological world and in particular he can be considered the father of epigenetics.[235] Now, in developmental studies it is quite clear that the 'form' of the organism (and therefore the formal cause here) and the final stage that it reaches in its maturity (a process that for Aristotle was ruled by final causation since it leads unavoidably to a species-specific final constitution of the organism) are strictly intertwined factors. This is not the place for discussing about the possible applicability of these notions to modern biology, but what is certain is the inapplicability of the notion of final causes to the physical world.

A Epistemological Remark

Notwithstanding these distinctions, the reader may consider with a certain suspicion that I am aiming at re-introducing notions like causal constraints and potentiality (on which we shall have something to say also in the following) that appear to have been rejected since centuries by modern science. So, it seems that the risk of anachronism is quite strong here. Nevertheless, I am deeply convinced that solutions that have been presented in the past (both in science and philosophy but likely more for the latter), if carefully thought and made free from internal inconsistencies, might always be useful for further developments in knowledge that were fully unpredictable in previous times. The whole worry of both scientific and philosophical ideas or solutions appears to me to be not about the ideas themselves but concerns the misapplication of these ideas to some classes of problems that transcend the original context in which those ideas arose. For instance, the notion of causal constraints cannot be applied to many traditional problems of CM as long as it acknowledges the separability principle. There is now a specific reason why ideas can always have a new life for new and *appropriate* contexts: it is likely that there is a finite amount of possible solutions to given problems, and this explain why recurrently we bounce again and again against the same ideas. There are plenty of scientific examples of equations or concepts that have been resurrected in new contexts (for instance, we have recently seen the reintroduction in cosmology of the cosmological constant, in the form of dark energy, that its proponent, A. Einstein, considered the biggest error of his life). The same is true for philosophical ideas.[236]

---

[234] See Pasnau (2004), Torcal (2013).

[235] As I have recalled in Auletta (2011a, Sect. 11.1.2; 2011c, Sect. 3.3.7).

[236] In fact, there is currently a flourishing of Aristotelian and neo-Aristotelian studies in philosophy.

Therefore, epistemologically speaking, I suggest to avoid rejection or demolition of an idea as such (provided that it is self-consistent and is not a pure truism), but I recommend to always target the *misapplications* of the notion in our critical scrutiny. If this is not sufficient to justify my attempt at re-introducing the notions of causal constraints and potentiality (indeed such a justification can only be found in the effectiveness of these concepts for solving our problem of the interpretation of QM), I hope to have at least provided sufficient reason for no rejection a priori of the same either.

Causal Constraints and Efficient Causes

The reason why Aristotle introduced the notion of formal cause is that he observed that in nature there are many regularities that cannot be explained through efficient causes only.[237] In fact, efficient or mechanical causation is not able to produce any kind of regularity: efficient causation is robust relative to *its consequences*, that is, an efficient cause produces whatever it is fated to produce (i.e. a *single* effect), so that, variations in the effect are determined by corresponding variations in the causes.[238] This is why I am calling mechanical causation also *efficient* causation (in other words, any mechanical causation is sufficiently efficient in bringing to certain effects), and, as a consequence, CM is deterministic in its own way (Sect. 2.4.1). For instance, by slightly changing the direction of a killer's gun, the bullet will go elsewhere and, hoping, will murder nobody. In such a way, if we admit variation at the input or in the initial conditions (as our experience testifies) but assume that there are no additional factors acting as constraints, no regularity would arise in our universe, which in fact we experience everyday and everywhere.

At the opposite, causal constraints are defined by Aristotle and C. Peirce as robust relative to the *initial conditions*. In other words, several variations in the antecedent conditions (in a certain tolerance window) will not affect the result that a causal constraint will contribute to produce. This is well-known in every natural phenomenon in which there is an attractor or a network.[239] In fact, a network can work and produce similar results even if several nodes are impaired (provided that no hub crushes). This property is called *degeneracy*, and it is an universal character of biological systems.[240] This is precisely what does not happen with a mechanical engine like a car's motor (considered as a token): the absence of a part will in most cases impair the performance of the whole. Now, a network (and also an attractor) is also a formal entity (a type) since it consists in certain (probabilistic) interdependencies among its nodes. For this reason, we should not mix its *activity*, which is displayed when the

[237] As pointed out in Auletta (2008, 2013), Stanzione (2013).

[238] As remarked in Peirce (1902). Unfortunately, here and elsewhere Peirce speaks of final causes and not of formal ones.

[239] On network theory I recommend (Barabási 2002; Barrat et al. 2008).

[240] On the subject see Edelman and Gally (2001). See also Auletta (2011a, Sect. 8.2.5) and literature quoted there.

network is at work (when both efficient causes and causal constraints are in play), and the *network* itself. In fact, these interdependencies could be considered each time as the *result* of this activity (through its dynamics the network updates the connections, especially in the case of informational networks like Internet, representing in this way an example of information processing). However, we should not understand the potency (of the network here) as a diminished form of ontology. The properties or characters of the network or its structure are as real as any other thing of the world. It is only its *activity* that cannot start without the contribution of a dynamical cause or factor acting on some parameters or characters of the structure.[241] As said, dynamical or efficient causes can activate the potential resources of causal constraints that otherwise are dormant. Of course, we fall back again into the problem of the kind of reality of quantum correlations (Sect. 2.1.2): we could in fact accept that a classical network is real, but how can we deal with entanglement? Let us examine this problem.

No Single Effect

Of course, in introducing causal constraints as a physical tool, there is also a price to pay: when we say that causal constraints are not efficient or dynamical causes, this means that they cannot target a specific effect. As said, this is at the opposite the quintessence of mechanical or efficient causation and is in fact the basis of our whole technology so far. In other words, causal constraints cannot produce a specific, single effect but rather can give rise to a number of effects in certain windows. This is precisely what happens with entanglement that restricts the space of the possible events but does not tell us whether we shall get, in the case of a singlet state, up-down or down-up (I recall that there is no possibility of prediction without a detection already occurred). Any correlation has such a character. If it were not so, we would be able to manipulate tokens (i.e. local events) or type–tokens by means of global types; but this contradicts the results of the previous section (see what has been said about ebits). This circumstance, however, is precisely what makes causal constraints robust relative to their initial conditions. This seems contradictory. However, as anticipated, here we no longer deal with single causal effects but with *equivalence classes* of effects (Sect. 3.3.7), i.e. with some general characters or aspects that all those possible effects *share*. For instance, any of the Bell states is good for teleportation, or any binary codification is good for quantum cryptography. In this way, formal constraints are types that thanks to type–tokens and tokens can not only have causal influence on other tokens but also influence other, and even give rise to new, types, as it is evident for entanglement swapping.

To remain in the framework of one of classical or everyday examples, note that, when we arrange a forest in order to build a fire barrier, we are not concerned with the specific dynamical sequence that can block the fire, nor with the details of the materials (for instance, differences between individual trees) involved, apart

[241] I have stressed this in Auletta (2011c, Sect. 3.2.5). See also Auletta and Torcal (2011).

from what is really pertinent to this problem's solution. There are many different situations in which our barrier will work well and this is the only issue about which we are concerned, that is, that the barrier will work well in most situations (again Aristotle original insight!). However, when we build an engine like a motor of a car, we are interested in ensuring that all components work precisely in a single way, with a single timing and a single effect, i.e. mechanically. Of course, if we enlarge our vision and consider the possible drivers of such a car (the potential buyers), we are interested that such a car will work in the mean so and so, and in this way will satisfy the needs or wishes of the group of people which could become customers. Again the individual differences among those potential customers are not the business of marketing. Thus, what can be considered equivalence class is *context-dependent*. But this is precisely why we are speaking of potential resources: they need *operative physical contexts* (in terms of both physical conditions and processes) to be activated (the same is obviously true for attractors).

Utility of the Notion of Causal Constraint

Thus, I have provided evidence that both information and constraints can have a causal role. Although the latter notion appears more general than the former one, the subsequent examination will show that is rather the other way around. Moreover, let us remark that, understood in this way, activated causal constraints satisfy Leibniz's principle of the closure of physical (efficient) causes (Sect. 2.3.3) because do not interfere and cannot interfere with dynamical causes (like exchanges of signals). Moreover, it should be noted that, while efficient or mechanical causality is subjected to Einstein's principle of the local closeness of effects or *Nahewirkung* (Sect. 2.3.2), it is not necessary so for causal constraints, what may help to reconcile relativity with QM. Thus, an enlarged (although corrected) concept of causation that takes into account causal constraints may be very interesting in order to solve the problem of causality raised in Sect. 2.4. In fact, having acknowledged that in most cases QM represents a major break down of mechanical causation, the conclusion may be drawn that quantum processes are deprived of causality if not of any nomological character. The quantum of action may be thought to introduce a full anarchy (and even irrationality) in physics because the continuity of processes is broken. Causal constraints might, at the opposite, set additional physical conditions that allow us to overcome this view, and so helping us to build a bridge between quantum random events and the emerging of classical physics. In particular, we have a deterministic (but probabilistic) evolution (and *not* a stochastic one) of quantum systems only because of these causal constraints that are present in any quantum state allowing its coherence. This establishes a connection between laws and causality also for QM (although we do not identify these notions). Moreover, the fact that the evolution is probabilistic is fully consonant with the notion of equivalence classes of effects. In other words, the introduction of causal constraints can be considered a natural consequence of the dissociation of the notion of causality from determinism, still characterising CM.[242]

[242]On this see D'Ariano et al. (2017, Sect. 5.1).

Born's Contribution

Allow me in conclusion to recall that the great physicist Max Born defended unambiguously the point of view supported here. In fact, he distinguishes between two kinds of causal relations: one temporal (and spatially localised) the other atemporal.[243] Both express *dependence* (what justifies in his view the term causality for both cases), but not of mathematical type (characterised, e.g. by the mathematical notion of function). The temporal, mechanical or efficient causality relies on the two following principles: antecedence (the cause must precede the effect) and contiguity (Einstein's notion of *Nahewirkung*). However, according to Born, the timeless meaning of causality is the fundamental one. In fact, in any causal inference of the temporal type, we implicitly refer to a timeless causality individuating an equivalence class of possible cases. If we would drop such a reference to what I have called causal constraints, we would get at most some occasional and short-living *regularity* but not causality. And for this reason one cannot identify causality with classical determinism either, since with the latter there is no dependence of classes of events on other classes of events (in fact, it concerns single events). Thus, the temporal (current) use of causality is correct only when it is supplemented by, e.g. some time and space specification. Born noted that timeless causality or causal constraints go well together with the reversibility of physical laws (that in Sect. 2.4.2 I have shown to consist in correlations among observables), and this is the reason why, according to Born, QM preservers the timeless understanding of causality but rejects classical determinism.[244]

3.4.3 Potentiality

Again a Historical Remark

I have already introduced the general notion of potentiality or of potency, when referred to the resources of a system (see the previous subsection but also Sect. 3.2.3 in relation to information and Sect. 3.3.2 in relation to observables). The modern rejection of the Aristotelian formal causes brought with itself the rejection of the notion of *potentiality*. There are specific reasons that led to this choice. Classical-mechanical systems are mechanically deterministic, so that by looking the present state of any system, i.e. all of its properties, what is strictly connected with the notion of *omnimoda determinatio* (Sects. 1.2.1, 1.2.4, and 2.1.2), and knowing the conditions at contour we are capable to both fully retrodict and predict its past and future states, respectively (Sect. 2.4). And this is why these systems are considered as pure token, and S. Laplace could consider in principle possible to know the future (and the past) of

[243] Born (1949, Chap. 2).

[244] Born (1949, Chaps. 3–4, 9).

the whole universe.[245] In other words, the knowledge of the current state of classical system fully exhausts the knowledge of its future (and past) evolution. On this basis, the notion of potentiality is not meaningful. This is why all modern scholars, as, e.g. 18th century German philosophers like Alexander Baumgarten (1714–1762) and Immanuel Kant,[246] took for granted that there is only actual reality (instantiation of properties) and all that is not real in this sense can only be considered at most possible (if not impossible), that is, properties that, theoretically speaking, we could (counterfactually) ascribe to a system if conditions were different.

Potentiality Is Not Possibility

However, as clarified in Sect. 3.2.3, the notion of possibility and potentiality are very different. A simple example from current life can help here. For an Italian lemon seed, it is possible to be brought to Great Britain but it is potentially a lemon tree. While the first thing has no relation with its current conditions (or a very weak one) whether internal or external, the second is a consequence of its biological constitution: indeed, it cannot become (at least, spontaneously) an apple tree or an ant while it can always be brought to any alternative location. In other words, we can speak of potentiality only when there are some *real resources* in the constitution of a certain system such that its further evolution can be considered to be a consequence of such a constitution,[247] as it is evident in the case of entanglement (Sect. 3.3.6). This will not say that it is a consequence of this constitution alone. Also in the case of a seed, additional (and not less real) environmental factors are necessary, like a certain light intensity or temperature.[248] In other words, the difference between possibility and potentiality is a question of whether there are *particular* physical conditions (tokens or type–tokens) or not. This is why I have said that we take what is possible as what would happen in other conditions or at least on conditions that are not present. On the other hand, I wish also to recall that even for many other classical situations it is impossible to deal with systems without assuming some kind of dispositional properties, as it was clear to the epistemologists Ernst Nagel (1901–1985) and Carl Hempel (1905–1997) (Sect. 3.3.2).

Now, when we focus on quantum systems, we find a situation that is to a certain degree quite similar to the example of the lemon seed. Also these systems, in fact, do not possess properties that allow us to say with certainty what will be the future development at the individual scale, at the opposite of what occurs for classical-mechanical systems. In the quantum case, the correlations involved in the superposition state or entanglement will contribute to determine the further evolution of the system, and so have a causal power. However, the information present in this way in the initial

[245] Laplace (1825).

[246] Baumgarten (1739), Kant (1763). For an examination of the problem see also Auletta (2004a), Auletta (2006a).

[247] Aristotle (1950, 201a 10–19).

[248] See Auletta (2011a, Sects. 8.2.1–8.2.4).

state of the system consists in correlations among an array of possible outcomes. This means, in agreement with the previous subsection, that this initial amount of potential information is insufficient by itself to produce a particular outcome but needs additional dynamical factors in order to give rise to particular results.

Not Very Popular But …

Of course, I am aware that the notion of potentiality is not very popular among modern scholars, and this for the mentioned reason. There are few exceptions. I recall here some of them. A scholar who considered the relation between potentiality and actuality was the mathematician and philosopher N. Whitehead, who said that actuality is a "decision" amid potentiality.[249] As a further reference, we could consider the notion of potentiality as derived from Bell's concept of *beable*,[250] although he seems to have used the notion in a different sense than that employed here. However, one of the scholars who has likely more stressed the relevance of this notion for QM was A. Shimony.[251] Indeed, also the concept of passion at a distance (introduced in Sect. 3.3.6) is consonant with the notion of potentiality. More recently, as recalled in Sect. 3.2.3, Zurek has introduced the notion of actionable information.[252]

Heisenberg's Juvenile Standpoint and Operationalist Epistemology

The interpretation that I am proposing here was formulated for the first time by one of the fathers of QM, namely W. Heisenberg: after the interactionist and instrumentalist stage of his youth (Sect. 3.1.1), he recovered the Aristotelian philosophy as a way to deal with the problems that are also central to us. In fact, the flaws of his previous and still immature point of view were two:

(i) *Instrumentalism*: the idea that the theory formulates the correct predictions and this is what does matter and
(ii) *Interactionism*: the idea that the theory can work only because we dynamically interact with quantum systems.

Point (ii) is clearly heritage of the Bohr's stress on observation of detection events (Sects. 2.4.2 and 3.3.2) and has been in fact disproved by the interaction-free measurement (Sect. 2.1.2). Point (i) has likely to have been dismissed by Heisenberg after an interview that he had with Einstein.[253] Let us discuss extensively these two intertwined aspects.

---

[249] Whitehead (1929, pp. 43–46 and 61). See also Epperson and Zafiris (2013, Chap. 4).
[250] Bell (1976, 1981). See also Dieks (1994).
[251] Shimony (1993): see the index of the volumes.
[252] Zurek (2013).
[253] As reported in Heisenberg (1969, Chap. 5).

In fact, instrumentalism and interactionism are strictly connected in a framework that is called "operationalist epistemology". In 1927, the American physicist and epistemologist P. W. Bridgman (1882–1961) introduced the idea that "the concept is synonymous with the corresponding set of operations".[254] In other words, according to this standpoint, theoretical concepts (and inferences) should play no role in physics if empty of operational meaning. Thus, it is not a surprise that he affirms: "whenever operations do not exist the concept is meaningless".[255] There are two interesting consequences of this view:

(a) Bridgman remarked that the operations for, e.g. measuring the length at astronomical scale are different from operations for measuring length at our scale, and from this, he concluded that Einstein's astronomical 'length' does not mean the same as our ordinary 'length'.[256] This consequence appears to me to be absurd, since length is length everywhere, we cannot change this concept according to different operational contexts. No single physical quantity would resist such a treatment, and obviously the correspondence principle would appear wrong (Sect. 1.3.2). However, without that assumption we would have made no progress in quantum physics. Another issue, obviously, is that the scale does matter, making relevant or irrelevant a certain degree of exactitude. But this is another kind of problem.

(b) As a consequence, Bridgman could not account for measures that are quite common in physics but do not (or at least did not) appear to be the result of operationalist procedures. It is striking, in fact, his embarrass in dealing with the problem of the (classical) length of the electron's diameter, since he was unable to determine an operational procedure for measuring it.[257] Obviously, current developments in the theory of quarks, for not speaking of strings, would certainly embarrass him much more.

It is worth mentioning in such a context that Bridgman developed his epistemology by referring to the operational definition of simultaneity provided by Einstein (in a way that was quite similar to the early Heisenberg).[258] However, later on, Einstein took a certain distance from this view, and it is not by chance that in 1949 Bridgman criticised him for having introduced theoretical concepts in its general theory of relativity,[259] and that Einstein replied by observing that it is not necessary to demand that all of the assertions of a physical theory can be independently interpreted and 'tested' operationally.[260]

Coming back to the interview of Heisenberg with Einstein, Heisenberg told him that he would like to build quantum theory without any consideration of the onto-

[254] Bridgman (1927, p. 5).
[255] Bridgman (1927, p. 25, 28).
[256] Bridgman (1927, p. 12).
[257] Bridgman (1927, pp. 22–23).
[258] Heisenberg (1927).
[259] Bridgman (1949).
[260] Einstein (1949b, p. 679).

logical substrate of the entities we deal with but by taking into account only those quantities that are formally described by the theory and we are able to observe. Einstein replied that we cannot build a physical theory with observable quantities only. Heisenberg told him that he adopted Einstein's early interpretation thinking to be faithful to his own methodology. In particular, as mentioned Heisenberg quoted Einstein's dissolution of the concept of simultaneity in relativity theory (Sect. 2.3.2) and took this as a kind of methodological guide for the introduction of the uncertainty relations as suppressing the simultaneous measurability of canonical conjugate observables.[261] Einstein (with the same spirit of his reply to Bridgman) replied that although he may have used this philosophy, it remained nonetheless nonsense (*Unsinn*).

Heisenberg's Later Standpoint

As a matter of fact, in the 1950s Heisenberg proposed an interpretation according to which the components of a quantum state are only potential before detection.[262] In this new context, Heisenberg asked what happens 'in reality' in quantum processes.[263] He affirms that the concept 'event' must be limited to observations and that any attempt at attributing reality to quantum phenomena independently from observation leads to contradictions. In this, he remained a true follower of Bohr, accepting a widely understood phenomenism (Sects. 3.1.1 and 3.2.3). Nevertheless, according to him, the wave function, at least in a measurement context, joints subjective and objective factors,[264] because it contains

1. Assertions on the tendencies of a given system (in the sense of Aristotle's potency), and these are completely objective and independent from the observer,[265] on the one hand, and
2. Statements about our knowledge of the system, on the other hand,

i.e. both ontological and epistemic aspects need to be considered, as already pointed out in Sect. 3.3.2. Now, Heisenberg's new point of view is that, when measuring, the interaction of the system with the whole universe is also involved in the dynamical process, but we *cannot control* these effects[266]:

> Therefore, the transition from the "possible" to the "actual" takes place during the act of observation. If we want to describe what happens in an atomic event, we have to realise that the word "happens" can apply only to the observation, not to the state of affairs between two

[261] Heisenberg (1927).

[262] Heisenberg (1958, Chap. 3). In my previous work I was too negative on this point although the distinction between possibility and potentiality was already pointed out (Auletta 2000, Sect. 29.2.2).

[263] See also Auletta and Tarozzi (2004, pp. 1680–81).

[264] It seems that W. Pauli somehow supported a view in this sense (Home and Whitaker 2007, p. 63).

[265] Shimony, who is supportive of the notion of potentiality, reminds us that H. Margenau had already used (in 1949) the notion of "latency" (Margenau 1950, Sect. 8.2; Shimony 1981).

[266] Heisenberg (1958, pp. 54–55; see also pp. 137–38).

> observations. It applies to the physical, not the psychical act of observation, and we may say that the transition from the "possible" to the "actual" takes place as soon as the interaction of the object with the measuring device, and thereby with the rest of the world, has come into play; it is not connected with the act of registration of the result by the mind of the observer. The discontinuous change in the probability function, however, takes place with the act of registration, because it is the discontinuous change of our knowledge in the instant of registration that has its image in the discontinuous change of the probability function.

The concept of *possible* is meant here to denote the possible outcomes of a measurement.[267] We may measure the distance of this approach relative to Wigner's interpretation (with which Heisenberg could previously share some aspects) that fully relies on the "psychical act of observation" and with a local registration of information that is dependent on the mind of the observer (Sect. 3.1.1). In another place Heisenberg adds[268]:

> The probability function, which covered a wide range of possibilities, is suddenly reduced to a much narrower range by the fact that the experiment has led to a definite result, that actually a certain event has happened. In the formalism this reduction requires that the so-called interference of probabilities, which is the most characteristic phenomenon of quantum theory, is destroyed by the partly indefinable and irreversible interactions of the system with the measuring apparatus and the rest of the world.

Note that here the interference of probabilities is considered "the most characteristic phenomenon of quantum theory", and its washing out as crucial for measurement (Sect. 2.1.2). Thus, Heisenberg was the first scholar to hypothesise that the environment (and a kind of selection connected with it) could be responsible of the so-called reduction of the wave packet (Sect. 3.1.3). Formulated in more general terms, we can say that irreversibility both in thermodynamics and QM is due to a renunciation of the demand that in principle the fate of every single particle can be determined (but depend on the correlations and interactions with a huge amount of other, uncontrollable, factors).[269] The convergences seen here and in the previous subsection between Heisenberg and Born are not by chance, and we can speak here of a common operational–ontological point of view.

Conclusions

Summarising, Heisenberg could use the Aristotelian concept of potency when speaking of the initial state of the system and affirm that it is the dynamical process during the measurement in interconnection with the environment that brings to actuality one of these components. It may be also noted that the environmental conditions themselves are only potential: indeed the only activity comes from the local interaction dynamics between apparatus and system or between system and detector. In other words, the complex of physical conditions, from the environment to the experimental set up, together with quantum correlations as a frame involved in those conditions, represent the potency out of which, and in presence of specific dynamical processes,

[267] See also Van Fraassen (1991).

[268] Heisenberg (1958, p. 142).

[269] As pointed out in Born (1949, Chaps. 6–7).

actual events might occur. I hope to have succeeded in purifying the notion of potentiality from any reduction to the epistemic version of the Copenhagen interpretation, as feared by Deutsch.[270] Let us call *operationalism* in short the standpoint of the later Heisenberg as it is purified from the previous instrumentalism.[271] A different issue is whether he also abandoned interactionism. Likely not, as his reply to M. Renninger about the interaction-free measurement shows,[272] as also recalled in Sect. 3.1.1.

3.5 Summary

- If knowledge is the factor that makes the difference during the measurement process, the solution must be subjective. However, this does not imply solipsism, but raises the problem of our agreement with reality.
- At the opposite, the problems of objectivism and the MWI are basis degeneracy and the lack of distinction between possible and actual. Nevertheless, the MWI opens a new understanding of quantum systems as replicators of the superposition of all possibilities. A coherent subjectivist interpretation of measurement seems to clash with the fact that our physical theories agree with reality as far as our experience can testify.
- An alternative explanation has been provided by decoherence. It raises the problem of a new ontology centred on the dichotomy local/global: here, interference terms are not washed out in absolute terms but only locally. The necessity arises to find an adequate ontological substrate.
- The solution to the Schrödinger cat's paradox is the high instability of the cat's state. This suggests a criticism of Bohr's classicism about the complementarity principle. Laws do not determine local events and thus the world is not built by laws only.
- Formalist interpretations are recurrent. This suggests that we can understand measurement as dealing with information but in objective terms. The adimensionality of information is not a problem but rather testifies its universal applicability.
- Information in a quantum context, in a first instance, can be defined as connection among events not yet happened. The possible events themselves make the difference. In the state of a quantum system that has not been subjected to local procedures there is an infinite amount of potential information.
- Entanglement, a form of mutual information, is causally influent since it reduces the space of events.
- Quantum systems display self-similarity, but information becomes potentially accessible when there are specific and local physical conditions.

---

[270] Deutsch (2011, Chap. 12).

[271] An operationalist standpoint in this sense has been supported for the first time in Poincaré (1905, Chap. 3). In particular, he tells us that "to localise an object means to represent the movements needed for reaching it" (Poincaré 1905, p. 67).

[272] As reported in Renninger (1960).

- The triplet possibility–potentiality–actuality corresponds to the triplet free particle, reversible steps of preparation–premeasurement, irreversible detection. To acquire information is to combine local selection events and information sharing.
- Observables are real, in agreement with EPR, since also classically we deal in many cases with dispositional properties, but we are accepting also Bohr's point of view that conditions set by ourselves are fundamental. The crucial aspect is that the context needs to be both epistemic and ontological.
- There is mutual influence among far away systems without exchange of signals. Schrödinger proved this with a conditional attribution of properties, but this also shows that he never dealt with real detection.
- HV theories were born with the aim to complete QM. If Bohm's quantum potential is not understood as an ordinary physical potential, it can be integrated in this interpretation with some caution.
- The Kochen–Specker theorem (on the line of Bohr's argument) shows that attributing properties to quantum systems without consideration of the physical context and detection events leads to conflicts with classical ascription of properties.
- Bell inequalities allow a quantification of the EPR thought experiment. The results of the experimental tests show perfect agreement with the predictions of QM.
- Entanglement swapping shows that no local interaction is necessary for information sharing, while quantum teleportation suggests that entanglement is reservoir of potential information.
- Quantum cryptography shows that codes cannot be manipulated by quantum-mechanical means.
- There is a connection between Tsirelson bound for QM and the principle of information causality, according to which nobody can acquire more information than that which has been classically transmitted. This shows that entanglement is sharing of outcomes but not of settings or code alphabets.
- To entangle two systems amounts to couple code alphabets, as displayed by quantum cryptography.
- To avoid a renunciation of the causality notion, we need to enlarge the notion of causality by introducing causal constraints like entanglement.
- The notion of potentiality is very relevant for dealing with physical contexts or the notion of observables. With this term, it is meant, e.g. what a network can do when solicited. By adopting this notion, the late Heisenberg overcomes instrumentalism and likely only in part interactionism that characterised his previous work.

References

Abramsly, S., and B. Coecke. 2009. Categorial Quantum Mechanics. http://www.indiana.edu/iulg/qliqc/HQL_fin.pdf

Adler, S.L. 2004. *Quantum Theory as an Emergent Phenomenon: The Statistical Mechanics of Matrix Models as the Precursor of Quantum Field Theory*. Cambridge: Cambridge University Press.

Aharonov, Y. and D. Bohm. 1959. Significance of Electromagnetic Potentials in the Quantum Theory. *Physical Review* 115: 485–91; Report in [Shapere/Wilczek 1989, 104–110].
Anagnostopoulos, G. (ed.). 2009. *A Companion to Aristotle*. Malden: Wiley-Blackwell.
Araki, H., and E.H. Lieb. 1970. Entropy Inequalities. *Communications in Mathematical Physics* 18: 160–70.
Aravind, P.K. 1997. Borromean Entanglement of the GHZ State. in [Cohen et al. 1997, 53–59].
Aristotle. 1950. *Physica*. Oxford: Clarendon; Also in [Aristotle Op., 184–267].
Armstrong, D.M. 1978. *A Theory of Universals*. Cambridge: Cambridge University Press.
Armstrong, D.M. 1983. *What is a Law of Nature*. Cambridge: Cambridge University Press.
Aspect, A., J. Dalibard, and G. Roger. 1982. Experimental Tests of Bell's Inequalities. *Physical Review Letters* 49: 1804–807.
Auletta, G., and S.-Y. Wang. 2014. *Quantum Mechanics for Thinkers*. Singapore: Pan Stanford Publication. https://doi.org/10.4032/9789814411721.
Auletta, G., G. Ellis and L. Jaeger. 2008. Top-Down Causation by Information Control: From a Philosophical Problem to a Scientific Research Program. *Journal of the Royal Society: Interface* 5: 1159–72. https://doi.org/10.1098/rsif.2008.0018
Auletta, G., M. Fortunato, and G. Parisi. 2009. *Quantum Mechanics*. Cambridge: Cambridge University Press; Revised paperback ed. 2013.
Auletta, G. 1994. *Determinismo e Contingenza*. Napoli, Morano: Saggio sulla filosofia leibniziana delle modalità.
Auletta, G. 2000. *Foundations and Interpretation of Quantum Mechanics: In the Light of a Critical-Historical Analysis of the Problems and of a Synthesis of the Results*. Singapore: World Scientific; Revised ed. 2001.
Auletta, G. 2004a. Critical Examination of the Conceptual Foundations of Classical Mechanics in the Light of Quantum Physics. *Epistemologia* 27: 55–82.
Auletta, S. 2004b. Interpretazioni Soggettivistiche Della Misurazione in Meccanica Quantistica. *Teorie e modelli* 9: 43–61.
Auletta, G. 2006a. The Ontology Suggested by Quantum Mechanics. In *Topics on General and Formal Ontology*, ed. P. Valore, 161–79. Monza: Polimetrica International Scientific Publisher.
Auletta, G. 2006b. The Problem of Information. in [Auletta 2006c, 109–127].
Auletta, G. (ed.). 2006c. *The Controversial Relations Between Science and Philosophy: A Critical Assessment*. Vatican City: Libreria Editrice Vaticana.
Auletta, G. 2007. Il dibattito tra Einstein, Bohr e Schrödinger: Causalità, determinazione e azione a distanza. *Il Protagora* 9: 181–98.
Auletta, G. 2008. How Many Causes Are There?21mo. secolo. *Scienza e tecnologia* 5: 41–48.
Auletta, G. 2009. Quantum Mechanics and Contingency. *Filozofia Przyrody i Nauk Przyrodniczych* 1: 7–16.
Auletta, G. 2011a. *Cognitive Biology: Dealing with Information from Bacteria to Minds*. Oxford: Oxford University Press.
Auletta, G. 2011b. Correlations and Hyper-Correlations. *Journal of Modern Physics* 2: 958–61. https://doi.org/10.4236/jmp.2011.29114, http://www.scirp.org/journal/PaperInformation.aspx?paperID=7135.
Auletta, G. 2011c. (in collaboration with I. Colagè, P. D'ambrosio, and L. Torcal). *Integrated Cognitive Strategies in a Changing World*. Rome: G and B Press.
Auletta, G. 2013. La sfida del darwinismo e la causalità nel mondo naturale. in [Auletta/Pons 2013, 123–37].
Auletta, G., and S. Pons Domenech (eds.). 2013. *Si può parlare oggi di una finalità dell'evoluzione? Riessioni loso che e teologiche alla luce della scienza contemporanea*. Rome: G and B Press.
Auletta, G., and G. Tarozzi. 2004. On the Physical Reality of Quantum Waves. *Foundations of Physics* 34: 1675–94.
Auletta, G. and L. Torcal. 2011. From Wave–Particle to Features–Event Complementarity. *International Journal of Theoretical Physics* 50: 3654–68. https://doi.org/10.1007/s10773-011-0833-8

Ballentine, L.E. 1970. The Statistical Interpretation of Quantum Mechanics. *Review of Modern Physics* 42: 358–81.
Barabási, A.-L. 2002. *Linked: How Everything is Connected to Everything Else and What It Means for Business, Science, and Everyday Life*. New York: Penguin.
Barnett, S.M., and S.J.D. Phoenix. 1989. Entropy as a Measure of Quantum Optical Correlation. *Physical Review* A40: 2404–409.
Barnett, S.M., and S.J.D. Phoenix. 1991. Information Theory, Squeezing, and State Correlations. *Physical Review* A44: 535–45.
Barrat, A., M. Barthélemy, and A. Vespignani. 2008. *Dynamical Processes on Complex Networks*. Cambridge: Cambridge University Press.
Barrow, J.D., and F.J. Tipler. 1986. *The Anthropic Cosmological Principle*. Oxford: Clarendon P.
Battail, G. 2014. *Information and Life*. Berlin: Springer.
Baumgarten, A.G. 1739. *Metaphysica*, 7th ed. Halle 1779; rep. Hildesheim–New York, Georg Olms, 1982.
Bekenstein, J.D. 2003. Information in the Holographic Universe. *Scientific American* 2892: 58–65. https://doi.org/10.1038/scientificamerican0803-58
Bell, J.S. 1964. On Einstein Podolsky Rosen Paradox. *Physics* 1: 195–200; Report in [Bell 1987, 14–21].
Bell, J.S. 1987. *Speakable and Unspeakable in Quantum Mechanics*. Cambridge: Cambridge University Press.
Bell, J.S. 1966. On the Problem of Hidden Variables in Quantum Mechanics. *Review of Modern Physics* 38: 447–52; Report in [Bell 1987, 1–13].
Bell, J.S. 1971. Introduction to the Hidden-Variable Question. In [D'Espagnat 1971, 171–81]; Report in [Bell 1987, 29–39].
Bell, J.S. 1976. The Theory of Local Beables. *Epistemological Letters*; Report in [Bell 1987, 52–62].
Bell, J.S. 1981. Quantum Mechanics for Cosmologists. In *Quantum Gravity 2*, ed. C.J. Isham, R. Penrose, and D. Sciama. Oxford: Clarendon; Report in [Bell 1987, 117–38].
Bell, J.S. 1990. Against Measurement. *Physics World* 3 (8): 33–40.
Bennett, C.H., and G. Brassard. 1984. Quantum Cryptography: Public Key Distribution and Coin Tossing. *Proceedings IEEE International Conference on Computer Systems*, 175–79. New York: IEEE.
Bennett, C.H., and S.J. Wiesner. 1992. Communication via One- and Two-Particle Operators on EPR States. *Physical Review Letters* 69: 2881–84.
Bennett, C.H., G. Brassard, C. Crepeau, R. Jozsa, A. Peres, and W.K. Wootters. 1993. Teleporting an Unknown Quantum State via Dual Classical and EPR Channels. *Physical Review Letters* 70: 1895–99.
Bennett, C.H. 1973. Logical Reversibility of Computation. *IBM Journal of Research and Development* 17: 525–32.
Berkeley, G. 1710. *A Treatise Concerning the Principles of Human Knowledge*. Dublin, 1710: London, 1734; Penguin 1988.
Berry, M.V. 1984. Quantal Phase Factors Accompanying Adiabatic Changes. *Proceedings of the Royal Society of London* A392: 45–57; Report in [Shapere/Wilczek 1989, 124–36].
Berry, M.V. 1987. The Adiabatic Phase and the Pancharatnam's Phase for Polarized Light. *Journal of Modern Optics* 34: 1401–407; Report in [Shapere/Wilczek 1989, 67–73].
Bialynicki-Birula, I., and J. Mycielski. 1976. Nonlinear Wave Mechanics. *Annals of Physics* 100: 62–93.
Bird, A. 2007. *Nature's Metaphysics: Laws and Properties*. Oxford: Oxford University Press.
Bohm, D. 1952. A Suggested Interpretation of the Quantum Theory in Terms of 'Hidden Variables. *Physical Review* 85: 166–93; Report in [Wheeler/Zurek 1983, 369–96].
Bohm, D. 1980. *Wholeness and the Implicate Order*. London: Routledge.
Bohm, D., and B.J. Hiley. 1993. *The Undivided Universe: An Ontological Interpretation of Quantum Theory*. London: Routledge.

Bohr, N. 1928. The Quantum Postulate and the Recent Development of Atomic Theory. *Nature* 121: 580–90; Report in [Bohr CW, VI, 148–58].
Bohr, N. 1935a. Quantum Mechanics and Physical Reality. *Nature* 136: 65; Report in [Bohr CW, VII, 290].
Bohr, N. 1935b. Can Quantum-Mechanical Description of Physical Reality Be Considered Complete? *Physical Review* 48: 696–702; Report in [Bohr CW, VII, 292–96].
Bohr, N. 1948. On the Notions of Causality and Complementarity. *Dialectica* 1: 312–19; Report in [Bohr CW, VII, 325–37].
Bohr, N. 1949. Discussion With Einstein on Epistemological Problems in Atomic Physics. in [Schilpp 1949, 201–241].
Boltzmann, L. 1896. *Vorlesungen über Gastheorie I*. Leipzig: J. A. Barth.
Boltzmann, L. 1898. *Vorlesungen über Gastheorie II*. Leipzig: J. A. Barth.
Born, M. 1949. *Natural Philosophy of Chance and Cause*. Oxford: Clarendon.
Bouwmeester, D., J.-W. Pan, K. Mattle, M. Eibl, H. Weinfurter, and A. Zeilinger. 1997. Experimental Quantum Teleportation. *Nature* 390: 575–79.
Braginsky, V.B., and F.Y. Khalili. 1992. *Quantum Measurement*. Cambridge: Cambridge University Press.
Braunstein, S.L., A. Mann, and M. Revzen. 1992. Maximal Violation of Bell Inequalities for Mixed States. *Physical Review Letters* 68: 3259–61.
Bricmont, J. 2016. *Making Sense of Quantum Mechanics*. Berlin: Springer.
Bridgman, P.W. 1927. *The Logic Of Modern Physics*. New York: McMillan.
Bridgman, P.W. 1949. Einstein's Theories and the Operational Point of View. in [Schilpp 1949, 333–54].
Brillouin, L. 1962. *Science and Information Theory*. New York: Academic.
Brukner, C., and A. Zeilinger. 1999. Operationally Invariant Information in Quantum Measurements. *Physical Review Letters* 83: 3354–57.
Brukner, C., and A. Zeilinger. 2001. Conceptual Inadequacy of the Shannon Information in Quantum Measurements. *Physical Review A* 63: 022113.
Brukner, C. and A. Zeilinger. 2005. Quantum Physics as a Science of Information. in [Elitzur et al. 2005, 47–61].
Brune, M., E. Hagley, J. Dreyer, X. Maître, A. Maali, C. Wunderlich, J.M. Raimond, and S. Haroche. 1996. Observing the Progressive Decoherence of the 'Meter' in a Quantum Measurement. *Physical Review Letters* 77: 4887–90.
Bub, J. 1989. On Bohr's Response to EPR: A Quantum Logical Analysis. *Foundations of Physics* 19: 793–805.
Carnap, R. 1928. *Der Logische Aufbau der Welt*. Meiner: Hamburg.
Cavalcanti, D., M.L. Almeida, V. Scarani and A. Acín. 2011. Quantum Networks Reveal Quantum Nonlocality. *Nature Communications* 2: 184. https://doi.org/10.1038/ncomms1193
Cerf, N.J., and C. Adami. 1997. Negative Entropy and Information in Quantum Mechanics. *Physical Review Letters* 79: 5194–97.
Chalmers, D.J. 1996. *The Conscious Mind. In Search of a Fundamental Theory*. Oxford: Oxford University Press.
Clauser, J.F., M.A. Horne, A. Shimony and R.A. Holt. 1969. Proposed Experiment to Test Local Hidden-Variable Theories. *Physical Review Letters* 23: 880–84; Report in [Wheeler/Zurek 1983, 409–413].
Cohen-Tannoudji, G. 1991. *Les constantes universelles*. Paris: Hachette.
Cohen, R.S., M.A. Horne, and J. Stachel (eds.). 1997. *Potentiality, Entanglement and Passion-at-a-Distance: Quantum Mechanical Studies for Abner Shimony*. Dordrecht: Kluwer.
Collins, H. 2010. *Tacit and Explicit Knowledge*. Chicago: University of Chicago Press.
Conway, J., and S. Kochen. 2006. The Free Will Theorem. *Foundations of Physics* 36: 1441–73. https://doi.org/10.1007/s10701-006-9068-6
Conway, J., and S. Kochen. 2009. The Strong Free Will Theorem. *Notices of the AMS* 56 (2): 226–32.

Croca, J.R. 1987. Neutron Interferometry Can Prove (or Refute) the Existence of de Broglie's Waves. *Foundations of Physics* 17: 971–80.
D'Ariano, M., G. Chiribella, and P. Perinotti. 2017. *Quantum Theory from First Principles: An informational approach*. Cambridge: Cambridge University Press.
de Broglie, L. 1956. *Une tentative d'Interpretation Causale et Non-linéaire de la Mécanique ondulatoire*. Paris: Gauthier-Villars.
de Laplace, P.S. 1796. *Exposition du système du monde*. Paris: Fayard.
de Laplace, P.-S. 1825. *Essai philosophique sur les probabilités*, 5th ed. Cambridge: Cambridge University Press.
De Morgan, A. 1847. *Formal Logic or the Calculus of Inference, Necessary and Probable*. London: Taylor and Walton; Chicago.
de Muynck, W.M. 2002. *Foundations of Quantum Mechanics, an Empiricist Approach*. New York: Kluwer Academic Publishers.
D'Espagnat, Bernard (ed.). 1971. *Foundations of Quantum Mechanics*. New York: Academic.
D'Espagnat, B. (ed.). 1995. *Veiled Reality*. Reading: Addison/Wesley.
Deutsch, D. 1966. Comment on Lockwood. *British Journal for Philosophy of Science* 47: 222–28.
Deutsch, D. 1983. Uncertainty in Quantum Mechanics. *Physical Review Letters* 50: 631–33.
Deutsch, D. 1997. *The Fabric of Reality: the Science of Parallel Universes and Its Implications*. London: Penguin Books.
Deutsch, D. 2010. Apart from Universes. in [Saunders et al. 2010, 542–52].
Deutsch, D. 2011. *The Beginning of Infinity: Explanations that Transform the World*. London: Penguin Books.
DeWitt, B.S. 1970. Quantum Mechanics and Reality. *Physics Today* 23.9: 30–40; Report in [Dewitt/Graham 1973, 155–65].
DeWitt, B.S. 1971. The Many-Universes Interpretation of Quantum Mechanics. in [D'Espagnat 1971]; Report in [Dewitt/Graham 1973, 167–218].
Dieks, D. 1994. Modal Interpretation of Quantum Mechanics, Measurements, and Macroscopic Behavior. *Physical Review* A49: 2290–299.
Eberhard, P.H. 1978. Bell's Theorem and the Different Concepts of Locality. *Nuovo Cimento* 46B: 392–419.
Eddington, A.S. 1934. On the Method of Theoretical Physics. *Philosophy of Science* 1: 163–69; Eng. Tr. of [Einstein 1930].
Einstein, A. 1949a. Autobiographisches. in [Schilpp 1949, 2–94].
Einstein, A. 1949b. Remarks Concerning the Essays Brought Together in This Co-operative Volume. in [Schilpp 1949, 665–88].
Eddington, A.S. 1939. *Philosophy of Physical Science*. Cambridge: Cambridge University Press.
Edelman, G.M., and J.A. Gally. 2001. Degeneracy and Complexity in Biological Systems. *Proceedings of the National Academy of Sciences USA* 98: 13763–68.
Einstein, A., B. Podolsky and N. Rosen. 1935. Can Quantum-Mechanical Description of Physical Reality be Considered Complete? *Physical Review* 47: 777-80; Report in [Wheeler/Zurek 1983, 138–41].
Einstein, A. 1918. Prinzipien der Forschung; Report in [Einstein 1993, 107–113].
Einstein, A. 1930. Zur Methodik der Theoretischen Physik; Report in [Einstein 1993, 113–19].
Elby, A., and J. Bub. 1994. Triorthogonal Uniqueness Theorem and its Relevance to the Interpretation of Quantum Mechanics. *Physical Review* A49: 4213–16.
Englert, B.-G., M.O. Scully, and H. Walther. 1994. The Duality in Matter and Light. *Scientific American* 271 (6): 56–61.
Epperson, M., and E. Zafiris. 2013. *Foundations of Relational Realism: A Topological Approach to Quantum Mechanics and The Philosophy of Nature*. Lexington Books.
Everett, H. III. 1957. Relative State' Formulation of Quantum Mechanics. *Review of Modern Physics* 29: 454–62. 1; Report in [Dewitt/Graham 1973, 141–49].
Everett, H.III. 1973. The Theory of the Universal Wave Function. In: [Dewitt/Graham 1973, 141–49].

Fano, V. (ed.). 2014. *Gino Tarozzi Philosopher of Physics: Studies in the Philosophy of Entanglement in his 60th Birthday*. Milano: Franco Angeli.
Ferrero, M., T.W. Marshall, and E. Santos. 1990. Bell's Theorem: Local Realism versus Quantum Mechanics. *American Journal of Physics* 58: 683–88.
Fine, A. 1981. Einstein's Critique of Quantum Theory: The Roots and Significance of EPR. In *After Einstein*, ed. P. Barker, and C.G. Shugart, 147–59. Memphis State University Press; Report in [Fine 1986, 26–39].
Furusawa, A., J.L. SøRensen, S.L. Braunstein, C.A. Fuchs, H.J. Kimble, and E.S. Polzik. 1998. Unconditional Quantum Teleportation. *Science* 282: 706–709.
Garuccio, A., and F. Selleri. 1984. Enhanced Photon Detection in EPR Type Experiments. *Physics Letters* 103A: 99–103.
Garuccio, A., V. Rapisarda, and J.-P. Vigier. 1982. Next Experimental Set-Up for the Detection of de Broglie Waves. *Physics Letters* 90A: 17–19.
Geroch, R. 1978. *General Relativity From A to B*. Chicago: University of Chicago Press.
Giulini, D., E. Joos, C. Kiefer, J. Kupsch, I.-O. Stamatescu, and H.D. Zeh (eds.). 1996. *Decoherence and the Appearance of a Classical World in Quantum Theory*. Berlin: Springer.
Gleason, A.M. 1957. Measures on the Closed Subspaces of a Hilbert Space. *Journal of Mathematics and Mechanics* 6: 885–93; Report in [Hooker 1975, 123–33].
Graham, N. 1973. The Measurement of Relative Frequency. in [Dewitt/Graham 1973, 229–53].
Greenberger, D.M., M.A. Horne and A. Zeilinger. 1989. Going Beyond Bell's Theorem. in [Kafatos 1989, 69–72].
Greenberger, D.M., M.A. Horne, A. Shimony, and A. Zeilinger. 1990. Bell's Theorem Without Inequalities. *American Journal of Physics* 58: 1131–43.
Griffiths, R.B. 1984. Consistent Histories and the Interpretation of Quantum Mechanics. *Journal of Statistical Physics* 36: 219–72.
Halliwell, J.J. 1993. Quantum-Mechanical Histories and the Uncertainty Principle: Information-Theoretic Inequalities. *Physical Review D* 48: 2739–52.
Halliwell, J.J. 2010. Macroscopic Superpositions, Decoherent Histories, and the Emergence of Hydrodynamic Behavior. in [Saunders et al. 2010, 99–117].
Hardy, L. 1992. On the Existence of Empty Waves in Quantum Theory. *Physics Letters* 167A: 11–16.
Harrigan, N. and R.W. Spekkens. 2010. Einstein, Incompleteness, and the Epistemic View of Quantum States. *Foundations of Physics* 40: 125–57. https://doi.org/10.1007/s10701-009-9347-0
Hartle, J.B., and S.W. Hawking. 1983. Wave Function of the Universe. *Physical Review D* 28: 2960–75.
Heisenberg, W. 1927. Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. *Zeitschrift für Physik* 43: 172–98; Report in [Heisenberg GW, A.I, 478–504].
Heisenberg, W. 1952. *Philosophic Problems of Nuclear Science*. Woodbridge (Connecticut): Ox Bow Press.
Heisenberg, W. 1958. *Physics and Philosophy*. New York: Harper.
Heisenberg, W. 1969. *Der Teil und Das Ganze*. München: Piper.
Hempel, C.G. 1953. *Methods of Concept Formation Science*. Chicago: University of Chicago Press. (NA)
Hiley, B.J. 1999. Active Information and Teleportation. In *Epistemological and Experimental Perspectives on Quantum Physics*, ed. Greenberger, et al., 113–26. Dordrecht: Kluwer.
Holevo, A.S. 1998. The Capacity of Quantum Channel with General Signal States. *IEEE Transactions on Information Theory* 44: 269–273.
Holland, P.R. 1993. *The Quantum Theory of Motion*. Cambridge: Cambridge University Press.
Home, D., and M.A.B. Whitaker. 2007. *Einstein's Struggles with Quantum Theory: A Reappraisal*. Berlin: Springer.
Horodecki, M., J. Oppenheim, and A. Winter. 2005. Partial Quantum Information. *Nature* 436: 673–76.

Howard, D. 1992. Einstein and Eindeutigkeit: A Neglected Theme in the Philosophical Background to General Relativity. In *Studies in the History of General Relativity*, ed. J. Eisenstaedt, and A.J. Kox. Birkhäuser: Bioston.
Huang, K. 1963. *Statistical Mechanics*, 2nd ed. New York: Wiley.
Husserl, E. 1931. *Méditations CartéSiennes*. Paris:
Jammer, M. 1974. *The Philosophy of Quantum Mechanics. The Interpretation of Quantum Mechanics in Historical Perspective*. New York: Wiley.
Joos, E., and H.D. Zeh. 1985. The Emergence of Classical Properties Through Interaction with the Environment. *Zeitschrift für Physik* B59: 223–43.
Kant, I. 1763. *Der einzige mögliche Beweisgrund zu einer Demonstration des Daseins Gottes*. in [Kant A, II, 63–164].
Kant, I. 1787. *Kritik der reinen Vernunft*, 2nd ed. 1787: III v. of [Kant A].
Khinchin, A.I. 1957. *Mathematical Foundations of Information Theory* (Engl. Tr.), New York: Dover.
Kochen, S., and E. Specker. 1965. Logical Structures Arising in Quantum Theory. In *Theory of Models*, ed. H.L. Addison, and A. Tarsky. Amsterdam, North-Holland. Report in [Hooker 1975, 263–76].
Krenn, G., and A. Zeilinger. 1996. *Entangled Entanglement. Physical Review* A54: 1793–97.
Landsman, K. 2017. *Foundations of Quantum Theory: From Classical Concepts to Operator Algebras*. Berlin: Springer.
Laughlin, R.B. 2005. *A Different Universe (Reinventing Physics from the Bottom Down)*. New York: Basic Books.
Leibniz, G.W. 1686. *Discours de Metaphysique*. in [Leibniz PS, IV, 427–63].
Leibniz, G.W. 1978. Philosophische Schriften (Ed. Gerhardt), Halle, 1875. rep. Olms: Hildesheim.
Leibniz, G.W. 1710. *Essais de Théodicée*; In [Leibniz PS, VI, 21–436]
Leibniz, G.W. 2019. Philosophische Schriften (Ed. Gerhardt), Halle, 1875; Report Olms: Hildesheim.
Lewis, D. 1986. *On the Plurality of Worlds*. Oxford: Blackwell.
Lieb, E.H. 1975. Some Convexity and Subadditivity Properties of Entropy. *Bulletin of the American Mathematical Society* 81: 1–13.
Lindblad, G. 1973. Entropy, Information and Quantum Measurement. *Communications in Mathematical Physics* 33: 305–322.
Lindblad, G. 1983. *Non-Equilibrium Entropy and Irreversibility*. Dordrecht: Reidel.
Ling, S., and C. Xing. 2004. *Coding Theory: A First Course*. Cambridge: Cambridge University Press.
Lockwood, M. 1966. Many Minds Interpretations of Quantum Mechanics. *British Journal for Philosophy of Science* 47: 159–88.
London, F., and E. Bauer. 1939. *La Théorie de L'observation en Mécanique Quantique*. Paris: Hermann; Eng. Trans. in [Wheeler/Zurek 1983, 217–59].
Lüders, G. 1951. Über die Zustandsänderung durch Meßprozeß. *Annalen der Physik* 8: 322–28.
Mana P.G.L. 2004. Consistency of the Shannon entropy in Quantum Experiments. *Physical Review* A69: 062108. https://doi.org/10.1103/PhysRevA.69.062108
Margenau, H. 1950. *The Nature of Physical Reality: A Philosophy of Modern Physics*. New York: McGraw-Hill, 1950; Woodbridge, Conn.: Ox Bow Press.
Marshall, T.W., and E. Santos. 1985. Local Realist Model for the Coincidence Rates in Atomic-Cascade Experiments. *Physics Letters* 107A: 164–68.
Marshall, T.W., E. Santos, and F. Selleri. 1983. Local Realism Has not Been Refuted by Atomic Cascade Experiments. *Physics Letters* 98A: 5–9.
Masanes, L., A. Acin, and N. Gisin. 2006. General Properties of Nonsignaling Theories. *Physical Review* A73: 012112-1–9.
Maudlin, T. 1994. Quantum Non-locality and Relativity: Metaphysical Intimations of Modern Physics. Blackwell.

Mehra, J., and H. Rechenberg. 1982. *The Historical Development of Quantum Theory*. New York: Springer.
Mittelstaedt, P., A. Prieur, and R. Schieder. 1987. Unsharp Particle-Wave Duality in a Photon Split-Beam Experiment. *Foundations of Physics* 17: 891–903.
Monroe, C., D.M. Meekhof, B.E. King, and D.J. Wineland. 1996. A Schrödinger Cat' Superposition State of an Atom. *Science* 272: 1131–36.
Munowitz, M. 2005. *Knowing: The Nature of Physical Law*. Oxford: Oxford University Press.
Nagel, E. 1961. The Structure of Science. Harcourt-Brace.
Nielsen, M.A., and I.L. Chuang. 2000. *Quantum Computation and Quantum Information*. Cambridge: Cambridge University Press.
Norsen, T. 2017. *Foundations of Quantum Mechanics: An Exploration of the Physical Meaning of Quantum Theory*. Berlin: Springer.
Opera. 1831. (ed. Immanuel Bekker), Berlin, G. Reimerum.
Ohya, M., and D. Petz. 1993. *Quantum Entropy and its Use*. Berlin: Springer.
Oppenheim, J. and S. Wehner. 2010. The Uncertainty Principle Determines the Nonlocality of Quantum Mechanics. *Science* 330: 1072. https://doi.org/10.1126/science.1192065
Ott, E. 1993. *Chaos in Dynamical Systems*. Cambridge: Cambridge University Press.
Ou, Z.Y., and L. Mandel. 1988. Violation of Bell's Inequality and Classical Probability in a Two-Photon Correlation Experiment. *Physical Review Letters* 61: 50–53.
Partovi, H.M. 1989. Irreversibility, Reduction and Entropy Increase in Quantum Measurements. *Physics Letters* 137A: 445–50.
Pasnau, R. 2004. Form, Substance, and Mechanism. *Philosophical Review* 113: 31–88.
Pawłowski, M., T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, and M.Z. Żukowski. 2009. Information Causality as a Physical Principle. *Nature* 461: 1101–104.
Pawłowski, M. and V. Scarani. 2016. Information Causality. in [Chiribella/Spekkens 2016, 423–38].
Paz, J.P., S. Habib, and W.H. Zurek. 1993. Reduction of the Wave Packet: Preferred Observable and Decoherence Time Scale. *Physical Review D* 47: 488–501.
Pearl, J. 1988. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Mateo: Morgan Kaufmann.
Peirce, C.S. 1877. The Fixation of Belief. *Popular Science Monthly* 12: 1–15; Report in [Peirce W, III, 242–57].
Peirce, C.S. 1891. The Architecture of Theories. *Monist* 1: 161–76; in [Peirce W, VIII, 98–110].
Peirce, C.S. 1898. *Reasoning and The Logic of Things: The Cambridge Conferences Lectures of 1898*. Manuscript; ed. K. L. Ketner. Cambridge: Harvard University Press.
Peirce, C.S. 1902. On Science and Natural Classes. in [Peirce EP, II, 115–32].
Peres, A., and W.H. Zurek. 1982. Is Quantum Theory Universally Valid? *American Journal of Physics* 50: 807–10.
Peshkin, M., and A. Tonomura. 1989. *The Aharonov-Bohm Effect*. Berlin: Springer.
Pitowsky, I. 1989. *Quantum Probability-Quantum Logic*. Berlin: Springer.
Planck, M. 1922. *The Origin and Development of the Quantum Theory*. Oxford: Clarendon Press.
Plesch, M. and V. Bužek. 2010. Efficient Compression of Quantum Information. *Physical Review* A81: 032317. https://doi.org/10.1103/PhysRevA.81.032317
Poincaré, Henri. 1905. *La Valeur de la science*. Paris: Flammarion.
Popescu, S., and D. Rohrlich. 1994. Quantum Nonlocality as an Axiom. *Foundations of Physics* 24: 379–85.
Popper, K.R. 1982. *Quantum Theory and the Schism in Physics*. London: Unwin Hyman Ltd; Routledge.
Rees, M. 1999. *Just Six Numbers*, Weidenfeld and Nicolson; London: Phoenix.
Reid, T. 1764. *An Inquiry into the Human Mind on the Principles of Common Sense, 1764*, 5th ed. Edinburgh: Bell and Co.
Renninger, M.K. 1960. Messungen ohne Störung des Meßobjekts. *Zeitschrift für Physik* 158: 417–21.
Rindler, W. 2001. *Relativity: Special, General, and Cosmological*. Oxford: Oxford University Press.

Rovelli, C. 1996. Relational Quantum Mechanics. *International Journal of Theoretical Physics* 35:1637–78. https://doi.org/10.1007/bf02302261
Rovelli, C. 2005. Relational Quantum Mechanics. in [Elitzur et al. 2005, 113–20].
Rozema, L.A., D.H. Mahler, A. Hayat, P.S. Turner and A.M. Steinberg. 2014. Quantum Data Compression of a Qubit Ensemble. *Physical Review Letters* 113: 160504. https://doi.org/10.1103/PhysRevLett.113.160504
Shapere, A., and F. Wilczek (eds.). 1989. *Geometric Phases in Physics*. Singapore: World Scientic.
Santos, E. 1991. Does Quantum Mechanics Violate the Bell Inequalities? *Physical Review Letters* 66: 1388–90; Errata: 3227.
Saunders, S., J. Barrett, A. Kent, and D. Wallace (eds.). 2010. *Many Worlds?*. Oxford: Oxford University Press.
Schlosshauer, M. 2007. *Decoherence and the Quantum-to-Classical Transition*. Berlin: Springer.
Schrödinger, E. 1935a. Die gegenwärtige Situation in der Quantenmechanick. I–III. *Naturwissenschaften* 23: 807–12, 823–28, 844–49; Report in [Schröodinger GA, III, 484–501].
Schrödinger, E. 1935b. Discussion of Probability Relations Between Separated Systems. *Proceedings of the Cambridge Philosophical Society* 32: 446–52; Report in [Schröodinger GA, I, 433–39].
Schrödinger, E. 1936. Probability Relations Between Separated Systems. *Proceedings of the Cambridge Philosophical Society* 31: 555–63; Report in [Schröodinger GA, I, 424–32].
Schrödinger, E. 1958. *Mind and Matter*. Cambridge: Cambridge University Press; Report in [Schrödinger 1992, 91–164].
Schrödinger, E. 1992. *What is Life? with Mind and Matter and Autobiographical Sketches*. Cambridge: Cambridge University Press (enlarged ed. of [?]).
Schumacher, B.W. 1990. Information from Quantum Measurements. in [Zurek 1990, 29–37].
Schuster, H.G. 1988. *Deterministic Chaos: An Introduction*, 2nd ed. Weinheim: VCH; 3rd ed. 2005.
Scully, M.O., B.-G. Englert, and H. Walther. 1991. Quantum Optical Tests of Complementarity. *Nature* 351: 111–16.
Selleri, F. 1969. On the Wave Function of Quantum Mechanics. *Lettere a Nuovo Cimento* 1: 908–910.
Serjeant, S. 2010. *Observational Cosmology*. Cambridge: Cambridge University Press.
Shafiee, A., F. Safinejad and F. Naqsh. 2006. Information and the Brukner–Zeilinger Interpretation of Quantum Mechanics: A Critical Investigation. *Foundations of Physics Letters* 19: 1–20. https://doi.org/10.1007/s10702-006-1845-0
Shannon, C.E. 1948. A Mathematical Theory of Communication. *Bell System Technical Journal* 27 (379–423): 623–56.
Shih, Y.H., and C.O. Alley. 1988. New Type of EPR-Bohm Experiment Using Pairs of Light Quanta Produced by Optical Parametric Down Conversion. *Physical Review Letters* 61: 2921–24.
Shimony, A. 1965. Quantum Physics and the Philosophy of Whitehead. In *Philosophy in America*, ed. Max Black. London: Allen and Unwin; Report in [Shimony 1993, II, 291–309].
Shimony, A. 1981. Réflections sur la philosophy de Bohr, Heisenberg et Schrödinger. *Journal de physique* 42: 81–95; Eng. Tr. in [*SHIMONY* 1993, II, 310–22] with the title "Reflections on the Philosophy of Bohr, Heisenberg, and Schrödinger".
Shimony, A. 1983. Controllable and Uncontrollable Non-Locality. In *Foundations of Quantum Mechanics in the Light of New Technology*, ed. S. Kamefuchi, et al., 225–30. Japan: Physical Society of Japan. Report in [Shimony 1993, II, 130–139].
Shimony, A. 1993. *Search for a Naturalistic Point of View*. Cambridge: Cambridge University Press.
Shull, C.G., D.K. Atwood, J. Arthur, and M.A. Horne. 1980. Search For a Nonlinear Variant of the Schrödinger Equation by Neutron Interferometry. *Physical Review Letters* 44: 765–68.
Smolin, L. 1997. *The Life of the Cosmos*. Oxford: Oxford University Press.
Spinoza, B. 1677. *Ethica*. Amsterdam; In [Spinoza O, II, 41–308].
Spivak, D.I. 2013. *Category Theory for Scientists*. Cambridge: MIT Press.
Stachel, J. 2017. It Ain't Necessarily So: Interpretations and Misinterpretations of Quantum Theory. in [Kastner et al. 2017, 53–74].

Stanzione, M. 2013. Evoluzionismo e teleologia: riconsiderazione storico-teorica del problema. in [Auletta/Pons 2013, 81–104].
Stapp, H.P. 1982. Mind, Matter, and Quantum Mechanics. *Foundations of Physics* 12: 363–99; Report in [Stapp 1993, 81–118].
Stapp, H.P. 1993. *Mind, Matter, and Quantum Mechanics*, 2nd ed, 2004. Berlin: Springer.
Tegmark, M., A. Aguirre, M.J. Rees, and F. Wilczek. 2006. Dimensionless Constants, Cosmology, and Other Dark Matters. *Physical Review D* 73: 023505.
'T Hooft, G. 2016. *The Cellular Automaton Interpretation of Quantum Mechanics*. Berlin: Springer.
Tegmark, M. 2003. Parallel Universes. *Scientific American*.
Timpson, C.G. 2013. *Quantum Information Theory and the Foundations of Quantum Mechanics*. Oxford: Oxford University Press.
Torcal, L. 2013. Complessità, direzionalità e finalismo. in [Auletta/Pons 2013, 235–55].
Tsirelson, B.S. 1980. Quantum Generalizations of Bell's Inequality. *Letters in Mathematical Physics* 4: 93–100.
van Fraassen, B.C. 1991. *Quantum Mechanics. An Empiricist View*. Oxford: Clarendon.
Vedral, V., M.B. Plenio, M.A. Rippin, and P.L. Knight. 1997. Quantifying Entanglement. *Physical Review Letters* 78: 2275–79.
von Neumann, J. 1932. *Mathematische Grundlagen der Quantenmechanik*. Berlin: Springer.
Wallace, D. 2010. Decoherence and Ontology. In [Saunders et al. 2010, 53–72].
Wallace, D. 2012. *The Emergent Multiverse: Quantum Theory According to the Everett Interpretation*. Oxford: Oxford University Press.
Wang, L.J., X.Y. Zou, and L. Mandel. 1991. Induced Coherence Without Induced Emission. *Physical Review* A44: 4614–22.
Wehrl, A. 1978. General Properties of Entropy. *Review of Modern Physics* 50: 221–60.
Wheeler, J.A. 1983. *Law Without Law*, 182–213. in [Wheeler/Zurek 1983].
Wheeler, J.A. 1990. Information, Physics, Quantum: The search for Links. in: [Zurek 1990, 3–28].
Whitehead, A.N. 1925. *Science and the Modern World*. New York: Macmillan.
Whitehead, A.N. 1929. *Process and Reality*. London: Macmillan Publication, corr. ed. 1978, 1979.
Wiesner, S.J. 1983. *Conjugate Coding. SIGACT News* 15: 78–88.
Wigner, E.P. 1960. The Unreasonable Effectiveness of Mathematics in the Natural Sciences. *Communications in Pure and Applied Mathematics* 13: 1–13.
Wigner, E.P. 1961. Remarks on the Mind–Body Question. In *The Scientist Speculates*, ed. I.J. Good, 284–302. London: Heinemann; Report in [Wheeler/Zurek 1983, 168–81].
Withaker, M.A.B. 2004. The EPR Paper and Bohr's Response. *Foundations of Physics* 34: 1305–40.
Wootters, W.K., and W.H. Zurek. 1979. Complementarity in the Double-Slit Experiment: Quantum Nonseparability and a Quantitative Statement of Bohr's Principle. *Physical Review* D19: 473–84; Report in [Wheeler/Zurek 1983, 443–54].
Zeh, H.D. 1970. On the Interpretation of Measurement in Quantum Theory. *Foundations of Physics* 1: 69–76; Report in [Wheeler/Zurek 1983, 342–49].
Zeilinger, A. 2005. The Message of the Quantum. *Nature* 438: 743.
Zou, X.Y., T.P. Grayson, L.J. Wang, and L. Mandel. 1992. Can an Empty de Broglie Pilot Wave Induce Coherence? *Physical Review Letters* 68: 3667–69.
Zurek, W.H. 1981. Pointer Basis of Quantum Apparatus: into What Mixture Does the Wave Packet Collapse? *Physical Review D* 24: 1516–25.
Zurek, W.H. 1982. *Environment-Induced Superselection Rules. Physical Review* D26: 1862–80.
Zurek, W.H. 1989a. Thermodynamic Cost of Computation, Algorithmic Complexity and the Information Metric. *Nature* 341: 119–24.
Zurek, W.H. 1989b. Algorithmic Randomness and Physical Entropy. *Physical Review* A40: 4731–51.
Zurek, W.H. 1991. Decoherence and the Transition from Quantum to Classical. *Physics Today* 44 (10): 36–44.
Zurek, W.H. 2004. Quantum Darwinism and Envariance. In [BARROW et al. 2004, 121–37].

Zurek, W.H. 2007. Quantum Origin of Quantum Jumps: Breaking of Unitary Symmetry Induced by Information Transfer in the Transition from Quantum to Classical. *Physical Review* A76: 052110-1–5.

Zurek, W.H. 2009. Quantum Darwinism. *Nature Physics* 5: 181–88. https://doi.org/10.1038/nphys1202

Zurek, W.H. 2010. Quantum Jumps, Born's Rule, and Objective Reality. In [Saunders et al. 2010, 409–32].

Zurek, W.H. 2013. Wave-Packet Collapse and the Core Quantum Postulates: Discreteness of Quantum Jumps from Unitarity, Repeatability, and Actionable Information. *Physical Review* A87.5: 052111. https://doi.org/10.1103/PhysRevA.87.052111

Zwolak, M., and W.H. Zurek. 2013. Complementarity of Quantum Discord and Classically Accessible Information. *Scientific Reports* 3: 1729. https://doi.org/10.1038/srep01729

Żukowski, M., A. Zeilinger, M.A. Horne and A.K. Ekert.,. 1993. Event-Ready-Detectors Bell Experiment via Entanglement Swapping. *Physical Review Letters* 71: 4287–90.

Chapter 4
Quantum Computation



The previous examination has raised an important issue: the effectiveness of quantum information. Here, I shall present quantum computation and its basic operations as the best way to see the 'machinery' of information at work: this will be also crucial for the foundational role of information (to be developed in the next part). I have stressed that the state of a quantum system encodes information (Sect. 3.2.3). Moreover, it changes with the time (according to the Schrödinger or Heisenberg equations). Therefore, it can be conceived as an information *processor* that maps some codified information to another kind of codified information.

After having introduced reversible computation, I shall present the main quantum gates and their capability to generate relevant kinds of states. Then, I shall deal with three main quantum algorithms: Deutsch's, Shor's, and Grover's. Subsequently, I shall deal with the Heisenberg picture for quantum computation and with quantum-gate teleportation. Finally, the relation energy time of computation is analysed.

4.1 A Reversible Multipath Computation

Quantum computation started with some pathbreaking papers of the American physicist Richard Feynman (1918–1988) and D. Deutsch pointing out that a computer working with the quantum principles could simulate any physical process describable in finite terms employing a finite number of qubits and a finite number of gates (information transformers), thus also circumventing some of the limitations of classical computers.[1] A related issue is that quantum computation provides a precise logical analysis of quantum processes and also a very intuitive graphical representation of them. In the following, this stuff will help us with the analysis of the operations performed during measurement, which is in fact a process of information acquisition.

[1] Feynman (1982, 1986), Deutsch (1985b).

© Springer Nature Switzerland AG 2019

G. Auletta, *The Quantum Mechanics Conundrum*,
https://doi.org/10.1007/978-3-030-16649-6_4

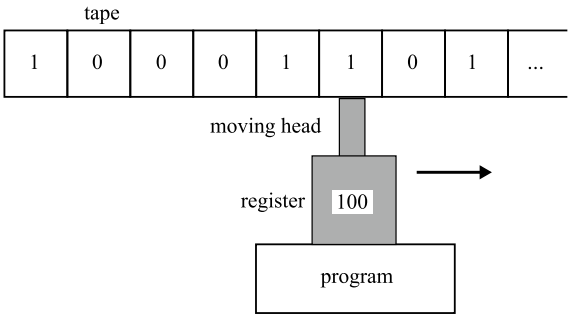


Fig. 4.1 Scheme of a Turing machine. A moving head goes (here from the left to the right) along a tape that is divided into cells (here, storing binary encoded information) with the task to read and eventually to modify (write) the content of each cell. A register stores the state of the Turing machine. Finally the program is the appropriate set of instructions. Adapted from Auletta and Wang (2014, p. 350)

Turing Machine

Classical computers are instantiations of the universal Turing machine. A *Turing machine*, after the name of the British computer scientist, mathematician, logician, cryptanalyst, philosopher and theoretical biologist Alan Turing (1912–1954),[2] is composed of four parts (Fig. 4.1):

(i) A tape that is divided into cells, one next to the other. Each cell can have a symbol from a finite alphabet (like binary numbers).
(ii) A head that can read and write symbols on the tape and move along the tape left or right one cell at a time.
(iii) A register that stores the state of the Turing machine, one of finitely many.
(iv) A table of a finite collection of instructions that tell the machine how operations are performed.

The action of the Turing machine is made up of discrete steps, and each step is determined by two initial conditions: the current state of the machine and the symbol currently scanned by the head. Given these two conditions, the machine performs the following operations in sequence: (a) assume the same or a new state of the machine. (b) Write a symbol into the scanned cell. (c) Move the head left or right along the tape or stop. The Turing machine is the logical structure of any ordinary computer. Moreover, a universal Turing machine is an abstract model capable to simulate any other Turing machine. Ordinary classical computers as instantiations of the Turing machine have electrical connections that not only provide for energy supply (as it is the case for any ordinary device performing some operations, like a traditional, non-computerised, washing machine) but also for logical connections. For instance, let us consider the circuit shown in Fig. 4.2. Negation can be implemented by an open

[2] Turing (1936).

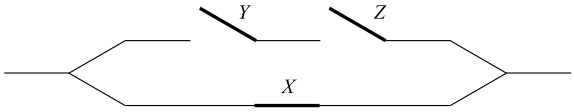


Fig. 4.2 Scheme of a logical circuit. Electric current flows from the right to the left. Logical gates (the transformers of the input information) are represented by thick lines for the sake of representation, while wires by thin lines. In order that electric current flows, it is necessary that the gates be closed (which corresponds to the value 1). The two main wires of the circuit are parallel. In other words, the current can take both paths simultaneously, which means that one path closed is sufficient for transmitting current. The wire at the top displays moreover a serial circuit, where both gates need to be closed for having a current. In specific, the circuit corresponds to the proposition $X + YZ$

gate (it inverts the truth value), a disjunction (logically represented by the symbol $+$) by two parallel gates (that is, two gates implemented in parallel wires), and a conjunction (represented by no symbol like the algebraic product) by serial gates: see Table 3.3 and 3.2 for their definition.

The Church–Turing thesis, after the names of A. Turing and the American mathematician and logician Alonzo Church (1903–1995), is that "if we can write an algorithm (a computational procedure), then it can be formulated as a Turing machine",[3] although there is undecidability/formal incompleteness, as we shall see now. The main idea here is that a completely random sequence represents a string that cannot be further compressed in a reliable way. In fact, in the case of a random sequence, "the smallest algorithm capable of specifying it has about the same number of bits of information as the series itself".[4] This defines algorithmic randomness. Note, however, that in formal systems (constituted by alphabet of symbols, grammar, list of axioms, rules of inference), no sequence can be proved to be random unless the complexity of the latter is less than that of the formal system. This means that there is the so-called *halting problem*: in many cases we do not know whether a Turing machine, when processing a sequence, will halt, and, if yes, when it will do that.[5]

Algorithms

An *algorithm* is a method in which a set of well-defined instructions, applied to a certain initial state or problem, can complete a required task, or solve a given problem. From a completely general viewpoint, let us consider an algorithm which takes an input of M digits and gives an answer after a certain number N of steps (or time). One says that an algorithm is *easy* or *efficient* if the time taken to execute it increases no faster than a polynomial function of the size of the input M (*polynomial-time solution*). On the contrary, solutions that are not efficient require a time that is

[3] Turing (1936), Turing (1937), Church (1936a, b), Church (1937). See also Chaitin et al. (2011, p. 25).

[4] Chaitin (1975).

[5] Chaitin et al. (2011, p. 75). On Rice's theorem on programs see Chaitin et al. (2011, p. 13).

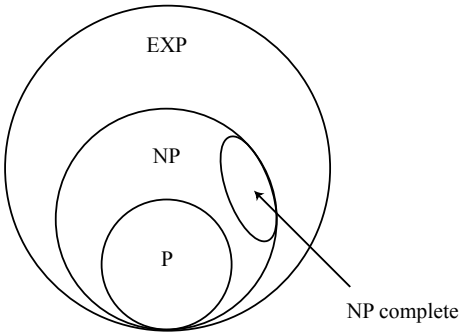


Fig. 4.3 Representation of computational complexity. In the 'space' of all possible problems (algorithms), P stands for the subset of algorithms that require a polynomial time calculation, NP for algorithms requiring polynomial time on a non-deterministic computer (roughly speaking a computer with an unbound degree of parallelism), EXP for exponential time calculation. Problems are designated "NP-complete" if their solutions can be quickly (polynomially) checked for correctness. Many kinds of optimization problems (e.g. that of the "traveling salesman") belong to the NP-complete category. A typical example of problem belonging to EXP is constituted by chess game. Adapted from Auletta et al. (2009, p. 660)

exponential in the size of the input. Many operations can be performed using an efficient algorithm, either logarithmic or polynomial (in most of the cases). Now, there is a class of operations whose algorithms are not efficient classically. We can pictorially represent the degrees of computational complexity as in Fig. 4.3. The Cook's theorem, after the name of the American computer scientist Stephen Cook, states that if NP-complete = P, then NP $\subset$ P. This is so because we can map any NP-problem to a NP-complete one with a polynomial-time procedure.

General Characters of Quantum Computation

There are some important characters of quantum computation that need to be stressed at the start: quantum systems can work as computers but *in a reversible way*, since all relevant transformations are unitary[6]; moreover, they can work much more efficiently than classical ones[7]:

- Thanks to the reversibility of the transformation, quantum computers can process information without an expenditure of energy up to the moment in which we select a specific output (that is, when we acquire information), an operation that is comparable to detection.[8] In other words, we can have a (unitarily) transformed

[6] On the problems of reversibility and computation see Landauer (1961), Bennett (1973), Bennett and Landauer (1985).

[7] See for instance Lloyd (2000). See also Auletta et al. (2009, Sects. 17.7–8), Auletta and Wang (2014, Sect. 11.4).

[8] Nielsen and Chuang (2000, Sect. 3.2.5). I strongly recommend this textbook on quantum computing.

output and nevertheless the whole process between input and output waste no energy provided that we do not acquire this information: in these conditions, there is a change in information but not in energy. Although also classical computers can be abstractly thought to be reversible in the absence of information erasure, this turns out to be quite impossible to realise in practice due to the lack of quantum correlations. Therefore, quantum computation shows that information is more basic than energy (although the latter is fundamental for signal exchanging), as anticipated in Sect. 3.2.4.

- Due to the multipath dynamics thanks to the superposition principle,[9] they can solve problems that are not easily solvable by classical computers, since it is like taking several computing paths simultaneously. A known classical problem for which there are no efficient solutions is the factorisation of numbers that can instead be dealt with quantum computing.

Regarding the first item, in Sect. 3.2.2 I have introduced information as a dimensionless and formal quantity (a type). In that context, it appeared (and still appears to many scholars) as a limitation. At the opposite, this is its greatest strength. In fact, this is precisely what guarantees to information universal applicability, from photons to black holes,[10] from biomolecules to human communication.[11] This is also what makes of free quantum systems types. In fact, without additional knowledge, no particular state of, for example energy, position or momentum can be assigned and the only thing that we can say of them is that they represent multiple ways of codifying the same amount of information (according to different possible observables with their eigenbasis).

A Physical Process

Quantum computing sets the problem of computation in new terms, making of information processing a *physical process* (and vice versa), although the abstract logical–mathematical component (type) can still be distinguished from the physical one (type-token) and in fact only its presence makes of the latter a computation.[12] For D. Deutsch, quantum computer are universal. Indeed, he recalls us that it is the quantum nature of this computation to make discrete and also digital treatment of information possible. In other words, it is what allows for linear codes (Fig. 3.7, Sect. 3.2.3). On the basis of the universality of quantum computation, Deutsch tells us that the universe is essentially a flow of information.[13] The reason, according to him, is that, since an ordinary (quantum) computer is a universal environment simulator, the universe can be reciprocally understood as an information processor or computer.

---

[9] In recall, here, the fundamental textbook on path integrals (Feynman and Hibbs 2005).

[10] As pointed out in Lloyd and Ng (2004).

[11] The universality of information has been also stressed in Wheeler (1990). See also Auletta (2006a).

[12] A thesis that has been perhaps overemphasised in Timpson (2013, Chap. 6). As we shall see, the basic structure of our universe is indeed logical-informational.

[13] Deutsch (1997, 2011). See also Lloyd (2002, 2006).

Table 4.1 Truth table for Toffoli gate. It is noted that X and Y are the control bits and are not changed under the transformation

INPUT			OUTPUT	
X	Y	Z	XY	$Z \nsim XY$
0	0	0	0	0
0	0	1	0	1
0	1	0	0	0
0	1	1	0	1
1	0	0	0	0
1	0	1	0	1
1	1	0	1	1
1	1	1	1	0

In order to understand the exact meaning of this, we need to take into account that quantum computation is not only information *processing* but makes use also of information *sharing* among different qubits (mutual information) and information *selection*, in a global dynamics. In other words, all relevant operations on information (on which I shall come back) are involved in computation. This is what allows us to understand that computation is a physical, and even the *basic physical*, process. I remind that the German–American physicist Rolf Landauer (1927–1999) proposed that "information is physical".[14] There are many ways to interpret this statement (and in fact, it is diversely interpreted). If the sense is that information is the basic 'stuff' of our physical universe, I agree, but if this means that information can be reduced to some other physical parameter, I do not.

Local Selection and Global Reversibility

Since information-selection procedures are local processes, it is possible to include information selection as a local procedure in a larger network that is still reversible. Let us consider an example of reversible computation that is also an example of how decoherence can work. The analysis is provided in pure logical terms and its results are valid for both classical and quantum computation. Let us consider the conjunction between X and Y. I recall that it is false only when X or Y is false (Table 3.2, Sect. 3.3). Now, consider not just two inputs but enlarge the number of inputs to three inputs, denoted by X, Y and Z. For our purposes, we shall make use of the so-called Toffoli gate,[15] after the name of the Italian computer scientist Tommaso Toffoli, which is displayed in Table 4.1. We note that, according to Table 3.4, Sect. 3.3, the exclusive disjunction (XOR) $Z \nsim XY$ in the last of the two output columns on the right is true only when either (i) $Z = 1$ and $XY = 0$ or (ii) $Z = 0$ and $XY = 1$. It is false only

[14] Landauer (1996).

[15] Fredkin and Toffoli (1982).

when either (i) $Z = 1$ and $XY = 1$ or (ii) $Z = 0$ and $XY = 0$. We have then the following inputs represented by the three input columns on the left:

(i) The second, fourth and sixth rows correspond to the case $Z = 1$ and $XY = 0$.
(ii) The seventh row corresponds to the case $Z = 0$ and $XY = 1$.
(iii) The latter row corresponds to $Z = 1$ and $XY = 1$.
(iv) The first, third and fifth rows correspond to the case $Z = 0$ and $XY = 0$.

Taking into account these inputs and the XY output, let us consider the last of the two output columns on the right:

(i) The values on the second, fourth and sixth rows are 1 since $Z = 1$ and $XY = 0$.
(ii) The value on the seventh row is again 1 since $Z = 0$ and $XY = 1$.
(iii) The value on the last row is 0 since $Z = 1$ and $XY = 1$.
(iv) The values on first, third and fifth rows are 0 since $Z = 0$ and $XY = 0$.

Therefore, if we only consider the cases in which both $Z = 1$ and $Z \nsim XY = 1$, we obtain precisely all the three cases in which $XY = 0$ (second, fourth and sixth rows). This subset of the outputs is therefore determined by setting $Z = 1$ and can be expressed as

$$Z = 1 \text{ AND } (XY)' \quad \text{or also} \quad Z = 1 \text{ AND } (X' + Y'), \tag{4.1}$$

where I have used de Morgan's law (3.130). It should be noted that we could obtain the same result by choosing an alternative subset, i.e. when both $Z = 0$ and $Z \nsim XY = 0$. Also, this choice covers precisely all three cases in which $XY = 0$. Therefore, with both choices, we have obtained the disjunction (OR) between X' and Y', which in itself is an irreversible operation but is embedded here in a larger reversible transformation. Note indeed that all outputs of the Toffoli gate are univocally mapped to the inputs: all truth values of X and Y are maintained, while the values of Z and $Z \nsim XY$ are also the same apart from the last two lines (when both X and Y are true) whose truth value is inverted. This transformation will be also described in quantum-mechanical terms in the following.

Decoherence

Quantum computation is the best example of spontaneous decoherence (Sect. 3.1.3) and therefore another piece of evidence that our controlled operations in labs are not essentially different from those occurring spontaneously in nature. In fact, the biggest problem in the realisation of effective quantum computers are not the algorithms, which are known more and more better, but the fact that spontaneous decoherence processes happen disturbing or even destroying that coherence that makes the algorithms working. In other words, quantum computers need to be protected. Significant progress have been made, especially by IBM, by building and planning processors that work between 17 and 50 qubits at the current stand of technology. 50 qubits would already represent a significant improvement on classical computation. It is

interesting to observe that the IBM researchers introduced the notion of *quantum volume* that accounts for number and quality of qubits, circuit connectivity and error rates of operations.

4.2 Quantum Gates

4.2.1 The Basic Gates

Wires and Gates

Both classical and quantum-mechanical computation requires a logical-machine able to process information. The Turing machine serves as an idealized model for logical-mathematical calculation (Fig. 4.1, Sect. 4.1). I recall that according to the Church–Turing thesis, any algorithm process can be simulated efficiently using a Turing machine, where I recall that with "efficiently", it is understood a computation in at most polynomial time.[16]

As mentioned, quantum computation like the classical one is performed through *circuits* built of wires and gates that instantiate logical relations: *wires* pass the information over while *gates* transform it unitarily (as said, all quantum gates are unitary). A logic gate is a Boolean function

$$f : \{0, 1\}^j \longmapsto \{0, 1\}^k \tag{4.2}$$

from some fixed number j of input (qu-)bits to some fixed number k of output (qu-)bits, and where the symbol ":" means the definition of the mapping on its left given by the transformation on its right. I shall introduce some examples of gates.[17]

Phase Gate

An elementary one-qubit gate is the *phase shift* gate, which on the computational basis states $|0\rangle$ and $|1\rangle$ acts as (see also Eq. (2.29))

$$\hat{U}_\phi |0\rangle = |0\rangle \ , \quad \hat{U}_\phi |1\rangle = e^{i\phi} |1\rangle \ . \tag{4.3}$$

As usually Eqs. (3.81), writing the basis states $|0\rangle$ and $|1\rangle$ in component form as

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{4.4}$$

[16] Turing (1936, 1937), Church (1936a). On this subject, see also Nielsen and Chuang (2000, Sect. 1.1 and Chap. 3).

[17] The theory of quantum gates and the basic examples were introduced in Deutsch (1985b).

respectively, we can define $\hat{U}_\phi$ in matrix form as

$$\hat{U}_\phi := \begin{bmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{bmatrix}, \tag{4.5}$$

which is precisely the unitary operator representing the phase shifter PS on the upper path in the Mach–Zehnder interferometer giving a phase shift ϕ to the component $|u\rangle$, as displayed in Fig. 2.2, Sect. 2.1.2.

Hadamard Gate

Another gate is the *Hadamard* gate that takes the form

$$\hat{U}_{\rm H} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}. \tag{4.6}$$

Note that, we obtain the Hadamard gate from the 2D Hilbert-space rotation matrix (see e.g. Fig. 1.9, Sect. 1.1.3)

$$\hat{R}(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \tag{4.7}$$

for $\theta = \pi/4$ and flipped by the Pauli-X matrix (Fig. 4.4), i.e.

$$\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}. \tag{4.8}$$

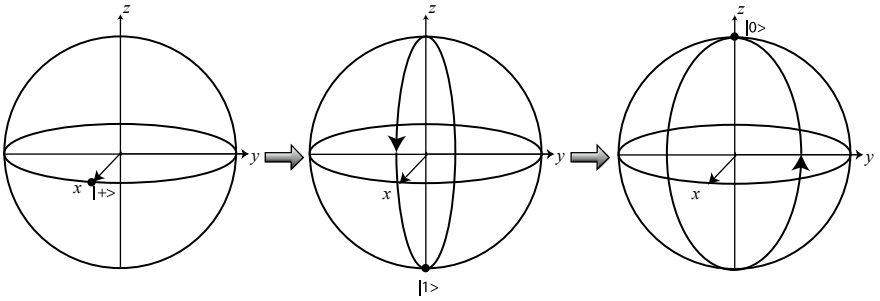


Fig. 4.4 Poincaré sphere representation of how the Hadamard gate acts on the state $|+\rangle$ (see also Fig. 3.6, Sect. 3.2.3). It can be thought first to induce a 90° CW rotation about the y axis (bringing to state $|1\rangle$), corresponding to the action of the rotation matrix (4.7) around y for $\theta = 90°$, and then a 180° CW rotation about the x axis (bringing to the state $|0\rangle$), correspond to the action of the Pauli X matrix. As previously noted (see comments to Eq. (3.80)), the angle θ on the Poincaré sphere is doubled relative to rotations in the Hilbert space

An example of Hadamard gate is given by the 50–50% beam-splitting unitary operator (see Eq. (2.27)). On the basis $\{|0\rangle\,, |1\rangle\}$, it works as Eq. (3.82)

$$|0\rangle \overset{\mathrm{H}}{\longmapsto} \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) = |+\rangle\,, \tag{4.9}$$

$$|1\rangle \overset{\mathrm{H}}{\longmapsto} \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) = |-\rangle\,, \tag{4.10}$$

where I have followed the convention to denote H the Hadamard matrix $\hat{U}_{\mathrm{H}}$. Being the Hadamard gate the inverse of itself, two subsequent applications of it bring back to the original state, what implies that this transformation is both unitary and Hermitian. I have already used maps in the previous chapters, but since we deal here with computation diagrams that represent *sequences* of operations, it is convenient to write all of the relevant transformations in terms of maps $\longmapsto$ and not of equations. In the latter case, in fact, we should use the opposite convention and write the sequence of operations from the right to the left.

Pauli Gates

Important gates are represented by the one-qubit Pauli spin matrices (1.324). I recall that the matrix $\hat{\sigma}_x$ is called in quantum computation the NOT gate for a single qubit and can be denoted by $\hat{U}_{\mathrm{NOT}}$. In Sect. 3.3.6, I have already mentioned that it is also called the Pauli-X gate (or X gate for short). I also recall that it represents a rotation of the Poincaré sphere around the x-axis by π radians, as well as the Pauli matrices $\hat{\sigma}_y$ and $\hat{\sigma}_z$ are here the Pauli-Y gate (or Y gate for short) and Pauli-Z gate (or Z gate for short), representing rotations of the Poincaré sphere around the y- and z-axes by π radians, respectively.

Transformations that can be performed through a symmetric beam splitter can be cast as follows, where the phase difference (Z) is included in the transformation (see also Eq. (4.6)):

$$\mathrm{H}\,|0\rangle = |+\rangle\,,\quad \mathrm{H}\,|1\rangle = |-\rangle\,, \tag{4.11a}$$

$$\mathrm{ZH}\,|0\rangle = |-\rangle\,,\quad \mathrm{ZH}\,|1\rangle = |+\rangle\,, \tag{4.11b}$$

$$\mathrm{XH}\,|0\rangle = |+\rangle\,,\quad \mathrm{XH}\,|1\rangle = -\,|-\rangle\,, \tag{4.11c}$$

$$\mathrm{XZH}\,|0\rangle = -\,|-\rangle\,,\quad \mathrm{XZH}\,|1\rangle = |+\rangle\,. \tag{4.11d}$$

The effect of the composed transformations is to move the element -1 in the matrix (4.8) in all the four positions. Of course, we can also consider different forms of phase difference, so that *a priori* we can say that any 50–50% beam splitter (that could also easily be generalised) induces the general transformation

$$|0\rangle\,, |1\rangle \longmapsto \frac{1}{\sqrt{2}}\left(|0\rangle \pm e^{i\phi}\,|1\rangle\right). \tag{4.12}$$

Recalling that the Pauli matrices constitute a complete set Eq. (1.326), we can express any unitary single-qubit gate by their means.[18] First, according to the Stone's theorem (see Eq. (1.300)), we need to build exponentials with the involved operators, which turn out to be rotations about the three Cartesian axes x, y, z by an arbitrary angle $\theta/2$ Eq. (4.7):

$$\hat{R}_x(\theta) = e^{-i\frac{\theta}{2}\mathrm{X}} = \cos\frac{\theta}{2}\hat{I} - i\sin\frac{\theta}{2}\mathrm{X} = \begin{bmatrix} \cos\frac{\theta}{2} & -i\sin\frac{\theta}{2} \\ -i\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{bmatrix}, \tag{4.13a}$$

$$\hat{R}_y(\theta) = e^{-i\frac{\theta}{2}\mathrm{Y}} = \cos\frac{\theta}{2}\hat{I} - i\sin\frac{\theta}{2}\mathrm{Y} = \begin{bmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{bmatrix}, \tag{4.13b}$$

$$\hat{R}_z(\theta) = e^{-i\frac{\theta}{2}\mathrm{Z}} = \cos\frac{\theta}{2}\hat{I} - i\sin\frac{\theta}{2}\mathrm{Z} = \begin{bmatrix} e^{-i\frac{\theta}{2}} & 0 \\ 0 & e^{i\frac{\theta}{2}} \end{bmatrix}, \tag{4.13c}$$

where I have used Euler formulae (1.181) and on the last row Eqs. (2.34); note that, the z component is a generalisation of the phase gate (4.5). This allows us to write any unitary matrix in terms of the parameters $\alpha, \beta, \gamma, \delta \in \mathbb{R}$:

$$\hat{U}(\alpha, \beta, \gamma, \delta) = \begin{bmatrix} e^{i\left(\alpha-\frac{\beta}{2}-\frac{\delta}{2}\right)}\cos\frac{\gamma}{2} & -e^{i\left(\alpha-\frac{\beta}{2}+\frac{\delta}{2}\right)}\sin\frac{\gamma}{2} \\ e^{i\left(\alpha+\frac{\beta}{2}-\frac{\delta}{2}\right)}\sin\frac{\gamma}{2} & e^{i\left(\alpha+\frac{\beta}{2}+\frac{\delta}{2}\right)}\cos\frac{\gamma}{2} \end{bmatrix}, \tag{4.14}$$

as the rows and columns of a unitary matrix are orthonormal. In fact, any combination of unitary operator is itself unitary. For instance, we have

$$\left(\hat{U}_1\hat{U}_2\right)\left(\hat{U}_1\hat{U}_2\right)^\dagger = \hat{U}_1\hat{U}_2\hat{U}_2^\dagger\hat{U}_1^\dagger = \hat{U}_1\hat{U}_1^\dagger = \hat{I}, \tag{4.15a}$$

$$\left(\hat{U}_1\hat{U}_2\right)^\dagger\left(\hat{U}_1\hat{U}_2\right) = \hat{U}_2^\dagger\hat{U}_1^\dagger\hat{U}_1\hat{U}_2 = \hat{U}_2^\dagger\hat{U}_2 = \hat{I}, \tag{4.15b}$$

where I have made use of the rule (1.82d), valid for all operators, and generalisation to all cases of $n > 2$ unitary operators is self-evident. Taking into account the explicit forms (4.13), we have:

$$\hat{U}(\alpha, \beta, \gamma, \delta) = e^{i\alpha}\begin{bmatrix} e^{-i\frac{\beta}{2}} & 0 \\ 0 & e^{i\frac{\beta}{2}} \end{bmatrix}\begin{bmatrix} \cos\frac{\gamma}{2} & -\sin\frac{\gamma}{2} \\ \sin\frac{\gamma}{2} & \cos\frac{\gamma}{2} \end{bmatrix}\begin{bmatrix} e^{-i\frac{\delta}{2}} & 0 \\ 0 & e^{i\frac{\delta}{2}} \end{bmatrix}, \tag{4.16}$$

or, in short,

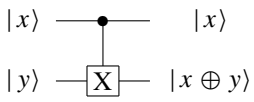
$$\hat{U}(\alpha, \beta, \gamma, \delta) = e^{i\alpha}\hat{R}_z(\beta)\hat{R}_y(\gamma)\hat{R}_z(\delta). \tag{4.17}$$

[18] See Nielsen and Chuang (2000, 174–77).

4.2.2 Generation of States

CNOT Gate

One of the most important two-qubit unitary gates is the controlled-NOT gate,[19] which can be represented by the circuit



(4.18)

The top line represents the *control* qubit, the bottom line the *target* qubit. The states $|x\rangle$ and $|y\rangle$ represent the computational basis states $|0\rangle$ and $|1\rangle$ and $\oplus$ denotes modulo-2 addition (3.203).[20] Note that, all gates here and in the following are represented by boxes while the computational trajectory of a single qubit is represented by a horizontal line (going from the left to the right). Controls or other interdependencies are represented by vertical lines. Control qubits are represented by dots. Note also that here and in the following circuits like (4.18) can be considered to be equivalents to equations (and this is why they are numbered together with them), and in fact, many proofs can be performed on that graphical basis only.

The CNOT gate leaves the control qubit unchanged and performs a X transformation on a target qubit depending on the state of the control qubit: if it is in the state $|0\rangle$, then it leaves the target qubit unchanged, but if it is in the state $|1\rangle$, then the quantum NOT gate is applied to the target qubit. Thus, using the current simplification of the notation CNOT $:= \hat{U}_{\text{CNOT}}$, we have

$$|00\rangle \overset{\text{CNOT}}{\longmapsto} |00\rangle\,, \quad |01\rangle \overset{\text{CNOT}}{\longmapsto} |01\rangle\,, \tag{4.19a}$$

$$|10\rangle \overset{\text{CNOT}}{\longmapsto} |11\rangle\,, \quad |11\rangle \overset{\text{CNOT}}{\longmapsto} |10\rangle\,, \tag{4.19b}$$

where the first qubit is the control qubit and the second the target qubit. In the computational basis of the two-qubit 4D Hilbert space $\mathcal{H}_1 \otimes \mathcal{H}_2$ (as a particular instance of (1.70) and an analogue of the basis (1.384))

$$|00\rangle := \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad |01\rangle := \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad |10\rangle := \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad |11\rangle := \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \tag{4.20}$$

the CNOT gate is defined in matrix form as

[19] Introduced in Barenco et al. (1995).

[20] In the literature a small circle is generally used for denoting the X gate (here depicted as a box). The reason for this choice here is that a qubit can control also other kinds of operations on a target qubit, so that controlling gate is a more general notion than the CNOT gate.

$$\text{CNOT} := \begin{bmatrix} 1\,0\,0\,0 \\ 0\,1\,0\,0 \\ 0\,0\,0\,1 \\ 0\,0\,1\,0 \end{bmatrix}. \tag{4.21}$$

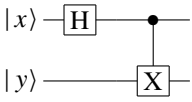
From the above expression, it is obvious that CNOT is both Hermitian and unitary (like the Hadamard or the Pauli gates). Since there can be diverse unitary operations controlled in the same way of a CNOT gate, as a convention, in all these cases, the unitary gate (e.g. a Hadamard transformation) acts only when the control qubit is 1.

Bell States

An important feature of the CNOT gate is its ability to entangle and disentangle states of a pair of qubits. Note that in the computational basis (4.20), the Bell states (3.170) will be given by

$$|\Psi^+\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\1\\1\\0\end{pmatrix}, \quad |\Psi^-\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\1\\-1\\0\end{pmatrix}, \quad |\Phi^+\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\0\\1\end{pmatrix}, \quad |\Phi^-\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\0\\-1\end{pmatrix}, \tag{4.22}$$

where I have dropped subscripts for the sake of generalisation. Let us now consider, for instance, a device composed of a Hadamard gate followed by a CNOT gate:



$$|\Psi^\pm\rangle, |\Phi^\pm\rangle \tag{4.23}$$

The output states are recognised, respectively, as the four Bell states $|\Phi^+\rangle$, $|\Psi^+\rangle$, $|\Phi^-\rangle$ and $|\Psi^-\rangle$, given by Eqs. (3.164) (see also Eqs. (3.164)), and therefore are maximally entangled states of the two qubits. The states $|x\rangle$ and $|y\rangle$ represent always the computational basis states $|0\rangle$ and $|1\rangle$. It is easy to see that the actions of this device on the unentangled product states $|00\rangle, |01\rangle, |10\rangle$ and $|11\rangle$ of the first (control) and second (target) qubits are, respectively, given by

$$|00\rangle \overset{\text{H}\otimes\hat{I}}{\longmapsto} \frac{1}{\sqrt{2}}(|00\rangle + |10\rangle) \overset{\text{CNOT}}{\longmapsto} \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle), \tag{4.24a}$$

$$|01\rangle \overset{\text{H}\otimes\hat{I}}{\longmapsto} \frac{1}{\sqrt{2}}(|01\rangle + |11\rangle) \overset{\text{CNOT}}{\longmapsto} \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle), \tag{4.24b}$$

$$|10\rangle \overset{\text{H}\otimes\hat{I}}{\longmapsto} \frac{1}{\sqrt{2}}(|00\rangle - |10\rangle) \overset{\text{CNOT}}{\longmapsto} \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle), \tag{4.24c}$$

$$|11\rangle \overset{\text{H}\otimes\hat{I}}{\longmapsto} \frac{1}{\sqrt{2}}(|01\rangle - |11\rangle) \overset{\text{CNOT}}{\longmapsto} \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle), \tag{4.24d}$$

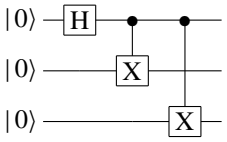
where $\mathrm{H} \otimes \hat{I}$ means a Hadamard transformation on the first qubit and identity on the second qubit. Since both H and CNOT are the inverse of themselves (i.e. they are Hermitian and unitary), the device can be also run backwards so as to disentangle the Bell states into unentangled product computational basis states. For example, we have

$$\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \overset{\text{CNOT}}{\longmapsto} \frac{1}{\sqrt{2}}(|00\rangle + |10\rangle) \overset{\mathrm{H}\otimes\hat{I}}{\longmapsto} |00\rangle \ . \tag{4.25}$$

In other words, when running backwards, this device also serves as a measurement device on each of the Bell states. This shows that coupling (entangling) and detecting are *inverse operations*, and it represents an evidence of an interesting complementarity between (detection) event and quantum correlations.

GHSZ States

Another combination of Hadamard and CNOT gates can generate the GHSZ state (3.233). It can be easily prepared using quantum gates through the circuit



$$\frac{1}{\sqrt{2}}(|000\rangle + |111\rangle) \tag{4.26a}$$

which is described by

$$\begin{aligned} |0\rangle\,|00\rangle \quad &\overset{\mathrm{H}_1}{\longmapsto} \quad \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)\,|00\rangle \\ &\overset{\text{CNOT}_{12}}{\longmapsto} \quad \frac{1}{\sqrt{2}}(|000\rangle + |110\rangle) \\ &\overset{\text{CNOT}_{13}}{\longmapsto} \quad \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle) \ , \end{aligned} \tag{4.26b}$$

where the Hadamard operation is on the first qubit and the targets of the CNOT transformations are, in succession, the second and the third qubit (the first figure denotes control while the second one the target qubit). Note that in the following, I shall also make use of a CNOT gate having the first qubit as target:

$$\text{CNOT}_{21} := \begin{bmatrix} 1\,0\,0\,0 \\ 0\,0\,0\,1 \\ 0\,0\,1\,0 \\ 0\,1\,0\,0 \end{bmatrix} . \tag{4.27}$$

Swap Gate

Another logic gate is the Swap gate that interchanges two qubits, whose matrix expression is defined as

$$\text{SWAP} := \begin{bmatrix} 1\,0\,0\,0 \\ 0\,0\,1\,0 \\ 0\,1\,0\,0 \\ 0\,0\,0\,1 \end{bmatrix}. \tag{4.28a}$$

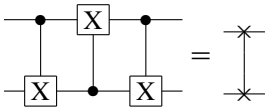
Note that, the SWAP gate acts as a sequence of CNOT, inverted CNOT, and CNOT gates:

$$\begin{aligned} |x, y\rangle &\overset{\text{CNOT}_{12}}{\longmapsto} |x, x \oplus y\rangle \\ &\overset{\text{CNOT}_{21}}{\longmapsto} |x \oplus (x \oplus y), x \oplus y\rangle = |y, x \oplus y\rangle \\ &\overset{\text{CNOT}_{12}}{\longmapsto} |y, (x \oplus y) \oplus y\rangle = |y, x\rangle\,, \end{aligned} \tag{4.28b}$$

where, due to rules (3.203), $\forall x, y = 0, 1$, we have

$$x \oplus x \oplus y = y \text{ and } x \oplus y \oplus y = x. \tag{4.28c}$$

Thus, it is represented by the circuit:



(4.28d)

4.2.3 *Relations Among Gates*

Phase-Shift Gates

The phase-shift gate (4.5) is related with other computational gates thanks to the generalisation (4.13c).[21] In fact, for $\phi = \pi/4$, we have the so-called $\pi/8$ gate or T gate

$$\text{T} := e^{\frac{\pi}{8}i} \begin{bmatrix} e^{-\frac{\pi}{8}i} & 0 \\ 0 & e^{\frac{\pi}{8}i} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & e^{\frac{\pi}{4}i} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{i} \end{bmatrix}, \tag{4.29}$$

where I have made use of the first Euler formulae (1.181), which for $\phi = \pi/2$ reduces to i. Now, the gate S that denotes the phase shifting of $\pi/2$ is

[21] On this subject, see Nielsen and Chuang (2000, Sect. 4.2).

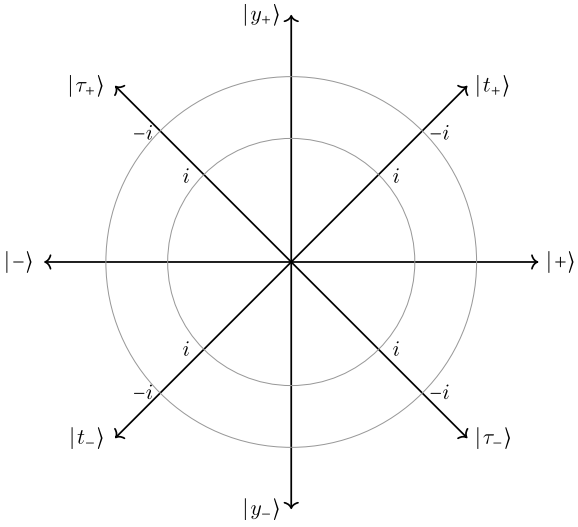


Fig. 4.5 Relative phase transformations on the x–y plane of the Poincaré sphere. In other words, the Poincaré sphere can be understood as formed by the intersection of the complex plan x-y with the real axis z. 90° CCW rotations give a factor i while 90° CW rotations a factor $-i$. We change the depicted reference axis by rotating of 45° either CCW (by $\sqrt{i}$) or CW (by $-\sqrt{i}$). Note that, the set $\{1, i, -1, -i\}$ constitutes a group (called C_4) with 1 as identity

$$\mathrm{S} := \mathrm{T}^2 = \begin{bmatrix} 1 & 0 \\ 0 & e^{\frac{\pi}{4}i} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & e^{\frac{\pi}{4}i} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix}. \tag{4.30}$$

Note in fact that the action of this gate is to rotate, for example states (3.82) into states (3.173), as well as the action of Gate T is to CCW rotate states (3.82) into

$$|t_+\rangle := \frac{1}{\sqrt{2}} \left(|0\rangle + \sqrt{i}\,|1\rangle \right) \text{ and } |t_-\rangle := \frac{1}{\sqrt{2}} \left(|0\rangle - \sqrt{i}\,|1\rangle \right), \tag{4.31}$$

as well as to CCW rotate States (3.173) into

$$|\tau_+\rangle := \frac{1}{\sqrt{2}} \left(|0\rangle + i\sqrt{i}\,|1\rangle \right) \text{ and } |\tau_+\rangle := \frac{1}{\sqrt{2}} \left(|0\rangle - i\sqrt{i}\,|1\rangle \right), \tag{4.32}$$

according to Fig. 4.5. Now, it is easy to see that the logic gate Z is in fact equal to S^2 and corresponds to a phase shift of π (the two components are completely out of phase (Figs. 1.4, Sect. 1.1.3, and 5.4, Sect. 5.1.2), as it happens for states (3.82)) and for this reason is also called (relative) *phase flip* (while the NOT gate X is called *bit flip*, since it interchanges the two basis elements $|0\rangle$ and $|1\rangle$). In fact, we have:

$$\mathrm{Z} = \mathrm{S}^2 = \begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{4.33}$$

Therefore, on the states (3.82a) and (3.82b) Z acts as

$$\mathrm{Z}\,|\pm\rangle = |\mp\rangle\,. \tag{4.34}$$

In other words, we have the sequence: $e^{\frac{\pi}{4}i} = \sqrt{i}$, $e^{\frac{\pi}{2}i} = i$, $e^{\pi i} = -1$, $e^{2\pi i} = 1$, satisfying the *Euler's identity*, after the name of L. Euler,

$$e^{\pi i} + 1 = 0, \tag{4.35}$$

which connects the constants π (fundamental for the circumference), the imaginary unit i, and the base of the natural logarithm e with the numbers 0 and 1. Of course, having $e^{2\pi i} = 1$, we go back to the initial position, since both exponential and logarithms, as trigonometric functions, are *multivalued* functions; for example, from $e^z = e^{z+2\pi i}$, with $z \in \mathbb{C}$, it follows that

$$\ln e^z = \ln e^{z+2\pi i}. \tag{4.36}$$

Similarly, from the definition (4.30) of the S gate, it follows that $\ln i = i\pi/2$, and since $e^{\ln i} = i$, we also have

$$i^i = e^{i \ln i} = e^{-\frac{\pi}{2}}. \tag{4.37}$$

Finally, it is interesting to note that the Y gate is given by

$$\mathrm{Y} = e^{\frac{\pi}{2}i}\mathrm{XZ} = i \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}. \tag{4.38}$$

Square Not

Another interesting circuit is represented by the whole interferometer set up

$$|x\rangle - \boxed{\mathrm{H}} - \boxed{\phi} - \boxed{\mathrm{H}} - \boxed{\phantom{\mathrm{H}}} \tag{4.39}$$

where the last gate represents the final detection. Detection destroys the qubit, and so closes the circuit. When considering the whole transformation that leads to the final state (2.31), this can be expressed in a compact matrix

$$\sqrt{\mathrm{NOT}} = \mathrm{H}\hat{U}_\phi\mathrm{H} = \frac{1}{2}\begin{bmatrix} 1+e^{i\phi} & 1-e^{i\phi} \\ 1-e^{i\phi} & 1+e^{i\phi} \end{bmatrix}, \tag{4.40}$$

where $|0\rangle$ replaces $|1\rangle$ and $|1\rangle$ replaces $|2\rangle$. Note that, the square-not gate is a unitary but not Hermitian matrix. When this gate is applied to states $|0\rangle$, $|1\rangle$ gives

$$\sqrt{\mathrm{NOT}}\,|0\rangle = \frac{1}{2}\begin{pmatrix} 1+e^{i\phi} \\ 1-e^{i\phi} \end{pmatrix}, \quad \sqrt{\mathrm{NOT}}\,|1\rangle = \frac{1}{2}\begin{pmatrix} 1-e^{i\phi} \\ 1+e^{i\phi} \end{pmatrix}, \tag{4.41}$$

which are represented by all superpositions on the equatorial plane of the Poincaré sphere (in other words, the Hadamard gate is only a particular case of the square-not gate).[22]

Toffoli Gate

The simplest three-qubit gate is the controlled-controlled-NOT (CCNOT) or quantum Toffoli gate. Its action on three qubits is that if the first two (control) qubits are in the state $|1\rangle$, then it applies a NOT on the third (target) qubit, otherwise the initial state of the three qubits remains unchanged. Therefore, we have

$$|000\rangle \overset{\text{CCNOT}}{\longmapsto} |000\rangle\,, \quad |001\rangle \overset{\text{CCNOT}}{\longmapsto} |001\rangle\,, \tag{4.42a}$$

$$|010\rangle \overset{\text{CCNOT}}{\longmapsto} |010\rangle\,, \quad |011\rangle \overset{\text{CCNOT}}{\longmapsto} |011\rangle\,, \tag{4.42b}$$

$$|100\rangle \overset{\text{CCNOT}}{\longmapsto} |100\rangle\,, \quad |101\rangle \overset{\text{CCNOT}}{\longmapsto} |101\rangle\,, \tag{4.42c}$$

$$|110\rangle \overset{\text{CCNOT}}{\longmapsto} |111\rangle\,, \quad |111\rangle \overset{\text{CCNOT}}{\longmapsto} |110\rangle\,, \tag{4.42d}$$

where the first two qubit are the control qubits and the third qubit is the target qubit. In the computational basis of the three-qubit ($2^3 = 8$ dimensional) Hilbert space (as a particular instance of (1.70))

$$|000\rangle := \begin{pmatrix} 1\\0\\0\\0\\0\\0\\0\\0 \end{pmatrix}, \quad |001\rangle := \begin{pmatrix} 0\\1\\0\\0\\0\\0\\0\\0 \end{pmatrix}, \quad |010\rangle := \begin{pmatrix} 0\\0\\1\\0\\0\\0\\0\\0 \end{pmatrix}, \quad |011\rangle := \begin{pmatrix} 0\\0\\0\\1\\0\\0\\0\\0 \end{pmatrix}, \tag{4.43a}$$

$$|100\rangle := \begin{pmatrix} 0\\0\\0\\0\\1\\0\\0\\0 \end{pmatrix}, \quad |101\rangle := \begin{pmatrix} 0\\0\\0\\0\\0\\1\\0\\0 \end{pmatrix}, \quad |110\rangle := \begin{pmatrix} 0\\0\\0\\0\\0\\0\\1\\0 \end{pmatrix}, \quad |111\rangle := \begin{pmatrix} 0\\0\\0\\0\\0\\0\\0\\1 \end{pmatrix}. \tag{4.43b}$$

the quantum Toffoli gate is defined in matrix form as

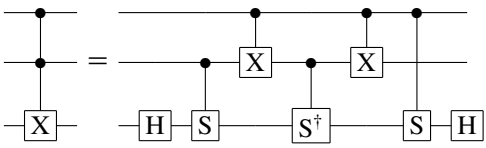
[22] Reference paper is Deutsch (1989).

$$\text{CCNOT} := \begin{bmatrix} 1&0&0&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0 \\ 0&0&1&0&0&0&0&0 \\ 0&0&0&1&0&0&0&0 \\ 0&0&0&0&1&0&0&0 \\ 0&0&0&0&0&1&0&0 \\ 0&0&0&0&0&0&0&1 \\ 0&0&0&0&0&0&1&0 \end{bmatrix}. \tag{4.44}$$

As the name suggests, the quantum Toffoli gate is the quantum analogue of the classical Toffoli gate introduced in Table 4.1.

Decomposition of Gates

Note that, any unitary gate of whatever dimension can be composed with a good approximation from just the CNOT gate and single-qubit (controlled) gates. For instance, the 8D CCNOT (Toffoli) gate on the left can be built by a combination of unitary operations shown on the right:



(4.45)

Note that, the network is symmetric apart from the S gate that is first controlled by qubit 2 and then by qubit 1. In explicit mapping, this gives:

$$\begin{array}{llll}
|000\rangle & \frac{1}{\sqrt{2}}(|000\rangle + |001\rangle) & \frac{1}{\sqrt{2}}(|000\rangle + |001\rangle) & \frac{1}{\sqrt{2}}(|000\rangle + |001\rangle) \\
|001\rangle & \frac{1}{\sqrt{2}}(|000\rangle - |001\rangle) & \frac{1}{\sqrt{2}}(|000\rangle - |001\rangle) & \frac{1}{\sqrt{2}}(|000\rangle - |001\rangle) \\
|010\rangle & \frac{1}{\sqrt{2}}(|010\rangle + |011\rangle) & \frac{1}{\sqrt{2}}(|010\rangle + |01\rangle\, i\, |1\rangle) & \frac{1}{\sqrt{2}}(|010\rangle + |01\rangle\, i\, |1\rangle) \\
|011\rangle \overset{\mathrm{H}_3}{\longmapsto} & \frac{1}{\sqrt{2}}(|010\rangle - |011\rangle) \overset{\mathrm{C_2S_3}}{\longmapsto} & \frac{1}{\sqrt{2}}(|010\rangle - |01\rangle\, i\, |1\rangle) \overset{\mathrm{CNOT}_{12}}{\longmapsto} & \frac{1}{\sqrt{2}}(|010\rangle - |01\rangle\, i\, |1\rangle) \\
|100\rangle & \frac{1}{\sqrt{2}}(|100\rangle + |101\rangle) & \frac{1}{\sqrt{2}}(|100\rangle + |101\rangle) & \frac{1}{\sqrt{2}}(|110\rangle + |111\rangle) \\
|101\rangle & \frac{1}{\sqrt{2}}(|100\rangle - |101\rangle) & \frac{1}{\sqrt{2}}(|100\rangle - |101\rangle) & \frac{1}{\sqrt{2}}(|110\rangle - |111\rangle) \\
|110\rangle & \frac{1}{\sqrt{2}}(|110\rangle + |111\rangle) & \frac{1}{\sqrt{2}}(|110\rangle + |11\rangle\, i\, |1\rangle) & \frac{1}{\sqrt{2}}(|100\rangle + |10\rangle\, i\, |1\rangle) \\
|111\rangle & \frac{1}{\sqrt{2}}(|110\rangle - |111\rangle) & \frac{1}{\sqrt{2}}(|110\rangle - |11\rangle\, i\, |1\rangle) & \frac{1}{\sqrt{2}}(|100\rangle + |10\rangle\, i\, |1\rangle)
\end{array}$$

$$
\begin{array}{l}
\frac{1}{\sqrt{2}}(|000\rangle + |001\rangle) \\
\frac{1}{\sqrt{2}}(|000\rangle - |001\rangle) \\
\frac{1}{\sqrt{2}}(|010\rangle + |011\rangle) \\
\xrightarrow{C_2S_3^\dagger} \quad \frac{1}{\sqrt{2}}(|010\rangle - |011\rangle) \\
\quad \frac{1}{\sqrt{2}}(|110\rangle + |11\rangle\,(-i\,|1\rangle)) \\
\quad \frac{1}{\sqrt{2}}(|110\rangle - |11\rangle\,(-i\,|1\rangle)) \\
\quad \frac{1}{\sqrt{2}}(|100\rangle + |10\rangle\,i\,|1\rangle) \\
\quad \frac{1}{\sqrt{2}}(|100\rangle - |10\rangle\,i\,|1\rangle)
\end{array}
\quad
\begin{array}{l}
\frac{1}{\sqrt{2}}(|000\rangle + |001\rangle) \\
\frac{1}{\sqrt{2}}(|000\rangle - |001\rangle) \\
\frac{1}{\sqrt{2}}(|010\rangle + |011\rangle) \\
\xrightarrow{CNOT_{12}} \quad \frac{1}{\sqrt{2}}(|010\rangle - |011\rangle) \\
\quad \frac{1}{\sqrt{2}}(|100\rangle + |10\rangle\,(-i\,|1\rangle)) \\
\quad \frac{1}{\sqrt{2}}(|100\rangle - |10\rangle\,(-i\,|1\rangle)) \\
\quad \frac{1}{\sqrt{2}}(|110\rangle + |11\rangle\,i\,|1\rangle) \\
\quad \frac{1}{\sqrt{2}}(|110\rangle - |11\rangle\,i\,|1\rangle)
\end{array}
\xrightarrow{C_1S_3}
$$

$$
\begin{array}{ll}
\frac{1}{\sqrt{2}}(|000\rangle + |001\rangle) & \frac{1}{2}(|000\rangle + |001\rangle + |000\rangle - |001\rangle) = |000\rangle \\
\frac{1}{\sqrt{2}}(|000\rangle - |001\rangle) & \frac{1}{2}(|000\rangle + |001\rangle - |000\rangle + |001\rangle) = |001\rangle \\
\frac{1}{\sqrt{2}}(|010\rangle + |011\rangle) & \frac{1}{2}(|010\rangle + |011\rangle + |010\rangle - |011\rangle) = |010\rangle \\
\frac{1}{\sqrt{2}}(|010\rangle - |011\rangle) & \xrightarrow{H_3} \frac{1}{2}(|010\rangle + |011\rangle - |010\rangle + |011\rangle) = |011\rangle \\
\frac{1}{\sqrt{2}}(|100\rangle + |101\rangle) & \frac{1}{2}(|100\rangle + |101\rangle + |100\rangle - |101\rangle) = |100\rangle \\
\frac{1}{\sqrt{2}}(|100\rangle - |101\rangle) & \frac{1}{2}(|100\rangle + |101\rangle - |100\rangle + |101\rangle) = |101\rangle \\
\frac{1}{\sqrt{2}}(|110\rangle + |11\rangle\,(-|1\rangle)) & \frac{1}{2}(|110\rangle + |111\rangle - |110\rangle + |111\rangle) = |111\rangle \\
\frac{1}{\sqrt{2}}(|110\rangle + |11\rangle\,(-|1\rangle)) & \frac{1}{2}(|110\rangle + |111\rangle + |110\rangle - |111\rangle) = |110\rangle
\end{array}
$$

I have used the convention that C_1S_3 denotes the gate S acting on qubit 3 and controlled by qubit 1, and similarly in other cases. Note that only for the latter two lines, there is an inversion of the truth value of the third qubit. This confirms that any quantum computation can be performed by combinations of (controlled) X,Y,Z gates in agreement with Eqs. (4.13)–(4.16).

4.3 Quantum Algorithms for Computing

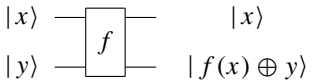
4.3.1 Deutsch's Algorithm

Allow me to give now some specific examples of the efficiency of quantum computation. I recall that a Boolean function (4.2) mapping the truth values $\{0, 1\}^j$ to the truth values $\{0, 1\}^k$ is called *constant* if it always returns the same value, i.e. either $f(0) = f(1) = 0$ or $f(0) = f(1) = 1$, and *balanced* if it returns 1 for half of the input domain and 0 for the other half. As an example of constant Boolean function consider, for whatever X, the disjunction (conjunction) with its negation, which is always true (false). As an example of balanced Boolean function consider, for whatever X, its XOR with whatever $Y \neq X, X'$.

Suppose now that there is a device that can evaluate the function f and that it is allowed to run only once. The so-called *Deutsch problem*, after the name of D. Deutsch, is to ask whether, under these conditions, it is possible to determine if the function f is constant or balanced.[23] It is easy to see that classically it is impossible to answer this question (at least two runs are needed). However, this can be done quantum-mechanically. Let us consider first the quantum implementation of a Boolean function as a quantum gate. For a given Boolean function f, its quantum implementation is the gate (which will be referred to as the Boolean gate) whose action is defined by the following unitary transformation on qubit $|x\rangle$:

$$|x\rangle\,|y\rangle \overset{\hat{U}_f}{\longmapsto} |x\rangle\,|f(x)\oplus y\rangle\,, \tag{4.46}$$

where $|x\rangle\,,|y\rangle = |0\rangle\,,|1\rangle$, and $\oplus$ is the modulo-2 addition (3.203). This is displayed in the circuit

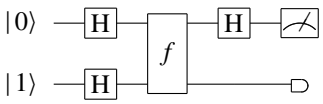


(4.47)

The power of a quantum device lies largely in the fact that due to linearity, we can input not just the computational basis states $|0\rangle$ and $|1\rangle$, but an arbitrary superposition of them. Thus, in its generality, the transformation $\hat{U}_f$ can be implemented by

$$\sum_j c_j \left|x_j\right\rangle |y\rangle \overset{\hat{U}_f}{\longmapsto} \sum_j c_j \left|x_j\right\rangle \left|f(x_j)\oplus y\right\rangle, \tag{4.48}$$

where $\left|x_j\right\rangle$ are arbitrary qubit states and c_j's $\in \mathbb{C}$ are coefficients satisfying the normalisation condition. This is *Deutsch algorithm*. It is interesting to note that in the above expression the state created by the Boolean gate is an entangled state. Now, consider the two-qubit quantum device



(4.49)

in which the input state is set to be $|01\rangle$. The D-shaped gate at the end of the second row denotes that we are discarding rather than measuring the second qubit: in fact, a measurement on the output state of the first qubit is sufficient for obtaining the requested answer. The two qubits are first transformed separately by the two Hadamard gates

$$|01\rangle \overset{\mathrm{H}\otimes\mathrm{H}}{\longmapsto} \mathrm{H}\,|0\rangle\,\mathrm{H}\,|1\rangle = \frac{1}{2}\left(|0\rangle + |1\rangle\right)\left(|0\rangle - |1\rangle\right). \tag{4.50}$$

[23] Deutsch (1985a, b).

For later convenience, the resultant state of Eq. (4.50) can be rewritten as

$$|\Psi\rangle = \frac{1}{2}\left(|00\rangle + |10\rangle - |01\rangle - |11\rangle\right). \tag{4.51}$$

The state $|\Psi\rangle$ is then processed by the Boolean gate, leading to four possible results depending on the nature of the function f. If f is a constant function, then we obtain

$$\begin{aligned}
\hat{U}_f(0,0) : |\Psi\rangle &\longmapsto \frac{1}{2}(|0\rangle\,|0\oplus 0\rangle + |1\rangle\,|0\oplus 0\rangle - |0\rangle\,|0\oplus 1\rangle - |1\rangle\,|0\oplus 1\rangle) \\
&= \frac{1}{2}[(|0\rangle + |1\rangle)\,|0\rangle - (|0\rangle + |1\rangle)\,|1\rangle] \\
&= \frac{1}{2}(|0\rangle + |1\rangle)(|0\rangle - |1\rangle) \\
&= |+\rangle\,|-\rangle\,, \qquad (4.52a) \\
\hat{U}_f(1,1) : |\Psi\rangle &\longmapsto \frac{1}{2}(|0\rangle\,|1\oplus 0\rangle + |1\rangle\,|1\oplus 0\rangle - |0\rangle\,|1\oplus 1\rangle - |1\rangle\,|1\oplus 1\rangle) \\
&= \frac{1}{2}\left(|01\rangle + |11\rangle - |00\rangle - |10\rangle\right) \\
&= -\frac{1}{2}(|0\rangle + |1\rangle)(|0\rangle - |1\rangle) \\
&= -|+\rangle\,|-\rangle\,, \qquad (4.52b)
\end{aligned}$$

where use has been made of the shorthand notation $\hat{U}_f(0,0) = \hat{U}_{f(0)=f(1)=0}$ and $\hat{U}_f(1,1) = \hat{U}_{f(0)=f(1)=1}$ for the first and second qubit. Similarly, if f is a balanced function, we have

$$\begin{aligned}
\hat{U}_f(0,1) : |\Psi\rangle &\longmapsto \frac{1}{2}(|0\rangle\,|0\oplus 0\rangle + |1\rangle\,|1\oplus 0\rangle - |0\rangle\,|0\oplus 1\rangle - |1\rangle\,|1\oplus 1\rangle) \\
&= \frac{1}{2}\left(|00\rangle + |11\rangle - |01\rangle - |10\rangle\right) \\
&= \frac{1}{2}(|0\rangle - |1\rangle)(|0\rangle - |1\rangle) \\
&= |-\rangle\,|-\rangle \qquad (4.52c) \\
\hat{U}_f(1,0) : |\Psi\rangle &\longmapsto \frac{1}{2}(|0\rangle\,|1\oplus 0\rangle + |1\rangle\,|0\oplus 0\rangle - |0\rangle\,|1\oplus 1\rangle - |1\rangle\,|0\oplus 1\rangle) \\
&= \frac{1}{2}\left(|01\rangle + |10\rangle - |00\rangle - |11\rangle\right) \\
&= -\frac{1}{2}(|0\rangle - |1\rangle)(|0\rangle - |1\rangle) \\
&= -|-\rangle\,|-\rangle\,, \qquad (4.52d)
\end{aligned}$$

where $\hat{U}_f(0,1) = \hat{U}_{f(0)=0,f(1)=1}$ and $\hat{U}_f(1,0) = \hat{U}_{f(0)=1,f(1)=0}$. According to circuit (4.49), the first qubit of each of the above four results (two for the constant case and two for the balanced case) is further processed by the final Hadamard gate, leading to the results

$$\pm|+\rangle\,|-\rangle \xmapsto{\mathrm{H}\otimes\hat{I}} \pm|0\rangle\,|-\rangle \quad \text{if } f \text{ is constant,} \tag{4.53a}$$

$$\pm|-\rangle\,|-\rangle \xmapsto{\mathrm{H}\otimes\hat{I}} \pm|1\rangle\,|-\rangle \quad \text{if } f \text{ is balanced.} \tag{4.53b}$$

This shows clearly that, apart from an irrelevant global phase factor (i.e. the overall sign of the output state), a measurement on the first qubit immediately gives the requested answer: the output $|0\rangle$ tells us that the function is constant, while the output $|1\rangle$ tells us that the function is balanced.

4.3.2 Shor's Theorem

The most known problem for which a fast algorithm is not classically known, is that of the factorisation of large integer numbers: the number of computational steps increases exponentially with the binary dimension of the input. Fortunately, the problem of factorising a number can be transformed in that of finding the period of a periodic function, and the latter problem can be solved quantum-mechanically. This is the important result of the *Shor's theorem*,[24] after the name of the American computer scientist Peter Shor: by (i) evaluating a wave function in a superposition of exponentially many arguments, each one representing a value of the requested periodic function, (ii) computing a parallel Fourier transform on the superposition, and finally (iii) sampling the Fourier power spectrum, one obtains the searched period (Fig. 4.6).

We choose a number $\xi \in \mathbb{R}$ at random and find a period r such that $\xi^r \equiv 1 \bmod N$, where N is the number to be factorized. The operation modulo N is a generalisation of the previous modulo 2 and in our case means that $\xi^r - 1$ needs to be a multiple of N. Now we choose a smooth number $q \in \mathbb{R}$ (a number with small prime factors) such that $N^2 < q < 2N^2$ and prepare the input state

$$|\psi_i\rangle = \frac{1}{\sqrt{q}} \sum_{x=0}^{q-1} |x, 0\rangle \ . \tag{4.54}$$

In other words, we are preparing a superposition with a number q of coefficients and equal $1/q$ probability for each possible outcome. From this state, with an opportune quantum-computational step, one obtains the entangled state

[24] See Shor (1994). See also Auletta et al. (2009, Sect. 17.8.2).

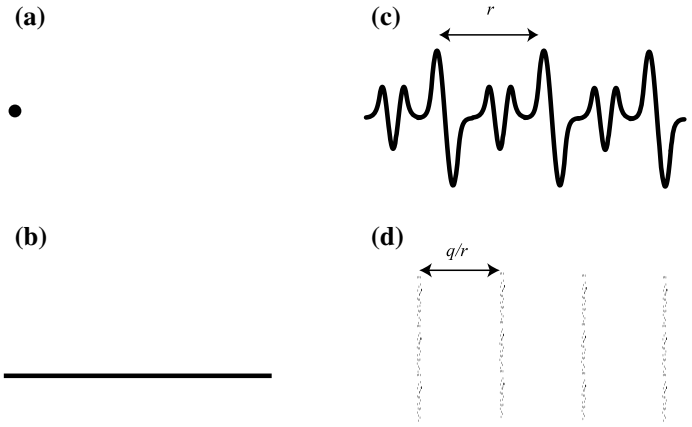


Fig. 4.6 Shor's superfast quantum Fourier sampling uses quantum interference to measure the period r of a periodic function $f(x)$. The period may be exponentially larger than the number of qubits involved in the computation. **a** The computer starts in the state $|x, f(x)\rangle = |0, 0\rangle$. **b** The x-register is put in a superposition of all possible values Eq. (4.54). **c** The value $f(x)$ is computed in the y-register simultaneously for all x values Eq. (4.55). **d** A Fourier transform of the x register into the y register is performed (see Eq. (4.57)). Measuring the Fourier transform of x yields a result $c = \lambda q/r$ from which the period r can be deduced (see Eq. (4.59)). Adapted from Auletta et al. (2009, p. 664)

$$\hat{U}\,|\psi_i\rangle \overset{\hat{U}}{\longmapsto} |\psi_f\rangle = \frac{1}{\sqrt{q}} \sum_{x=0}^{q-1} |x,\, f(x)\rangle \;, \tag{4.55}$$

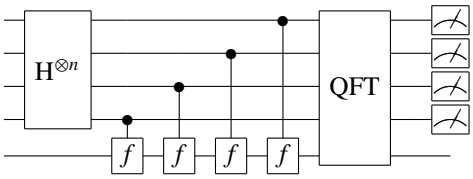
with $f(x) = \xi^x \mod N$. Now, we Fourier transform this pure state making use of the quantum or discrete Fourier transform (see Eqs. (1.175)–(1.177) for the continuous case)

$$\hat{U}_{\mathrm{QFT}} : |x\rangle \longmapsto \frac{1}{\sqrt{q}} \sum_{y=0}^{q-1} e^{\frac{2\pi i}{q} xy} |y\rangle \;. \tag{4.56}$$

Thus, by inserting the transformed state into Eq. (4.55), we obtain

$$|\tilde{\psi}_f\rangle = \frac{1}{q} \sum_{y=0}^{q-1} \sum_{x=0}^{q-1} e^{\frac{2\pi i}{q} xy} |y \;,\, f(x)\rangle. \tag{4.57}$$

This can be schematically summarised (for some steps and some qubits) by the following circuit (where the controlled f denotes the transformation (4.55)):



(4.58)

Each of the f's applications depend on some input x. It is interesting to note that classically one needs about $N \lg(N) = n2^n$ exponential steps to Fourier transform $N = 2^n$ discrete inputs (since $\log_b x^a = a \log_b x$), whereas on a quantum computer, the Fourier transform takes about $\log^2(N) = n^2$ polynomial steps. We can now measure both arguments of the superposition (4.57), obtaining a certain value c of y, and some $\xi^k \bmod N$ (k being any number between 0 and r) as value of $f(x)$. The probability for such a result will be

$$\wp(c, \xi^k) = \left| \frac{1}{q} \sum_{x=0}^{q-1}{}' e^{\frac{2\pi i}{q} xc} \right|^2 , \tag{4.59}$$

where the prime indicates a restricted sum over values which satisfy $\xi^x = \xi^k$. The probability $\wp(c, x^k)$ is periodic in c with period q/r. A measurement gives, with high probability, $c = \lambda q/r$—where λ is an integer corresponding to one of the peaks shown in Fig. 4.6c–d. But since we know q, we can determine r with few trials.

4.3.3 Grover's Algorithm

The Algorithm

The third example is provided by the Grover's algorithm (after the name of the Indian–American computer scientist Lov Grover) for searching in some databases. This algorithm makes use of an unitary operator O called oracle and of the oracle qubit $|o\rangle$.[25] If $|x\rangle$ represents a possible solution, we have the map

$$|x\rangle\,|o\rangle \overset{\mathrm{O}}{\longmapsto} |x\rangle\,|o \oplus f(x)\rangle\ , \tag{4.60}$$

where $f(x)$ represents the search function that takes values 1 or 0 depending on whether x is in fact a solution or not. Suppose that the oracle qubit starts in the state $|-\rangle$. If $|x\rangle$ is not a solution to the search problem, applying the oracle to the state $|x\rangle\,|-\rangle$ does not change the state, while if x is a solution then $|0\rangle$ and $|1\rangle$ are interchanged by the action of the oracle giving rise to the final state $-|x\rangle\,|-\rangle$. The action of the oracle can then be described as

[25] Grover (1996), Nielsen and Chuang (2000, Sect. 6.1).

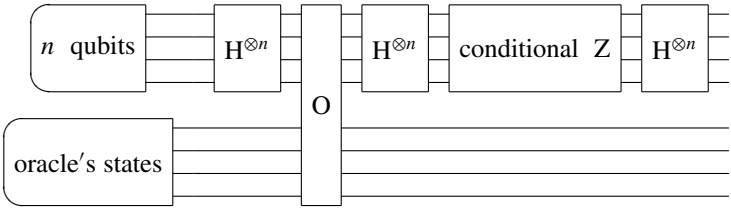
$$|x\rangle\,|-\rangle \overset{\mathrm{O}}{\longmapsto} (-1)^{f(x)}\,|x\rangle\,|-\rangle\,, \tag{4.61}$$

where it is noted that the state of the oracle remains unchanged and therefore can be omitted in the following. For solving an N item search problem with M solutions, it turns out that we need to apply the search oracles of the order of $\sqrt{N/M}$ times for obtaining a solution on a quantum circuit.

Let us suppose an n-qubit input state $|0\rangle^{\otimes n}$, with $n = \lg N$. We first Hadamard transform the state into the symmetric superposition state

$$|0\rangle^{\otimes n} \overset{\mathrm{H}^{\otimes n}}{\longmapsto} |\psi\rangle = \frac{1}{\sqrt{N}} \sum_{x=o}^{N-1} |x\rangle\,. \tag{4.62}$$

Having prepared the qubits in this state, we apply the quantum search algorithm $\hat{G}$ called the Grover's operator, displayed in the following circuit (the two D-shaped multigates on the left represent some preparation devices that gives the labeled qubits as inputs):



(4.63)

It consists of four steps (following the initial preparation and the first Hadamard transformation):

- Apply the oracle O:

$$|x\rangle \overset{\mathrm{O}}{\longmapsto} (-1)^{f(x)}\,|x\rangle\,. \tag{4.64}$$

- Apply again the Hadamard transform $\mathrm{H}^{\otimes n}$.
- Perform a conditional phase flip, with every computational basis state except $|0\rangle$ receiving a phase shift of -1

$$|x\rangle \longmapsto -(-1)^{\delta_{x0}}\,|x\rangle\,, \tag{4.65}$$

  where δ_{x0} is the delta function for $x, 0$.
- Apply a third time the Hadamard transform $\mathrm{H}^{\otimes n}$.

Since the phase shift (4.33) can be written as

$$\mathrm{Z} = 2\,|0\rangle\,\langle 0| - \hat{I}, \tag{4.66}$$

given that, according to Eq. (4.4), we have

$$2\,|0\rangle\,\langle 0| = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}, \tag{4.67}$$

the last three steps can be formulated as

$$\mathrm{H}^{\otimes n}\left(2\,|0\rangle\,\langle 0| - \hat{I}\right)\mathrm{H}^{\otimes n} = 2\,|\psi\rangle\,\langle\psi| - \hat{I}, \tag{4.68}$$

where $|\psi\rangle$ is given by Eq. (4.62). Thus, the Grover's algorithm can be simply written as a combination of the action of the oracle followed by the above operation:

$$\hat{G} = \left(2\,|\psi\rangle\,\langle\psi| - \hat{I}\right)\mathrm{O}. \tag{4.69}$$

A Geometric Representation of Grover's Algorithm

It is possible to build an insightful geometric representation.[26] Suppose that

$$\left|s'\right\rangle = \frac{1}{\sqrt{N-M}}\sum_{x}{}' \,|x\rangle \tag{4.70a}$$

represents a superposition restricted to the states $|x\rangle$'s that are not solutions to the search problem while

$$|s\rangle = \frac{1}{\sqrt{M}}\sum_{x}{}'' \,|x\rangle \tag{4.70b}$$

a superposition restricted to the states that are in fact solutions to the search problem. This allows to rewrite the state (4.62) as

$$|\psi\rangle = \sqrt{\frac{N-M}{N}}\left|s'\right\rangle + \sqrt{\frac{M}{N}}\,|s\rangle\ . \tag{4.71}$$

The effect of O can be understood as a reflection about the vector $\left|s'\right\rangle$ in the plane defined by $\left|s'\right\rangle$ and $|s\rangle$ (see Fig. 4.7) for arbitrary coefficients $c, c' \in \mathbb{C}$ satisfying the normalisation condition:

$$\mathrm{O}(c'\left|s'\right\rangle + c\,|s\rangle) = c'\left|s'\right\rangle - c\,|s\rangle\ . \tag{4.72}$$

Similarly, also the compound operation (4.68) performs a reflection on the same plane but about the vector $|\psi\rangle$. Now, it turns out that the application of two reflections along different directions is equivalent to a rotation. Let choose (see also Eq. (3.80))

[26] Nielsen and Chuang (2000, Sect. 6.1).

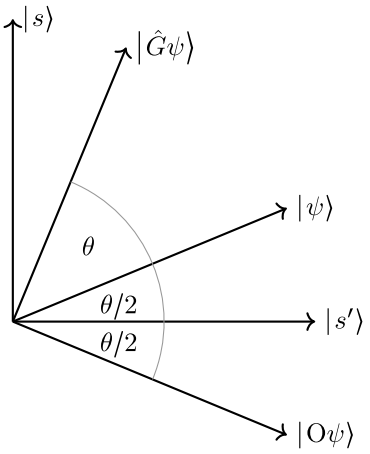


Fig. 4.7 Geometric representation of the Grover's algorithm: the vectors $|\psi\rangle$ and $\mathrm{O}|\psi\rangle$ make both an angle $\theta/2$ with $|s'\rangle$ while $\hat{G}|\psi\rangle$ an angle $3\theta/2$ with it. Adapted from Nielsen and Chuang (2000, p. 253)

$$\cos\theta = \sqrt{\frac{N-M}{N}}, \text{ so that } |\psi\rangle = \cos\frac{\theta}{2}|s'\rangle + \sin\frac{\theta}{2}|s\rangle\ . \tag{4.73}$$

Then, the two reflections corresponding to the rotation (4.69) performed by $\hat{G}$ take $|\psi\rangle$ to

$$\hat{G}|\psi\rangle = \cos\frac{3\theta}{2}|s'\rangle + \sin\frac{3\theta}{2}|s\rangle\ , \tag{4.74}$$

which is in fact a rotation of $|\psi\rangle$ by an angle θ, bringing it closer to the solution $|s\rangle$. After k iterations of the algorithm, the result comes closer and closer to the good solution $|s\rangle$ according to

$$\hat{G}^k|\psi\rangle = \cos\left(\frac{2k+1}{2}\theta\right)|s'\rangle + \sin\left(\frac{2k+1}{2}\theta\right)|s\rangle\ . \tag{4.75}$$

4.4 Heisenberg Picture and Teleportations

4.4.1 Heisenberg Picture for Computation

Clifford Group

It is possible to make use of the Heisenberg picture for computation (Sect. 2.1.1).[27] In other words, we shall consider not the evolution of qubits under the action of

[27] First studied in Gottesman (2008).

Table 4.2 The Clifford group. Since, according to Eq. (4.38), $Y = i\mathrm{XZ}$, we can always resolve its transformations in the following ones thanks to the map (4.78). It may be noted that X and Z behave opposite under the CNOT transformation

$\mathrm{SXS}^\dagger = \mathrm{Y}$	$\hat{U}_\phi \mathrm{Z} \hat{U}_\phi^\dagger = \mathrm{Z}$
$\mathrm{HXH} = \mathrm{Z}$	$\mathrm{HZH} = \mathrm{X}$
$\mathrm{CNOT}_{12}(\mathrm{X} \otimes \hat{I})\mathrm{CNOT}_{12} = \mathrm{X} \otimes \mathrm{X}$	$\mathrm{CNOT}_{12}(\mathrm{Z} \otimes \hat{I})\mathrm{CNOT}_{12} = \mathrm{Z} \otimes \hat{I}$
$\mathrm{CNOT}_{12}(\hat{I} \otimes \mathrm{X})\mathrm{CNOT}_{12} = \hat{I} \otimes \mathrm{X}$	$\mathrm{CNOT}_{12}(\hat{I} \otimes \mathrm{Z})\mathrm{CNOT}_{12} = \mathrm{Z} \otimes \mathrm{Z}$

gates but the evolution of other gates (operators) under such an action. We shall use a modification of the general formula (2.8a). Suppose to have an operation $\hat{O}$ and some unitary transformation $\hat{U}$. Since $\hat{U}^\dagger \hat{U} = \hat{I}$, we can write

$$\hat{U}\,\hat{O}\,|\psi\rangle = \hat{U}\,\hat{O}\hat{U}^\dagger \left(\hat{U}\,|\psi\rangle\right), \tag{4.76}$$

which allows us to have the map

$$\hat{O} \longmapsto \hat{U}\,\hat{O}\hat{U}^\dagger, \tag{4.77}$$

which is of the same form of (2.18). The reason is that the "evolved" operator $\hat{U}\,\hat{O}\hat{U}^\dagger$ acts on an already "evolved" state $\hat{U}\,|\psi\rangle$. In other words, what Eq. (4.76) tells us is that, after the transformation $\hat{U}$, the operator $\hat{U}\,\hat{O}\hat{U}^\dagger$ acts on the resulting state in just the same way as the operator $\hat{O}$ did on the initial state. Similarly in the case of multiplication (sequence) of operators (operations):

$$\hat{O}\hat{O}' \longmapsto \hat{U}\,\hat{O}\hat{O}'\hat{U}^\dagger = \left(\hat{U}\,\hat{O}\hat{U}^\dagger\right)\left(\hat{U}\,\hat{O}'\hat{U}^\dagger\right). \tag{4.78}$$

Let us now consider quantum gates acting on quantum gates. The gates H, S, and CNOT, through their actions on the Pauli gates, constitute a so-called Clifford group, after the name of the British mathematician William K. Clifford (1845–1879), as displayed in Table 4.2. The reiteration of any transformation in a Clifford group must either bring to the negative of the initial state or be identities.[28] In fact, apart from the action of the S gate on Z and of CNOT on both $\hat{I} \otimes \mathrm{X}$ and $\mathrm{Z} \otimes \hat{I}$, which all act as identity transformations, a reiteration of the operation performed by the S (and its adjoint) gate on X and Y from the left and right gives $-\mathrm{X}$ and $-\mathrm{Y}$, respectively, while reiteration of the H operation on X, Y, Z from the left and right acts as the identity operator and the same for the reiterated CNOT operation (each one always from left and right) on both $\mathrm{X} \otimes \hat{I}$ and $\hat{I} \otimes \mathrm{Z}$. Recall the expression (4.21) for the CNOT gate and take into account that we have

[28] see Penrose (2004, Sect. 11.5).

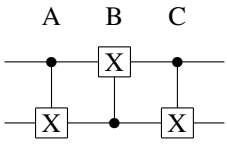
$$\hat{I}\otimes \mathrm{X}=\begin{bmatrix}0&1&0&0\\1&0&0&0\\0&0&0&1\\0&0&1&0\end{bmatrix},\quad \mathrm{X}\otimes\hat{I}=\begin{bmatrix}0&0&1&0\\0&0&0&1\\1&0&0&0\\0&1&0&0\end{bmatrix},\quad \mathrm{X}\otimes\mathrm{X}=\begin{bmatrix}0&0&0&1\\0&0&1&0\\0&1&0&0\\1&0&0&0\end{bmatrix},\tag{4.79}$$

$$\hat{I}\otimes \mathrm{Z}=\begin{bmatrix}1&0&0&0\\0&-1&0&0\\0&0&1&0\\0&0&0&-1\end{bmatrix},\ \mathrm{Z}\otimes\hat{I}=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&-1&0\\0&0&0&-1\end{bmatrix},\ \mathrm{Z}\otimes\mathrm{Z}=\begin{bmatrix}1&0&0&0\\0&-1&0&0\\0&0&-1&0\\0&0&0&1\end{bmatrix}.$$

I suggest the reader to check the validity of these matrices on some input qubits.

Some Examples

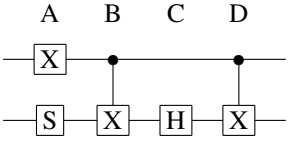
Consider the basis (4.20) and the circuit



(4.80)

The three operations are such that: A and C are CNOT_{12} while B is CNOT_{21} (see also circuit (4.28d)). It performs the transformations on the input gates displayed in Table 4.3, which are all transformations of the kind (4.77). Note that in each of these circuits a state entering as an eigenstate of the input operator remains the eigenstate with the same eigenvalue of the output operator. In fact, for the previous circuit (a SWAP operator), we have that $|00\rangle$ and $|11\rangle$ remain unchanged while $|01\rangle$ and $|10\rangle$ are interchanged.

Another interesting circuit is



(4.81)

Table 4.3 The transformations displayed in circuit (4.80). Consider that $|00\rangle$ and $|01\rangle$ are eigenstates of both $\hat{I}\otimes\mathrm{X}$ and $\hat{I}\otimes\mathrm{Z}$ with eigenvalue $+1$ as well as $|10\rangle$ and $|11\rangle$ are eigenstates of both $\mathrm{X}\otimes\hat{I}$ and $\mathrm{Z}\otimes\hat{I}$ again with eigenvalue $+1$

$\mathrm{X}\otimes\hat{I}$		$\mathrm{X}\otimes\mathrm{X}$		$\hat{I}\otimes\mathrm{X}$		$\hat{I}\otimes\mathrm{X}$
$\hat{I}\otimes\mathrm{X}$	A	$\hat{I}\otimes\mathrm{X}$	B	$\mathrm{X}\otimes\mathrm{X}$	C	$\mathrm{X}\otimes\hat{I}$
$\mathrm{Z}\otimes\hat{I}$	$\longmapsto$	$\mathrm{Z}\otimes\hat{I}$	$\longmapsto$	$\mathrm{Z}\otimes\mathrm{Z}$	$\longmapsto$	$\hat{I}\otimes\mathrm{Z}$
$\hat{I}\otimes\mathrm{Z}$		$\mathrm{Z}\otimes\mathrm{Z}$		$\mathrm{Z}\otimes\hat{I}$		$\mathrm{Z}\otimes\hat{I}$

Table 4.4 The transformations displayed in circuit (4.81)

$\mathrm{X}\otimes\hat{I}$	$\mathrm{Z}\otimes\hat{I}$	$\mathrm{Z}\otimes\hat{I}$	$\mathrm{Z}\otimes\hat{I}$	$\mathrm{Z}\otimes\hat{I}$
$\hat{I}\otimes\mathrm{X}$ $\overset{\mathrm{A}}{\longmapsto}$	$\hat{I}\otimes\mathrm{Y}$ $\overset{\mathrm{B}}{\longmapsto}$	$\mathrm{Z}\otimes\mathrm{Y}$ $\overset{\mathrm{C}}{\longmapsto}$	$-\mathrm{Z}\otimes\mathrm{Y}$ $\overset{\mathrm{D}}{\longmapsto}$	$-\hat{I}\otimes\mathrm{Y}$
$\mathrm{Z}\otimes\hat{I}$	$\mathrm{X}\otimes\hat{I}$	$\mathrm{X}\otimes\mathrm{X}$	$\mathrm{X}\otimes\mathrm{Z}$	$-\mathrm{Y}\otimes\mathrm{Y}$
$\hat{I}\otimes\mathrm{Z}$	$\hat{I}\otimes\mathrm{Z}$	$\mathrm{Z}\otimes\mathrm{Z}$	$\mathrm{Z}\otimes\mathrm{X}$	$\mathrm{Z}\otimes\mathrm{X}$

which performs the transformations displayed in Table 4.4.

4.4.2 Generalised Teleportation

We shall deal now again with teleportation for addressing a Heisenberg picture teleportation: teleportation of quantum logic gates.[29] Suppose that we have the usual qubit (3.168)

$$|\psi\rangle = c_0\,|0\rangle + c_1\,|1\rangle \tag{4.82}$$

to teleport, where I have dropped the subscript "1". In Sect. 3.3.6 we have considered $\left|\Psi^-\right\rangle$ Eq. (3.170a) to be the initial state of the entangled systems 2–3. In fact, teleportation can be performed with any vector of the Bell basis. For making this generalisation let us introduce the following formalism. We assume the state (3.170a) as basic for the systems 2–3 and write the other three by making use of the Pauli spin matrices:

$$\begin{aligned}
|\psi\rangle_1\otimes\left|\Psi^-\right\rangle_{23} = |\psi\rangle_1\otimes\left(\hat{I}\left|\Psi^-\right\rangle_{23}\right) &= -\frac{1}{2}\Big[\left|\Psi^-\right\rangle_{12}\otimes\left(\hat{I}\,|\psi\rangle_3\right) - \left|\Phi^-\right\rangle_{12}\otimes\left(\mathrm{X}\,|\psi\rangle_3\right)\\
&\quad + \left|\Phi^+\right\rangle_{12}\otimes\left(i\mathrm{Y}\,|\psi\rangle_3\right) + \left|\Psi^+\right\rangle_{12}\otimes\left(\mathrm{Z}\,|\psi\rangle_3\right)\Big],\\
|\psi\rangle_1\otimes\left|\Phi^-\right\rangle_{23} = |\psi\rangle_1\otimes\left(\mathrm{X}\left|\Psi^-\right\rangle_{23}\right) &= \frac{1}{2}\Big[\left|\Phi^-\right\rangle_{12}\otimes\left(\hat{I}\,|\psi\rangle_3\right) - \left|\Psi^-\right\rangle_{12}\otimes\left(\mathrm{X}\,|\psi\rangle_3\right)\\
&\quad - \left|\Psi^+\right\rangle_{12}\otimes\left(i\mathrm{Y}\,|\psi\rangle_3\right) + \left|\Phi^+\right\rangle_{12}\otimes\left(\mathrm{Z}\,|\psi\rangle_3\right)\Big],\\
|\psi\rangle_1\otimes\left|\Phi^+\right\rangle_{23} = |\psi\rangle_1\otimes\left(i\mathrm{Y}\left|\Psi^-\right\rangle_{23}\right) &= \frac{1}{2}\Big[\left|\Phi^+\right\rangle_{12}\otimes\left(\hat{I}\,|\psi\rangle_3\right) + \left|\Psi^+\right\rangle_{12}\otimes\left(\mathrm{X}\,|\psi\rangle_3\right)\\
&\quad + \left|\Psi^-\right\rangle_{12}\otimes\left(i\mathrm{Y}\,|\psi\rangle_3\right) + \left|\Phi^-\right\rangle_{12}\otimes\left(\mathrm{Z}\,|\psi\rangle_3\right)\Big],\\
|\psi\rangle_1\otimes\left|\Psi^+\right\rangle_{23} = |\psi\rangle_1\otimes\left(\mathrm{Z}\left|\Psi^-\right\rangle_{23}\right) &= -\frac{1}{2}\Big[\left|\Psi^+\right\rangle_{12}\otimes\left(\hat{I}\,|\psi\rangle_3\right) + \left|\Phi^+\right\rangle_{12}\otimes\left(\mathrm{X}\,|\psi\rangle_3\right)\\
&\quad + \left|\Phi^-\right\rangle_{12}\otimes\left(i\mathrm{Y}\,|\psi\rangle_3\right) + \left|\Psi^-\right\rangle_{12}\otimes\left(\mathrm{Z}\,|\psi\rangle_3\right)\Big],
\end{aligned} \tag{4.83}$$

where I have used the equalities (see also Eq. (4.38))

[29] Reference paper is Gottesman and Chuang (1999).

$$iXZ = -Y,\ iXY = -Z,\ iYZ = -X, \tag{4.84a}$$

$$iZX = Y,\ \ iYX = Z,\ \ iZY = X, \tag{4.84b}$$

in agreement with commutation relations (1.325c). Of course, we can choose any of the four Bell states as the standards and get the other 'representations' always thanks to the Pauli matrices. A similar procedure could also be applied to entanglements swapping.

Now, we choose for the initial state of qubits 2–3 the third one above, i.e. the state (3.170d):

$$\left|\Phi^+\right\rangle = \frac{1}{\sqrt{2}}\left(|00\rangle + |11\rangle\right), \tag{4.85}$$

which can be prepared as usually (see the first of Eqs. (4.24)):

$$|00\rangle_{23} \overset{H_2}{\longmapsto} |+\rangle_2 |0\rangle_3 \overset{CNOT_{23}}{\longmapsto} \frac{1}{\sqrt{2}}\left(|00\rangle + |11\rangle\right)_{23}. \tag{4.86}$$

Then, we proceed in explicit computational form as follows:

$$\begin{aligned} |\psi\rangle \left|\Phi^+\right\rangle \overset{CNOT_{12}}{\longmapsto} \quad & \frac{1}{\sqrt{2}}\left[c_0 |0\rangle \left(|00\rangle + |11\rangle\right) + c_1 |1\rangle \left(|10\rangle + |01\rangle\right)\right] \\ \overset{H_1}{\longmapsto} \quad & \frac{1}{2}\left[|00\rangle \left(c_0 |0\rangle + c_1 |1\rangle\right) + |01\rangle \left(c_0 |1\rangle + c_1 |0\rangle\right)\right. \\ & \left. + |10\rangle \left(c_0 |0\rangle - c_1 |1\rangle\right) + |11\rangle \left(c_0 |1\rangle - c_1 |0\rangle\right)\right]. \end{aligned} \tag{4.87}$$

Now, a measurement performed not on the Bell states but using the basis (4.20) for systems 1–2, whose results xy are random, leaves the qubit belonging to Bob in the state $\hat{U}_{xy}\left|\psi'\right\rangle$, where we have

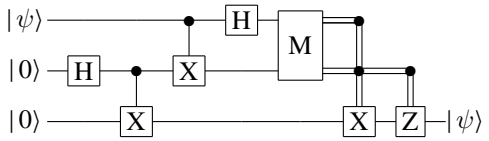
$$\hat{U}_{00} = \hat{I},\ \ i.e.\ \ \left|\psi'\right\rangle \longmapsto \left|\psi'\right\rangle, \tag{4.88}$$

$$\hat{U}_{01} = X,\ \ i.e.\ \ \left|\psi'\right\rangle \longmapsto \left|\psi' \oplus 1\right\rangle, \tag{4.89}$$

$$\hat{U}_{10} = Z,\ \ i.e.\ \ \left|\psi'\right\rangle \longmapsto (-1)^{\psi'} \left|\psi'\right\rangle, \tag{4.90}$$

$$\hat{U}_{11} = iY,\ \ i.e.\ \ \left|\psi'\right\rangle \longmapsto i(-1)^{\psi'} \left|\psi' \oplus 1\right\rangle. \tag{4.91}$$

The modulo-2 addition and exponentiation need to be performed on the two components of the qubit separately. Bob needs only to apply the inverse $\hat{U}_{xy}^{\dagger} = \hat{U}_{xy}$ of one of the previous transformations (according to Alice's classical communication) in order to recover $|\psi\rangle$:

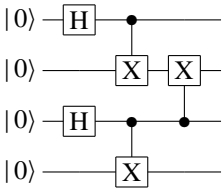


(4.92)

Practically, the chosen basis for systems 1–2 works as a 2-bit code, as anticipated in Sect. 3.3.6. The box M denotes the joint measurement on systems 1–2. Note that, double lines carry classical information (from Alice to Bob). The classical output of M on the first row is x while that of the second row is y. The final unitary operation of Bob (third row) is a controlled operation: if Alice gets $|01\rangle$, it is the second qubit to control the Z operation; if the Alice's result is $|10\rangle$, it is the first qubit to control the X gate; if it is $|11\rangle$, both qubits control the gates (I recall that $|00\rangle$ is mapped to the identity transformation). This confirms that such operations can be performed also without any human intervention (apart from the measurement).

4.4.3 Logic Gate Teleportation

With these tools we can now teleport a state together with a logic gate (that is called *logic gate teleportation*).[30] Suppose to have two EPR pairs (4.85) for qubits 1–2, 3–4. After creation of states $\left|\Phi^+\right\rangle_{12}$ and $\left|\Phi^+\right\rangle_{34}$, thanks to the double combination of H and CNOT gates on the left, let us pass qubits 2 and 3 a CNOT gate as displayed by the circuit



(4.93)

where a quadriparticle state is produced by the final CNOT_{32} gate:

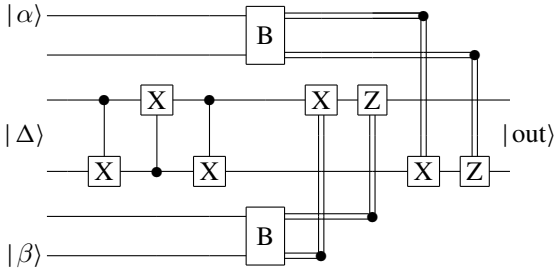
$$\begin{aligned}\left|\Phi^+\right\rangle_{12}\left|\Phi^+\right\rangle_{34} &\overset{\text{CNOT}_{32}}{\longmapsto} \frac{1}{2}\left[(|00\rangle + |11\rangle)_{12}|00\rangle_{34} + (|01\rangle + |10\rangle)_{12}|11\rangle_{34}\right] \\ &= \frac{1}{\sqrt{2}}\left(\left|\Phi^+\right\rangle_{12}|00\rangle_{34} + \left|\Psi^+\right\rangle_{12}|11\rangle_{34}\right) \\ &:= |\Delta\rangle\ . \end{aligned} \tag{4.94}$$

I recall that the CNOT gate acts as identity on the control qubit and as bit flip (X gate) on the target qubit depending on the component of the former. Let us now have two further inputs

[30] Gottesman and Chuang (1999).

$$|\alpha\rangle = a\,|0\rangle + b\,|1\rangle\,, \tag{4.95}$$
$$|\beta\rangle = c\,|0\rangle + d\,|1\rangle\,, \tag{4.96}$$

with $a, b, c, d \in \mathbb{C}$ and $|a|^2 + |b|^2 = |c|^2 + |d|^2 = 1$. Let us now consider the circuit



(4.97)

The preparation steps (4.86) and (4.94) already displayed in circuit (4.93), are not shown here. The central four qubits represent the state $|\Delta\rangle$. Bell measurement on qubits 1–2 gives as classical outputs x_1, y_1 (first and second row, respectively) while Bell measurement on qubits 5–6 gives x_2, y_2 (sixth and fifth row, respectively) as outputs (where the labels have been rearranged for the six-qubit input state). Suppose that we like to teleport the state

$$|\text{out}\rangle = \text{CNOT}_{34}\,|\beta\rangle_3\,|\alpha\rangle_4\,, \tag{4.98}$$

where $|\beta\rangle$ is the control qubit and $|\alpha\rangle$ is the target qubit. To this purpose, let us write the whole input state of the circuit

$$\begin{aligned}\frac{1}{2}\;|\alpha\rangle\,(|0000\rangle + |1100\rangle + |0111\rangle + |1011\rangle)\,|\beta\rangle = \\ \frac{1}{2}(ac\,|000000\rangle + ad\,|000001\rangle + ac\,|001110\rangle + ad\,|001111\rangle \\ + ac\,|011000\rangle + ad\,|011001\rangle + ac\,|010110\rangle + ad\,|010111\rangle \\ + bc\,|100000\rangle + bd\,|100001\rangle + bc\,|101110\rangle + bd\,|101111\rangle \\ + bc\,|111000\rangle + bd\,|111001\rangle + bc\,|110110\rangle + bd\,|110111\rangle).\end{aligned} \tag{4.99}$$

Now, the purpose is to rewrite this state in order that measurements on qubits 1–2 and 5–6 are performed on the Bell observable. Using the fact that the basis (4.20) can be expressed in terms of the Bell states

$$|00\rangle = \frac{1}{\sqrt{2}}\left(\left|\Phi^+\right\rangle + \left|\Phi^-\right\rangle\right),\quad |11\rangle = \frac{1}{\sqrt{2}}\left(\left|\Phi^+\right\rangle - \left|\Phi^-\right\rangle\right), \tag{4.100a}$$
$$|01\rangle = \frac{1}{\sqrt{2}}\left(\left|\Psi^+\right\rangle + \left|\Psi^-\right\rangle\right),\quad |10\rangle = \frac{1}{\sqrt{2}}\left(\left|\Psi^+\right\rangle - \left|\Psi^-\right\rangle\right), \tag{4.100b}$$

we can rewrite the input state as

$$
\begin{aligned}
\frac{1}{8}[\ & ac(\left|\Phi^+\right\rangle|00\rangle\left|\Phi^+\right\rangle+\left|\Phi^+\right\rangle|00\rangle\left|\Phi^-\right\rangle+\left|\Phi^-\right\rangle|00\rangle\left|\Phi^+\right\rangle+\left|\Phi^-\right\rangle|00\rangle\left|\Phi^-\right\rangle)\\
& ad(\left|\Phi^+\right\rangle|00\rangle\left|\Psi^+\right\rangle+\left|\Phi^+\right\rangle|00\rangle\left|\Psi^-\right\rangle+\left|\Phi^-\right\rangle|00\rangle\left|\Psi^+\right\rangle+\left|\Phi^-\right\rangle|00\rangle\left|\Psi^-\right\rangle)\\
& ac(\left|\Phi^+\right\rangle|11\rangle\left|\Psi^+\right\rangle-\left|\Phi^+\right\rangle|11\rangle\left|\Psi^-\right\rangle+\left|\Phi^-\right\rangle|11\rangle\left|\Psi^+\right\rangle-\left|\Phi^-\right\rangle|11\rangle\left|\Psi^-\right\rangle)\\
& ad(\left|\Phi^+\right\rangle|11\rangle\left|\Phi^+\right\rangle-\left|\Phi^+\right\rangle|11\rangle\left|\Phi^-\right\rangle+\left|\Phi^-\right\rangle|11\rangle\left|\Phi^+\right\rangle-\left|\Phi^-\right\rangle|11\rangle\left|\Phi^-\right\rangle)\\
& ac(\left|\Psi^+\right\rangle|01\rangle\left|\Psi^+\right\rangle-\left|\Psi^+\right\rangle|01\rangle\left|\Psi^-\right\rangle+\left|\Psi^-\right\rangle|01\rangle\left|\Psi^+\right\rangle-\left|\Psi^-\right\rangle|01\rangle\left|\Psi^-\right\rangle)\\
& ad(\left|\Psi^+\right\rangle|01\rangle\left|\Phi^+\right\rangle-\left|\Psi^+\right\rangle|01\rangle\left|\Phi^-\right\rangle+\left|\Psi^-\right\rangle|01\rangle\left|\Phi^+\right\rangle-\left|\Psi^-\right\rangle|01\rangle\left|\Phi^-\right\rangle)\\
& ac(\left|\Psi^+\right\rangle|10\rangle\left|\Phi^+\right\rangle+\left|\Psi^+\right\rangle|10\rangle\left|\Phi^-\right\rangle+\left|\Psi^-\right\rangle|10\rangle\left|\Phi^+\right\rangle+\left|\Psi^-\right\rangle|10\rangle\left|\Phi^-\right\rangle)\\
& ad(\left|\Psi^+\right\rangle|10\rangle\left|\Psi^+\right\rangle+\left|\Psi^+\right\rangle|10\rangle\left|\Psi^-\right\rangle+\left|\Psi^-\right\rangle|10\rangle\left|\Psi^+\right\rangle+\left|\Psi^-\right\rangle|10\rangle\left|\Psi^-\right\rangle)\\
& bc(\left|\Psi^+\right\rangle|00\rangle\left|\Phi^+\right\rangle+\left|\Psi^+\right\rangle|00\rangle\left|\Phi^-\right\rangle-\left|\Psi^-\right\rangle|00\rangle\left|\Phi^+\right\rangle-\left|\Psi^-\right\rangle|00\rangle\left|\Phi^-\right\rangle)\\
& bd(\left|\Psi^+\right\rangle|00\rangle\left|\Psi^+\right\rangle+\left|\Psi^+\right\rangle|00\rangle\left|\Psi^-\right\rangle-\left|\Psi^-\right\rangle|00\rangle\left|\Psi^+\right\rangle-\left|\Psi^-\right\rangle|00\rangle\left|\Psi^-\right\rangle)\\
& bc(\left|\Psi^+\right\rangle|11\rangle\left|\Psi^+\right\rangle-\left|\Psi^+\right\rangle|11\rangle\left|\Psi^-\right\rangle-\left|\Psi^-\right\rangle|11\rangle\left|\Psi^+\right\rangle+\left|\Psi^-\right\rangle|11\rangle\left|\Psi^-\right\rangle)\\
& bd(\left|\Psi^+\right\rangle|11\rangle\left|\Phi^+\right\rangle-\left|\Psi^+\right\rangle|11\rangle\left|\Phi^-\right\rangle-\left|\Psi^-\right\rangle|11\rangle\left|\Phi^+\right\rangle+\left|\Psi^-\right\rangle|11\rangle\left|\Phi^-\right\rangle)\\
& bc(\left|\Phi^+\right\rangle|01\rangle\left|\Psi^+\right\rangle-\left|\Phi^+\right\rangle|01\rangle\left|\Psi^-\right\rangle-\left|\Phi^-\right\rangle|01\rangle\left|\Psi^+\right\rangle+\left|\Phi^-\right\rangle|01\rangle\left|\Psi^-\right\rangle)\\
& bd(\left|\Phi^+\right\rangle|01\rangle\left|\Phi^+\right\rangle-\left|\Phi^+\right\rangle|01\rangle\left|\Phi^-\right\rangle-\left|\Phi^-\right\rangle|01\rangle\left|\Phi^+\right\rangle+\left|\Phi^-\right\rangle|01\rangle\left|\Phi^-\right\rangle)\\
& bc(\left|\Phi^+\right\rangle|10\rangle\left|\Phi^+\right\rangle+\left|\Phi^+\right\rangle|10\rangle\left|\Phi^-\right\rangle-\left|\Phi^-\right\rangle|10\rangle\left|\Phi^+\right\rangle-\left|\Phi^-\right\rangle|10\rangle\left|\Phi^-\right\rangle)\\
& bd(\left|\Phi^+\right\rangle|10\rangle\left|\Psi^+\right\rangle+\left|\Phi^+\right\rangle|10\rangle\left|\Psi^-\right\rangle-\left|\Phi^-\right\rangle|10\rangle\left|\Psi^+\right\rangle-\left|\Phi^-\right\rangle|10\rangle\left|\Psi^-\right\rangle)].
\end{aligned}
\tag{4.101}
$$

Now, let us collect the terms according to the possible measurement results for both qubit pairs, 1–2 and 5–6. The first four cases are

$$
\begin{aligned}
&\left|\Phi^+\right\rangle(ac\,|00\rangle+ad\,|11\rangle+bc\,|10\rangle+bd\,|01\rangle)\left|\Phi^+\right\rangle,\\
&\left|\Phi^+\right\rangle(ac\,|00\rangle-ad\,|11\rangle+bc\,|10\rangle-bd\,|01\rangle)\left|\Phi^-\right\rangle,\\
&\left|\Phi^-\right\rangle(ac\,|00\rangle+ad\,|11\rangle-bc\,|10\rangle-bd\,|01\rangle)\left|\Phi^+\right\rangle,\\
&\left|\Phi^-\right\rangle(ac\,|00\rangle-ad\,|11\rangle-bc\,|10\rangle+bd\,|01\rangle)\left|\Phi^-\right\rangle.
\end{aligned}
\tag{4.102}
$$

In other words, if the result of the two Bell measurements is $\left|\Phi^+\right\rangle,\left|\Phi^+\right\rangle$, we should get the four terms in the first row. Note that, the terms between brackets repeat on each row and only the signs (phases) change. Thus, for the other combinations we only need to write the first rows and indicate the signatures:

$$
\begin{aligned}
&\left|\Phi^+\right\rangle(ac\,|11\rangle+ad\,|00\rangle+bc\,|01\rangle+bd\,|10\rangle)\left|\Psi^+\right\rangle,\\
&\left|\Psi^+\right\rangle(ac\,|10\rangle+ad\,|01\rangle+bc\,|00\rangle+bd\,|11\rangle)\left|\Phi^+\right\rangle,\\
&\left|\Psi^+\right\rangle(ac\,|01\rangle+ad\,|10\rangle+bc\,|11\rangle+bd\,|00\rangle)\left|\Psi^+\right\rangle,
\end{aligned}
\tag{4.103}
$$

where the remaining signatures are $-++-$, $++--$, $-++-$ for the second case, $+-+-$, $++--$, $+--+$ for the third case and $-+-+$, $++--$, $-++-$ for

the fourth and last case. Note that measurement on qubits 1–2 instructs qubit 4 while measurement on qubits 5–6 instructs qubit 3, as it is also evident from the above circuit. Now, the standard teleportation output should be

$$|\beta\rangle_3\,|\alpha\rangle_4 = (c\,|0\rangle + d\,|1\rangle)(a\,|0\rangle + b\,|1\rangle) = ac\,|00\rangle + ad\,|10\rangle + bc\,|01\rangle + bd\,|11\rangle\,, \tag{4.104}$$

what means that teleporting the quantum gate CNOT_{34}, according to Eq. (4.98), gives

$$ac\,|00\rangle + ad\,|11\rangle + bc\,|01\rangle + bd\,|10\rangle \tag{4.105}$$

as output. In fact, applying CNOT_{34} to this state we get the standard output. To this purpose, before the two detections, the six-qubit state need to undergo the SWAP operation (4.28) displayed on the left of circuit (4.97). Then, again taking no care of the phase differences, we have for the four cases:

$$\begin{aligned}
ac\,|00\rangle + ad\,|11\rangle + bc\,|10\rangle + bd\,|01\rangle &\overset{\text{SWAP}}{\longmapsto} ac\,|00\rangle + ad\,|11\rangle + bc\,|01\rangle + bd\,|10\rangle\,,\\
ac\,|11\rangle + ad\,|00\rangle + bc\,|01\rangle + bd\,|10\rangle &\overset{\text{SWAP}}{\longmapsto} ac\,|11\rangle + ad\,|00\rangle + bc\,|10\rangle + bd\,|01\rangle\,,\\
ac\,|10\rangle + ad\,|01\rangle + bc\,|00\rangle + bd\,|11\rangle &\overset{\text{SWAP}}{\longmapsto} ac\,|01\rangle + ad\,|10\rangle + bc\,|00\rangle + bd\,|11\rangle\,,\\
ac\,|01\rangle + ad\,|10\rangle + bc\,|11\rangle + bd\,|00\rangle &\overset{\text{SWAP}}{\longmapsto} ac\,|10\rangle + ad\,|01\rangle + bc\,|11\rangle + bd\,|00\rangle\,.
\end{aligned}$$

It is easy to see that the unitary operations that need to be performed by Bob for getting the output (4.105) are (again, I shall not consider Z and thus I shall only take X and the X-component of Y into account)

- Identity on both qubits 3 and 4 for the first case,
- X operation on both qubits 3 and 4 for the second case,
- X operation on qubit 4 (and identity on qubit 3) for the third case, and
- X operation on qubit 3 (and identity on qubit 4) for the fourth case.

In this way, the receiver, in analogy with the previous case, would get the state $\hat{U}_{x_1y_1}\hat{U}_{x_2y_2}\,|\beta\alpha\rangle$, where the two unitary operators are Pauli gates. However, having replaced the usual EPR pair by a swapped $|\Delta\rangle$, Bob should receive a state of the form

$$\mathrm{CNOT}_{34}\;\hat{U}_{x_1y_1}\hat{U}_{x_2y_2}\,|\beta\alpha\rangle\,. \tag{4.106}$$

This appears to be weird since the unitary operations $\hat{U}_{x_1y_1}\hat{U}_{x_2y_2}$, being the final corrections of the output, should be performed at the end of a teleportation protocol, and thus after the CNOT operation. However, thanks to transformation

$$\mathrm{CNOT}\,\hat{U} \longmapsto \hat{U}\,\mathrm{CNOT}\,\hat{U}\hat{U}^\dagger = (\hat{U}\,\mathrm{CNOT}\,\hat{U}^\dagger)\hat{U} = \hat{U}\,\mathrm{CNOT}, \tag{4.107}$$

we can also interchange this operation with those performed by Bob and write

$$\mathrm{CNOT}_{34}\;\hat{U}_{x_1y_1}\hat{U}_{x_2y_2} = \hat{U}_{x_1y_1}\hat{U}_{x_2y_2}\;\mathrm{CNOT}_{34}, \tag{4.108}$$

what means that the CNOT gate when commuted with a Pauli gate produces again a Pauli gate. Thus, as displayed in circuit (4.97), the unitary correction operations can be anticipated and Bob gets the result $\text{CNOT}_{34}\ |\beta\alpha\rangle$.

4.5 Energy and Time in Quantum Computation

Let us now come back to the issue of the relation information-energy (and time) introduced in Sect. 4.1. Although energy does not play a central role in quantum computation, it becomes a crucial quantity when we are interested in calculating the time of computation. The reason is that energy and time are connected by the uncertainty relation (1.292). The Canadian–American physicist and computer scientist Norman Margolus,[31] remarked that, since classically kinetic energy is a measure of how much motion a system has, for a classical system with energy E and lowest possible energy E_0, the most energy that could possibly be changed into kinetic energy is given by $E - E_0$. This is the energy that can actually be turned into motion of particles. When we introduce the notion of the maximum speed of state change for computation, it turns out that its average is also a measure of motion. However, in CM one has to do with both particle motion and 'potential' motion, whereas, in considering computational systems, all 'motion' is state change. Nevertheless, for isolated quantum systems (not subject to fields or forces of whatever size), what is a requirement for reversibility, the two notions coincide.

Since quantum systems perform information processing, it has been calculated that the minimal time that a state takes to evolve into one of its orthogonal states (one of the other state vectors of an orthonormal basis denoting a different state of computation) is given by

$$\tau_{\perp} = \frac{h}{4E}, \tag{4.109}$$

where h is the Planck constant and E the energy.[32] In fact, recalling Eqs. (1.198)–(1.199), assume that the initial state of the system at time t_0 is a discrete superposition of energy eigenstates (which here I denote with $|E_n\rangle$)

$$|\psi(t_0)\rangle = \sum_n c_n |E_n\rangle\ , \tag{4.110}$$

and at a later time t it evolves into

$$|\psi(t)\rangle = \sum_n c_n e^{-i\frac{E_n t}{\hbar}} |E_n\rangle \tag{4.111}$$

thanks to the unitary transformation (1.16). Let us now introduce the amplitude

[31] Margolus (2003).

[32] Reference paper is Margolus and Levitin (1998).

$$\vartheta(t) = \langle \psi(t) \mid \psi(t_0) \rangle = \sum_{n=0}^{\infty} |c_n|^2 \, e^{-i \frac{E_n t}{\hbar}} \tag{4.112}$$

denoting the overlap between initial and evolved states. Now, we look for the smallest value of t such that $\vartheta(t) = 0$, i.e. for which the two states are orthogonal. Now, using the second of the trigonometric Euler formulae (1.181), we get for the real part Eqs. (1.9)

$$\begin{aligned} \Re(\vartheta) &= \sum_{n=0}^{\infty} |c_n|^2 \cos\left(\frac{E_n t}{\hbar}\right) \\ &\geq \sum_{n=0}^{\infty} |c_n|^2 \left[1 - \frac{2}{\pi}\left(\frac{E_n t}{\hbar} + \sin\left(\frac{E_n t}{\hbar}\right)\right)\right] \\ &= 1 - \frac{2}{\pi}\frac{Et}{\hbar} + \frac{2}{\pi}\Im(\vartheta), \end{aligned} \tag{4.113}$$

where the second line is derived thanks to the mathematical inequality $\forall x$

$$\cos x \geq 1 - \frac{2}{\pi}(x + \sin x), \tag{4.114}$$

and the last line by noting that $\sum_{n=0}^{\infty} E_n = E$ (i.e. the sum on all eigenvalues of the energy is the energy). Obviously, for any value of t such that $\vartheta(t) = 0$, we have $\Re(\vartheta) = \Im(\vartheta) = 0$ and Eq. (4.113) becomes

$$0 \geq 1 - \frac{4Et}{h}, \tag{4.115}$$

so that the minimal time amount $t = \tau_\perp$ is given by Formula (4.109). In conclusion, the maximum number of orthogonal states which a system can pass through in a certain time interval corresponds to the maximal number of operations that a computer can perform in that time.

4.6 Summary

- Quantum systems can be understood basically as computers performing reversible computation as long as no information selection is performed. However, *computation* appears to be a larger concept than information processing since it includes selection operations.
- Several kinds of reversible quantum gates have been introduced.
- The fundamental quantum algorithms are Deutsch's algorithm for solving with a single computation run the problem of determining whether a Boolean function

is constant or balanced, Shor's algorithm, especially efficient for factorisation of large numbers, and Grover's algorithm, especially suitable for search in databases.
- It is possible to teleport not only quantum states but also quantum gates.
- Time and energy of quantum computation have been calculated.

References

Auletta, Gennaro. 2006a. The Problem of Information. In: [Auletta 2006b, 109–127].

Auletta, Gennaro. 2006b. *The Controversial Relations Between Science and Philosophy: A Critical Assessment*. Vatican City: Libreria Editrice Vaticana.

Auletta, G., M. Fortunato, and G. Parisi. 2009. *Quantum Mechanics*. Cambridge: Cambridge University Press; rev. paperback ed. 2013.

Auletta, G. and S.-Y. Wang. 2014. *Quantum Mechanics for Thinkers*. Singapore: Pan Stanford Publishing. https://doi.org/10.4032/9789814411721.

Barenco, A., D. Deutsch, A. Ekert and R. Jozsa. 1995. Conditional Quantum Physics and Logic Gates. *Physical Review Letters* 74: 4083–4086.

Bennett, Charles H. 1973. Logical Reversibility of Computation. *IBM Journal of Research and Development* 17: 525–32.

Bennett, C.H. and R. Landauer. 1985. The Fundamental Physical Limits of Computation. *Scientific American* 253.1: 48–56. https://doi.org/10.1038/scientificamerican0785-48.

Chaitin, Gregory J. 1975 Randomness and Mathematical Proof. *Scientific American* 232.5: 47–52.

Chaitin, G.J., N. da Costa, and F.A. Doria. 2011. *Gödel's Way: Exploits into an Undecidable World*. United Kingdom: Taylor and Francis.

Church, Alonzo. 1936a. An Unsolvable Problem of Elementary Number Theory. *American Journal of Mathematics* 58: 345–363.

Church, Alonzo. 1936b. A Note on the Entscheidungsproblem. *Journal of Symbolic Logic* 1: 40–41. https://doi.org/10.2307/2269326.

Church, Alonzo. 1937. Review: A. M. Turing, On Computable Numbers, with an Application to the Entscheidungsproblem. *Journal of Symbolic Logic* 2: 42–43. https://doi.org/10.2307/2268810.

Deutsch, David. 1985a. Quantum Theory as a Universal Physical Theory. *International Journal of Theoretical Physics* 24: 1–41.

Deutsch, David. 1985b. Quantum Theory, the Church-Turing Principle and the Universal Quantum Computer. *Proceedings of the Royal Society of London* A400: 97–117.

Deutsch, David. 1989. Quantum Computational Networks. *Proceedings of the Royal Society of London* A425: 73–90.

Deutsch, David. 1997. *The Fabric of Reality: The Science of Parallel Universes and its Implications*. London: Penguin Books.

Deutsch, David. 2011. *The Beginning of Infinity: Explanations that Transform the World*. London: Penguin Books.

Feynman, Richard P. 1982. Simulating Physics with Computers. *International Journal of Theoretical Physics* 21: 467–488.

Feynman, Richard P. 1986. Quantum Mechanical Computers. *Foundations of Physics* 16: 507–531.

Feynman, R.P., and A.R. Hibbs. 2005. *Quantum Mechanics and Path Integrals*, New York: McGraw-Hill, 1965; Emended edition by Daniel F. New York, Dover: Styer.

Fredkin, E. and T. Toffoli. 1982. Conservative Logic. *International Journal of Theoretical Physics* 21: 219–253.

Gottesman, Daniel. 2008. The Heisenberg Representation of Quantum Computers. http://arxiv.org/quant-ph/9807006v1.

Gottesman, D. and I.L. Chuang. 1999. Quantum Teleportation is a Universal Computational Primitive. *Nature* 402: 390–393. https://doi.org/10.1038/46503.
Grover, Lov K. 1996. A Fast Quantum Mechanical Algorithm for Database Searching. In *Proceedings STOC 1996*, 212–219. Philadelphia.
Landauer, Rolf. 1961. Irreversibility and Heat Generation in the Computing Process. *IBM Journal of Research and Development* 5: 183–191.
Lloyd, Seth. 1996. The Physical Nature of Information. *Physics Letters* 217A: 188–193.
Lloyd, Seth. 2000. Ultimate Physical Limits to Computation. *Nature* 406: 1047–1054. https://doi.org/10.1038/35023282.
Lloyd, Seth. 2002. Computational Capacity of the Universe. *Physical Review Letters* 88: 237901.
Lloyd, Seth. 2006. *Programming the Universe: A Quantum Computer Scientist Takes on the Cosmos*. New York: A. A. Kopf.
Lloyd, S. and Y.J. Ng. 2004. Black Hole Computers. *Scientific American* 291.5: 52–61. https://doi.org/10.1038/scientificamerican1104-52.
Margolus, Norman. 2003. Looking at Nature as a Computer. *International Journal of Theoretical Physics* 42: 309–327.
Margolus, N. and L.B. Levitin. 1998. The Maximum Speed of Dynamical Evolution. *Physica* D120: 188–195.
Nielsen, M.A., and I.L. Chuang. 2000. *Quantum Computation and Quantum Information*. Cambridge: Cambridge University Press 2000, 2002, 2011.
Penrose, Roger. 2004. *The Road to Reality: A Complete Guide to the Laws of the Universe*. London: J. Cape, Vintage.
Shor, Peter W. 1994. *Algorithms for Quantum Computation: Discrete Log and Factoring*. In *Proceedings of the 35th Annual Symposium on the Foundations of Computer Science*, 124. Los Alamos: IEEE Computer Society Press.
Timpson, Christopher G. 2013. *Quantum Information Theory and the Foundations of Quantum Mechanics*. Oxford: Oxford University Press.
Turing, Alan M. 1936. On Computable Numbers, with an Application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society* 42: 230–265.
Turing, Alan M. 1937. On Computable Numbers, with an Application to the Entscheidungsproblem: A Correction. *Proceedings of the London Mathematical Society* 43: 544–546.
Wheeler, John A. 1990. Information, Physics, Quantum: The Search for Links. In: [Zurek 1990, 3–28].
Zurek, Wojciech H. 1990. *Complexity, Entropy and the Physics of Information*. Redwood City: Addison-Wesley.

Chapter 5
Ontological Ascription and Operations



In the previous chapter, we have deepened some formal issues developed in Chap. 3. Now, we shall deepen the related conceptual issues. I have stressed that notions like causal constraints and potentiality need to be effective in solving our problems in interpreting QM. In general, I have piecewise proposed a schematic interpretational framework that needs now to be constructively developed, improved and tested. In the first section, a preliminary examination of the concepts of quantum events and quantum features is developed, two notions that are problematic to deal with in a classical framework. The crucial notion of quantum discord will allow us to understand the distinction between quantum and classical correlations (and information). A detailed analysis of the measurement process, with its three basic steps (preparation, premeasurement and detection), follows. Finally, it is shown that Einstein's, Bohr's, Schrödinger's, and Heisenberg's interpretations converge and can be integrated in a view that corrects some overstress (mostly to attribute to the scarcity of experimental evidence at that time and to the work of their successors and interpreters) on particular aspects of their proposals.

The leitmotiv of this chapter could be taken from the famous book of Wheeler et al. on gravitation: theory, concept, law and method of measurement are born into the world in union.[1] Let us explore these interconnections.

5.1 Quantum Events and Features

5.1.1 Quantum Events

Experience and Its Interpretation

Almost nobody doubts about the reality of events like detections (Sect. 2.4.2). In fact, this would comport the rejection of the whole of experimental physics. In the first

[1] Misner et al. (1970, p. 71).

© Springer Nature Switzerland AG 2019

G. Auletta, *The Quantum Mechanics Conundrum*,
https://doi.org/10.1007/978-3-030-16649-6_5

approximation, an event is in fact what in principle can be observed by everybody.[2] A click of a detector is thus considered as an undeniable reality and so also our perception or experience of it, although, as we have seen, the meaning and range of such experience can be interpreted in different ways. When we raise the issue of the role and meaning of experience, there are several points to be scrutinised. Before dealing with an analysis of this concept, preliminary issue concerns precisely the general role of interpretation in making experience. This appears to be a crucial problem as far as we are dealing with the interpretation of QM. As it is evident when considering the history of science, many experiences that in the past appeared beyond any doubt (not only phenomenologically but also in a scientific framework) have been subsequently shown to be based on some mistake (like the perception of the absence of motion of the Earth[3] and of time as an absolute and universal parameter). So, some scholars could be tempted to delegitimate at least ordinary experience as an instance for judging about scientific theories and their implications (and in fact this is also what currently happens). The problem here is that our single thread with reality is represented ultimately by experience, and *controlled* experience, although certainly more reliable, necessarily shadows in *ordinary* experience and ultimately has its justification in it, as Bohr well understood. In fact, there are always steps in a controlled experience that are in fact ordinary experience, like the observer looking at the graduated scale of some measuring apparatus (and any attempt at reporting such steps under control will always leave a residue of ordinary experience). So, it would be better to carefully distinguish between *experiences* as such and the way in which we can *falsely interpret* them (what in fact happens also in controlled contexts).[4] For instance, we feel the Earth as at rest, but this is due to the conservation of motion and the effect of the Earth's gravitation that make happen everything on its surface as it would be at rest. There is nothing wrong with this experience. Moreover, we feel time as universal because, at our scale, relativistic effects are insignificant. Again, there is nothing wrong with this experience. In both cases, scientific theories are able to explain *why* we make such experiences and interpret them in a way that we can say to be not wrong as such but simply made under a certain (limited) perspective. In fact, in many cases we still make use of the "old" (pre-Copernican and pre-relativistic) physics when those effects are in fact negligible.

Thus, the epistemological principle that I would propose here is to (i) never deny experience, nor its interpretation if not in the presence of specific and cogent facts known otherwise that put into discussion our current interpretation of it, and (ii) when we do so, we should always be able to explain, in the new theoretical framework, why we were led to those perspective-like or partial interpretations. This makes also evident that interpretation, although fallible, is congruently connected (and corrected) with (by) the progress of scientific knowledge, and so it constitutes a fundamental aspect of the latter.

---

[2] Geroch (1978, p. 37).

[3] On the Copernican revolution I recommend the classical (Kuhn 1957).

[4] See Leibniz (1710, Discours Preliminaire, Sect. 65).

Relativistic and Probabilistic Events

Said this, when we deal with the notion of event we immediately run in a serious difficulty. In fact, there are two major and apparently incompatible definitions of this notion:

- One is that accepted by SR (and therefore ultimately due to Einstein): here, an event ideally denotes a point in space-time (which, for a given inertial frame of reference, can be specified by position and time), i.e. with no extension in neither space nor time.[5] This is the notion of event that was discussed in Sect. 2.3.2.
- The second definition is in the context of probabilistic theories. I report here the excellent Wikipedia's definition[6]: in probability theory, an event is a set of outcomes of an experiment (a subset of the sample space) to which a probability is assigned. According to Sect. 2.1.2, a single outcome may be an element of different events, and different events in an experiment are usually not equally likely, since they may include very different groups of outcomes.

The first definition treats an event as something determined by its space-time coordinates and therefore its frame of reference. Although also the complexity of an event is taken into account, so that it can in fact extend for a space-time interval, it is implicitly assumed that we could lead complex events to elementary relativistic ones that just represent (idealised) points in space-time. Is this justified? We should not forget that the role of the observer is crucial in relativity theory and consists in recording coincidence events, i.e. events occurring at the same space-time point like reading a clock when two particles collide. In this way, we are able to fix the proper time of the (collision) event. In other words, space-time localisation requires an act of observation. In such a case, we are considering the time that the clock's signal takes to arrive at our eye as negligible small so that we can assume that our observation and the event itself share the same reference frame. However, this is again an idealisation. It is true that we can establish some procedures for synchronising the clocks and so get some coincidence.[7] However, properly speaking, synchronisation of clocks can be done only in the same inertial frame of reference (which is therefore presupposed) and makes sense only if the clocks remain synchronised. In all these cases, even for a single clock (reading the time displayed by that clock), the coincidence between observed event and event of observation is defined as exchange of light signals between clock and eye, and therefore, for the issue of observation, is in fact defined through our act of perception, and, although we can record the collision event by some machine, as anticipated there will be always the reading of the value on the graduate scale of the recording machine. It is only in the limit of an infinite chain of such operations, enabling us to recalculate the proper time of the event that it can be said that we have a full coincidence between the event and its observation,

[5] Misner et al. (1970, p. 6); Geroch (1978, Chap. 1).

[6] https://en.wikipedia.org/wiki/Event_(probability_theory). An extensive treatment of this subject can be found in Leon–Garcia (2008).

[7] Rindler (2001, p. 42).

but this is another idealisation. This suggests that "space-time is indeed a coherent organisation of inferences based on a causal structure for events. The clock itself is just a sequence of events—a light pulse bouncing between two mirrors. The closest are the mirrors, the more precise is the clock, and the more refined is the coordinate system."[8] Of course, there will be other observations of the same event that by definition occur in other reference frames, and they represent the event as observed from other 'points of view' (which explains the relativity of synchronisation). Clearly, there are appropriate invariances and covariances among all these reference frames. Nevertheless, the determination of proper space-times is crucial also for computing invariances like the space-time distance (2.79). Thus, in any case we cannot fix the space-time of an event without some coarse-graining.

The second definition, at the opposite, treats an event as being a set of outcomes having a certain probability. Here, a single outcome is both fine-grained by definition and does not occur with certainty. In other words, the single outcome is treated here as a kind of limit of the fine graining of the event. Moreover, it is implicitly assumed that it is random. Thus, it is the second definition that is in agreement with the notion of quantum-mechanical event (in fact, we shall discover that QM can be reformulated as a general probabilistic theory). Note that, although also quantum-mechanical events can be coarse-grained, as it is the case for joint detection, I assume that behind any event there are always one or more totally fine-grained events like the absorption or emission of a photon. When I shall speak of *quantum event*, I refer to this case when not otherwise specified. We shall see below the exact meaning of this term. Before that we need still to deal with some objections.

Einstein's Point of View

This duality of definitions is another manifestation of the fundamental incompatibility of QM with relativity that many physicists take for granted. In the following I shall try to show that this is not fully correct. By now, I note that the main source of difficulty is the fact that, if events happen randomly, they are not predetermined. Instead of, according to relativity, events must be fully determined as far as they are assumed to be perfectly localised, i.e. occupying a point in space-time. In other words, appears here again the already noted difference between classical and quantum physics: see e.g. Fig. 2.7, Sect. 2.2.2. Thus, it is natural to deepen our understanding of Einstein's opinion. In a paper published 1 year before the EPR contribution, he writes[9]:

> I still believe in the possibility of giving a model of reality, a theory, that is to say, which shall represent events themselves and not merely the probability of their occurrence. On the other hand, it seems to be certain that we have to give up the notion of an absolute localisation of the particles in a theoretical model.

[8] D'Ariano and Tosini (2013).

[9] Einstein (1934, pp. 168–69).

Although still animated by the idea that it is possible to have a description (likely a prediction) of singular events, he renounces here to the tenet of perfect localisation, although is not specified whether there is, as a necessary condition of localisation, some indeterminacy in the conditions of the occurring of events (no predetermination) or in the events themselves (no determination), where the latter is stronger than the former since encompassing it. Note that the German physicist Gerhard Hegerfeldt proved that a strict space-time localisation of particles subjected to both quantum-mechanical and relativistic constraints contradicts the possibility of generalised causal connections.[10] In fact, all the talk about 'particles' and their space-time localisation should be rephrased in terms of properties of, and interactions among, quantised fields.[11] We shall consider later the consequences of this standpoint. In an interesting study, the American historian and philosopher of physics D. Howard has pointed out that Einstein, facing subsequent experiments on EPR, would have likely rather rejected determinism than separability, thus accepting a form of indeterminacy.[12] Now, in both of the above cases (i.e. determination and predetermination), to admit some indeterminacy cannot be completely disjoint from assuming some form of event randomness. I shall come back on Einstein and show that his position was in fact more complex as ordinarily assumed. Indeed, the causal structure that SR imposes on causal processes is based on the relative relations among events satisfying Lorentz transformations. Now, the classical identification of causality with laws (Sect. 2.4.1) is based on the full dominance of physical laws on happenings. However, if events are acknowledged as random, there is no need that those relations also determine the constitution of the events themselves, and this could explain, at least in part, Einstein's worry.

Uncontrolled Events

Nevertheless, a considerable problem is the following. Even if almost all scholars are willing to admit the reality of detection events happening in our labs, they are at least sceptical about their occurrence elsewhere in the natural world. The reason is simple: in our laboratories they are *controlled* events and therefore experimental facts, elsewhere they are uncontrolled, and it is always possible to think that, if the wave functions of quantum systems evolve according to the laws of QM when not subjected to experimental procedures, no event at all should occur (see, for instance, Wigner's standpoint as summarised in Sect. 3.1.1 but also the delayed-choice experiment reported in Sect. 2.4.3). The notion of event, according to this view, would make sense only when there is an observing agency. Note that this is not the case in a relativistic context where, as mentioned, observations in terms of coincidence events are crucial for the theory but nobody assumes that relativistic

[10] Reference papers are Hegerfeldt (1974, 1985).

[11] As elegantly proved in Malament (1995).

[12] Howard (1992).

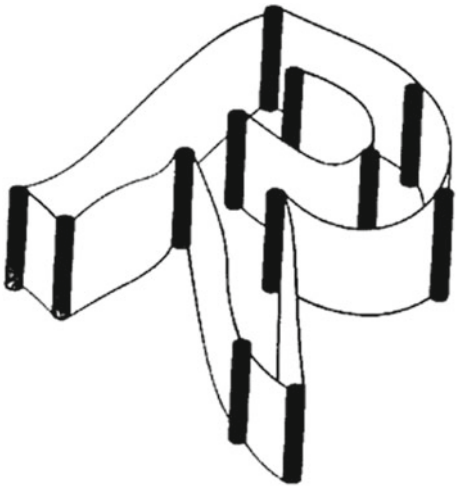


Fig. 5.1 The notion of reality according to Wheeler (1983). It is made of inferences (threads) connecting observations (stakes)

events only occur when there is an observer. The problem is important because we are in fact interested not only in what the theory may (probabilistically) predict but also in what is actually going on.[13]

QM Anachronistic?

Unfortunately, this seems to have been also the point of view of Bohr and even Wheeler (Sect. 3.1.1). In fact, the latter scholar said that what we call reality is the thread represented by our inferences that connects the different observations: he spoke of reality as a construction of imagination and theory (Fig. 5.1).[14] As previously said, I agree that our theories and models of reality are indeed "a construction of imagination and theory", and that inferences play a crucial role.[15] However, nothing prevents us to assume that these theories can lead us in our experience of a reality that does not depend entirely on those theories. To assume that things stand not so would have some unreasonable consequences. I have already mentioned that we could not explain how we come to correct these theories in the light of further experience. But there are also scientific objections to that standpoint: for instance, we should wait for the appearance of *homo sapiens* on the Earth (or other rational beings elsewhere in the universe) for getting (observed) quantum events. In fact, at cosmological scale, the evolution of man happened very late (and is likely that also other intelligent beings were not somewhere in our universe in the first billions of years). Then, how could the structure of matter, a solar system, a planet and the life evolving on it been

[13] 'T Hooft (2016, pp. 30–31).

[14] Wheeler (1983).

[15] Carnap (1928); Margenau (1950, Sect. 4.5). One has spoken of the "man–made" character of physical (scientific) theories (Rindler 2001, p. 190).

formed at all without events occurring also spontaneously? Is really all that only the construction of our imagination? In fact, how could *homo sapiens* have appeared at all as potential observer (provided with imagination) without these conditions? These structures, events and phenomena are even the quintessence of the classical or semi-classical world which we ordinarily experience. A consequent epistemic–phenomenic interpretation of events needs to provide a reasonable answer to these questions. In fact, the danger of this point of view is a potential conflict of QM with other scientific disciplines. Even more, although considered a leading discipline, here QM could appear to be not sufficiently updated relative to the enormous recent progresses in cosmology, chemistry, evolutionary biology, to quote here only some of the fields involved when dealing with these issues. In other words, an idealistic interpretation of quantum physics could risk scientific anachronism, and to simply affirm that it is the most basic theory of the world is of no help here: in science only rational arguments are valid and never those based on whatever kind of authority. On the other hand, we have evidence that the structure of matter can be explained only through the deterministic–probabilistic laws of QM (indeed we have today consolidated fields like quantum molecular physics and quantum chemistry[16]), but, certainly, if no event existed apart from those occurring and observed in our labs, it seems to me that we would never be able to explain the relative localisation of protons and electrons for not speaking of molecules, and without this we could get no structure of matter at all. Therefore, we need to understand Wheeler's statement correctly and, if we do not like to incur extreme (strong) forms of subjectivism–idealism or quantum objectivism, i.e. express a simple denial of reality as we perceive it, we need to admit that, also independently from our lab procedures, events (i.e. results of whatever physical interaction among *open* systems that could in principle be observed but did not need an observer to occur at all) *happen everywhere*.[17] As a matter of fact, we everywhere observe quantum events happening spontaneously, like a photon absorbed or released by an electron (e.g. in the photoelectric effect). As mentioned, also the spontaneous building of molecules is a fact well described by quantum chemistry. This is what decoherence predicts (Sect. 3.1.3).

Minimal Definition of Quantum Event

In conclusion, if this is right and we desire to go a little bit further than the simple removal of the problem, we need to generalise the notion of quantum event disentangling it from our experimental procedures, although the latter need to be still our 'Polar star' when dealing with such generalisations.

Let us establish some minimal conditions that the notion of quantum (and probabilistic) event, understood in this way, needs to satisfy the following:

[16] On these subject I suggest the following textbooks (Atkins and De Paula 2006; Atkins and Friedman 2005). I shall come back on these problems.

[17] Joos and Zeh (1985).

- It is a random event.[18] In other words, in *identical physical conditions* (for instance, of experimental kind) quantum systems can give *different outputs*, e.g. out of photons in an arbitrary identical state of linear polarisation superposition some pass a vertical filter and some do not.
- Of these happenings, the only thing that we can compute is its *probability* to occur, and this is therefore *irreducible*.
- Events when happening are absolutely determined, *actually determined* and represent therefore a kind of choice that is both the source of novelty and an embryonic form of being due to their potential further consequences. Recalling that all relativistic events can be considered as coarse-grained, I assume that only a quantum event is a perfectly fine graining of probabilistic events. It is an *outcome*, i.e. a singular variation in whatever physical conditions.

Some comments seem appropriate here. The first comment is that *a priori* indeterminism (or randomness) of events is more fundamental than uncertainty relations,[19] as remarked in Sect. 1.2.4.

The second comment is that probabilities are objective as often announced (Sects. 2.1.2 and 3.2.3). Any series of quantum detections, being unpredictable and likely algorithmically uncomputable (Sect. 3.2.4), represents a denial of the classical assumption that events happen with certitude, i.e. with either probability 1 or probability 0. So, any of them (apart from the trivial case of the measurement of an observable on a system's state that is already an eigenstate of that observable) has an objective probability $0 < \wp < 1$ to occur that we can know if we have prepared the system in a certain state. This seems to imply that probabilities are only theoretic, as far as they are computed, thanks to quantum theory (essentially, Schrödinger equation and Born's probability postulate: Sect. 1.1.3). This is certainly true for the computation of these probabilities but does not imply, in any way, that these probabilities exist only on the basis of our computation. At the opposite, we are confident that quantum theory is excellent for computing their *real* values. The confusion between objective probability (and incertitude), on the one hand, and subjective evaluation (and incertitude), on the other, is very common. I shall consider a classical example for clarifying this point. Suppose that we need to calculate the probability to get a six with a dice but we know nothing about a dice (neither that it has six faces). A priori, the value that we can assign to this event is 1/2. But suppose now that we subsequently have the opportunity to examine the dice. Then, we shall assign a probability of 1/6. If we like to avoid the incorrect conclusion that increase in knowledge can imply increase in incertitude, we need to distinguish between the two kinds of incertitude. *Subjective* incertitude relies on the partiality of information that we have about some possible events (that could be in themselves uncertain), while *objective* incertitude is dictated by the objective conditions that are exhaustively known (like knowing the state of a prepared and premeasured quantum system but ignoring the

[18] The first serious discussions of randomness are in Poincaré (1907), Borel (1920).

[19] Elitzur (1992).

possible detection outcome).[20] This issue is really crucial as far as I have pointed out that the objectivity of information depends on that of probabilities.

The third comment is that any event, once happened, is a source of further effects: *no event without effects of some kind*. For instance, a detector that absorbs a photon can give rise to a light pulse which in turn can have further effects, like being perceived by our optical–receptor system: see e.g. the reaction chain in the Schrödinger cat experiment (Sect. 2.2.5). This is also true from a relativistic standpoint. In fact, although relativity does not agree with such modality of the event's occurrence, it agrees that, by definition, an occurred event is determined and therefore is and determines in turn some kind of variation. If we do not postulate this, we should admit that events occur that could never be observed. Thus, actuality (occurring events) is different from potentiality as far as it closes a process of determination and eventually opens a new one (or new ones). In traditional philosophical language, a random event represents an *ex post* necessity: it was not necessary for itself that it happened (in fact, possibilities but also potentialities do not determine events), but it is necessary once happened[21]: it cannot be washed out or made unhappened.

Moreover, a quantum event has no alternatives: indeed, before occurring it is not an event, and, when occurs, it erases any possible alternative: in the old language of philosophy, *determinatio negatio est*.[22] In fact, those alternatives concern the level of the possibility but not that of actuality. The main concern of the MWI is to avoid the rule of thumb, according to which (in Deutsch's words) "whenever a measurement is made, all the histories but one cease to exist".[23] I insist that we should carefully distinguish between the plane of possibilities, where, if we agree that they represent a level of reality, nothing changes, and the level of the actual events happening in our physical universe, a problem that is fully independent of the fact whether there are other physical universes or not. As mentioned in Sect. 3.1.2, I think that to mix these two very different issues is one of the major sources of trouble here. Now, from the point of view of our universe, in fact all other alternatives are made locally inaccessible, and this is something that also supporters of the MWI should not deny.[24]

5.1.2 Quantum Features

What They Are

Physically speaking, interference terms often manifest themselves in interference phenomena, from single systems up to entangled ones. These interference effects are

[20] On this distinction see Poincaré (1899).

[21] A problem already analysed in such terms in ancient philosophy: (Aristotle 2019, 9, 19a).

[22] Letter of June 2, 1674 to Jarig Jelles (Spinoza 1972, IV, p. 240).

[23] Deutsch (2011, Chap. 12).

[24] "Once the decision is made, it becomes 'real' and there is nothing random about it anymore" (Laughlin 2005, p. 44). See also Laughlin (2005, p. 130).

due to correlations. Now, as we shall discover, quantum-mechanical correlations have both a pure quantum and a classical part. Let us introduce the term *quantum features* as a general way to denote the quantum part of the correlations, something, as we shall discover, that is present in any quantum system and quantum effect, also for single systems. And here we deal with an enormous difficulty: classical correlations are among *different* systems (and thus could depend on the nature or properties of those systems), while quantum features are present also in the simple case of superposition and, as remarked in Sect. 2.1.2, appear to be a relation without *relata*. This is much more difficult to admit as something real. In other words, when I have provided evidence for the reality of correlation as causal constraints (Sect. 3.4.2), this does not imply the reality of quantum features, since one could assume that only classical correlations among systems are real physical interdependencies.

Bénard Cells Displaying Classical Correlations

Thus, we deal here with the problem of capturing what is the specificity of quantum features relative to other kind of correlations and what is their role. In fact, correlations are ubiquitous in nature and, as it is clear so far, exist also for classical systems, especially when we consider chaotic or complex systems (but also the structures of molecules). A known example is represented by the Bénard cells, after the name of the French physicist Henri Bénard (1874–1939). Consider a thin layer of liquid between two large parallel plates (Fig. 5.2). If the system is in equilibrium, with the liquid and the two plates at the same temperature, and the liquid is motionless, then it is homogeneous. Suppose now that the upper plate is heated slowly. The heat will pass from the upper plate to the liquid and will be transferred through the liquid to its bottom layer by the process of thermal conduction. In thermal conduction, there is no bulk motion of the liquid but rather a greater thermal motion of the molecules that causes the transfer of heat from the warmer layers to adjacent cooler layers. However, as the temperature of the upper layer is increased, a stage is reached (at a

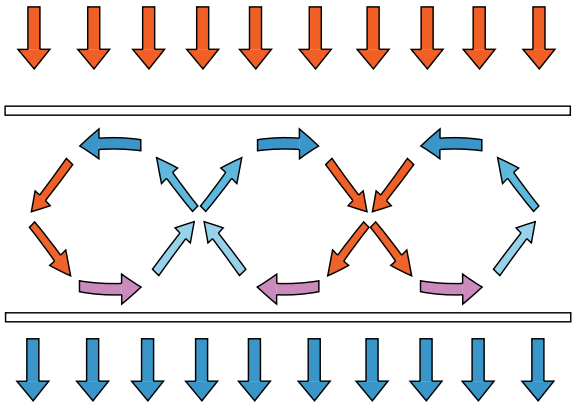


Fig. 5.2 Bénard cells between two metal layers: heat comes from above. Each cell runs either clockwise or counterclockwise (this depends on the long-ranging correlations). They assume a hexagonal form. Adapted from Auletta (2011, p. 188)

critical temperature) where the liquid overcomes its viscosity (the internal friction which opposes movement) and begins to undergo bulk motion. This results in a transport of heat by convection currents. The currents are not random but rather they lead to the formation of patterns, and often one sees small convection cells, called Bénard cells. They assume a hexagonal form and each one runs either clockwise or counterclockwise. This depends precisely on long-ranging (classical) correlations, so that all cells around any one need to rotate into the opposite direction relative to the latter. In fact, what is remarkable is that this organisation arises simultaneously in the whole fluid, which denotes correlations among different small fluid volumes and not exchange of signals.

Entanglement as a Combination of Classical and Quantum Correlations

Now, what is the difference between classical and quantum correlations? As mentioned, classical correlation is interdependency among variables or parameters pertaining to *different* systems or subsystems (for instance, among the rotation direction of the different Bénard cells). However, quantum superposition (that only expresses quantum features) is not among variables pertaining to different subsystems but among components (which can be associated with possible values of the *same* observable) of the *same* system as anticipated in Sect. 3.1.1: for instance, between spatial components like 'lower' and 'upper path' in an interferometer, as displayed by Eq. (2.25). This is why we have the phenomenon of self-interference (Sect. 1.2.1). This difference also clarifies why this phenomenon is unknown classically. Now, entanglement, as anticipated, is a quantum correlation among different (and even distant) subsystems, and thus shows a *combination* of classical and quantum correlations (Fig. 5.3). This is why there is a characteristic difference between superposition and entanglement: superposition is basis-dependent (it is relative to the observable with which we describe the system: see Sect. 1.2.2) and it is a character *intrinsic* to the single quantum system, while entanglement is not dependent on that description

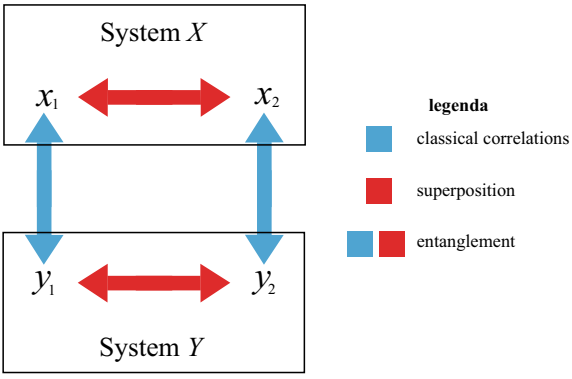


Fig. 5.3 Particular case in which the entanglement of two degrees of freedom is among two subsystems and each observable has two eigenvectors only. Although a particular case, the figure is very enlightening and therefore also easily generalisable. Adapted from Auletta and Wang (2014, p. 197)

and is therefore a *character of the compound state*: if two subsystems are entangled, they will be so for many other observables, although, operationally speaking, in order to entangle systems, we need to perform an operation through a specific observable (like measuring the Bell observable: see Sect. 3.3.6). Suppose indeed that the composite system at a certain time is described by the state vector[25] (see Eq. (1.392))

$$|\Phi\rangle_{12} = \frac{1}{\sqrt{2}}(|h\rangle_1 \otimes |h\rangle_2 + |v\rangle_1 \otimes |v\rangle_2), \tag{5.1}$$

which displays entanglement of horizontal and vertical polarisation of two photons. Now the corresponding density matrix is

$$\begin{aligned}\hat{\rho}_{12} &= |\Phi\rangle \langle\Phi|_{12} \\ &= \frac{1}{2}(|h\rangle \langle h|_1 \otimes |h\rangle \langle h|_2 + |v\rangle \langle v|_1 \otimes |v\rangle \langle v|_2 \\ &\quad + |h\rangle \langle v|_1 \otimes |h\rangle \langle v|_2 + |v\rangle \langle h|_1 \otimes |v\rangle \langle h|_2),\end{aligned} \tag{5.2}$$

where for the sake of simplicity I have put on the second row the analogous of classical terms. Note that the components $|h\rangle \langle v|_1 \otimes |h\rangle \langle v|_2$ and $|v\rangle \langle h|_1 \otimes |v\rangle \langle h|_2$, in the third row, represent the features or the interference terms, of the kind (3.13), not between states of different systems but between different state components *of the same* system. This conclusion could be put in question by pointing out that the state (5.2) has a particular form. As a matter of fact, if we take another kind of entanglement, e.g.

$$\begin{aligned}\hat{\rho}'_{12} = \frac{1}{2}(&|h\rangle \langle h|_1 \otimes |v\rangle \langle v|_2 + |v\rangle \langle v|_1 \otimes |h\rangle \langle h|_2 \\ &+ |h\rangle \langle v|_1 \otimes |v\rangle \langle h|_2 + |v\rangle \langle h|_1 \otimes |h\rangle \langle v|_2),\end{aligned} \tag{5.3}$$

things may appear different. It is clear now that each time we have a photon 1 polarised vertically, photon 2 is polarised horizontally and vice versa. Therefore, the two photons show a kind of antiparallelism. This concerns, however, the classical correlations between the two systems. In fact, when the single subsystems are considered, we have precisely the *same* cross terms (e.g. $|h\rangle \langle v|_1$) as in the density matrix (5.2). Thus the notion of features as *intrinsic correlations* makes perfect sense and is the ground of self-interference. In other words, we can solve the problem raised in Sect. 2.1.2 and say that such intrinsic correlations are in fact more basic and independent of the existence of *relata*, if we understand with *relata* some kind of physical system.[26] As I have remarked that quantum features are intrinsic to the system, we could call this modality of correlation a *self-correlation* (to do not mix with the notion of autocorrelation used in some fields of physics and signal theory).

[25] A quick introduction to the subject is in Auletta and Wang (2014, Sect. 7.9).

[26] Kuhlmann (2013).

Pro

Resuming, while almost all scholars have considered events as real, to admit that interference terms have a real substrate is certainly harder. The followers of the subjective interpretation of measurement and most followers of the Copenhagen interpretation would say that they represent a pure mathematical aspect without any ontological significance at all (Sects. 3.1.1 and 3.1.4). In fact, this is not unprecedented in the history of science. We may recall here that in the classical theory of electromagnetism both the scalar and vector potentials were useful tools of calculation but deprived of any physical significance in that context.[27]

Nevertheless, there are also some scholars who would, at least partly, accept the idea that quantum interference terms denote something that is somehow real. Let us consider some of them. We may recall the supporters of the MWI, who, following in turn Schrödinger's objectivism, consider interference and correlations as crucial for the description of any quantum system (Sects. 3.1.2 and 3.2.1). I recall that Schrödinger had an abiding faith in the powerfulness and objectivity of the formalism of QM and was very well aware that the object of this formalism is represented precisely by quantum correlations.[28] In fact, he considered quantum correlations as what was the distinctive trait of quantum physics relative to previous physical theories,[29] and the same for Heisenberg, as reported in Sect. 3.4.3. Even more, to Schrödinger these correlations are precisely what makes physical reality. And this is well understandable: a so crucial aspect of quantum phenomena cannot be considered a fiction or a pure mathematical tool. Another school that would partly agree would be the Bohmian one with its notion of quantum potential, provided that it is not understood as an 'ordinary' physical potential but as the quantum aspect of correlations making classical trajectories of the particles impossible, as I have reminded in Sects. 3.3.4 and 3.4.2. Finally, most of those who follow de Broglie have strongly insisted to attribute some ontological reality to the quantum waves, and in fact quantum features have to do with the 'wave-like' aspect of quantum systems.[30]

Contra

The fact remains that this still represents a (although authoritative) minority of physicists. The reason for that is simple: if quantum events are irritating due to their random and contingent nature but, at least when observed, are deeply rooted in the experimental tradition of physics, features simply clash with this history that has elevated the reductionist methodology to a basic postulate concerning the nature of physical

[27] On this subject and its quantum-mechanical subtleties, see Auletta et al. (2009, Sect. 13.8). I shall come back on these problems.

[28] Schrödinger (1935a).

[29] Schrödinger (1935b).

[30] For a review, see Auletta and Tarozzi (2004b). See also Auletta (2014a).

systems: accordingly, everything should be studied by reducing it to elementary and local components, so that everything can find a bottom-up explanation.[31] Moreover, events are the content of our immediate and local experience, at least in principle. The fact is that any detection is by definition a singular happening, i.e. a single click of a detector that, at least in principle, can be experienced by an observer. When the detector clicks, this is because a quantum impinges on it. Therefore, what immediately appears in detection is the corpuscular nature of quantum systems. In fact, interference terms are locally washed out in this very act of information acquisition (Sect. 3.1.3). So, nobody can have direct experience of quantum features, and, as a consequence, many believe that such quantum terms do not exist at all.[32]

Direct Experience and Inferences

The epistemological background of this scepticism is simple: to have knowledge of reality seems to imply the possibility (at least in principle) to make *direct experience* of it, where

- With "experience" is meant what is not based on an inferential process but relies on perception[33] (this is what makes experience *singular* and thus different from reasoning), and
- With "direct" I intend any way through which we have access to sensory data (also through instruments) as opposite to what we know because it is reported to us by others (and this is what makes of experience essentially an *individual* happening).[34]

So far, I agree with Wheeler (see previous subsection). The latter point is less relevant to the discussion here since it is customary to accept experiences performed by others (like experiments in which we were not directly involved) when we trust the team, the journal in which those results are reported, and everything appears correct under our critical scrutiny. The first point is less easy to deal with. Of course, Wheeler was aware that, also in our direct experience of reality, inferences play a role. Nevertheless, as pointed out in Sect. 3.1.1, to speak of experience makes sense only if the external reality imposes on us the perception of certain facts that we finally trust as evident.[35] In other words, in any direct experience there is something that is irreducible to inference or general reasoning. At the opposite, we can reconstruct the profile of interference only by comparing many detections and by varying the phase difference between the components of systems prepared in the same state (Fig. 5.4).[36] However, such a reconstruction is by definition an inference and not an experience (precisely

[31] On the problem of reductionist methodology see Auletta (2012).

[32] As pointed out in Auletta and Torcal (2011).

[33] See also Margenau (1950, Sect. 15.1).

[34] Poincaré (1897).

[35] As stressed in Peirce (1885). See also Auletta (2011, Sects. 1.2.4, 2.2.3).

[36] On this problem see also Schlosshauer (2007, Sect. 2.5.3).

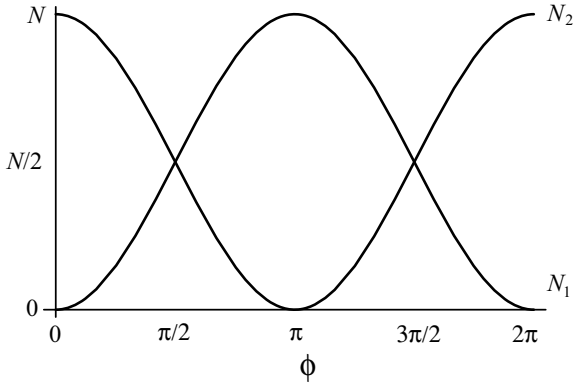


Fig. 5.4 The two curves show the statistical results of photon counting at detectors D1 and D2 in an interferometry experiment (see e.g. Fig. 2.2, Sect. 2.1.2). N_1 and N_2 denote the number of photons counted at detectors D1 and D2, respectively. Of course, in the ideal case (100% of detectors efficiency) for each value of the relative phase ϕ, we have $N_1(\phi) + N_2(\phi) = N$, where N denotes the total number of photon detected

because we need to compare and evaluate different detections that as such do not need to happen in the same conditions and be observed by the same scholar). This is strictly related to the problem that we cannot measure the state of a single system in a single shot (see also Sect. 2.1.1).[37] Thus, so think many scholars, at most we could attribute a statistical significance to such interference terms.[38]

Einstein's Standpoint

The fact is that many of the things to which we attribute reality cannot, strictly speaking, be direct objects of experience, but their reality is inferred. Nobody has directly seen electrons for not speaking of quarks (what we see are certain tracks during collisions), or black holes or even the Big Bang. Nevertheless, apart from strong idealists, few would doubt that they are real. In this context, I recall that Einstein, on the outline of what reported in Sect. 3.4.3, said that we cannot build a physical theory with observable quantities only[39]:

[37] D'ariano and Yuen (1996).

[38] On the more general statistical interpretation of QM see Ballentine (1970). For historical reconstruction see Jammer (1974, Chap. 10) and for further references see also Auletta (2000, Sect. 6.5.2).

[39] As reported in Heisenberg (1969, Chap. 5): "vom prinzipiellen Standpunkt aus ist es ganz falsch, eine Theorie nur auf beobachtbare Größen gründen zu wollen. Denn es ist ja in Wirklichkeit genau umgekehrt. Erst die Theorie entscheidet darüber, was man beobachten kann. Sehen Sie, die Beobachtung ist ja im allgemeinen ein sehr komplizierter Prozeß. Der Vorgang, der beobachtet werden soll, ruft irgendwelche Geschehnisse in unserem Meßapparat hervor. Als Folge davon laufen dann in diesem Apparat weitere Vorgänge ab, die schließlich auf Umwegen den sinnlichen Eindruck und

> From the point of view of the principles, it is absolutely false to claim to ground a theory only on observable quantities. In fact, in reality it is quite the opposite. It is first the theory that decides on what one can observe. See, observation is in general a very complicated process. The process that should be observed provokes some happening in our measuring apparatus. As a consequence, other processes go on in this apparatus, which finally, through indirect ways, provoke the sensory impressions and fix these results in our consciousness. Along this long way from the happening up to its fixation in our consciousness we should know how nature does work, must, at least from a practical point of view, know the laws of nature, if we will claim that we have in fact observed something. Only the theory, i.e. the knowledge of the laws of nature allows us to make an inference from the sensory impression to the happening that is at its root. When somebody claims that he can observe something, one should rather and more precisely say: although we dare formulate new laws of nature that do not harmonise with current ones, we guess that these previous laws work so well from the happening object of observation up to our consciousness, that we can trust them and therefore speak of an observation.

This remarkable paragraph, in agreement with what has been examined in the previous subsection, shows that finally Bohr and Einstein were more convergent than it is generally assumed. Clearly, the crucial difference (what could be called an Aristotelian point of view) is that according to Einstein external stimuli set the whole process in motion. Such point of view was clearly shared by the late Heisenberg[40]:

> What Kant had not foreseen was that these a priori concepts can be the conditions for science and at the same time can have only a limited range of applicability. When we make an experiment we have to assume a causal chain of events that leads from the atomic event through the apparatus finally to the eye of the observer; if this causal chain was not assumed, nothing could be known about the atomic event. Still we must keep in mind that classical physics and causality have only a limited range of applicability.

Note also that Einstein is perfectly clear about the necessity of inferences (as well as about their (inverse) Bayesian character, on which I shall come back). I also incidentally note that he says that we cannot change our theory and laws without starting from previous theories and laws.

Indirect Experience

This raises the general problem of how we can conceive inferences in a way that makes justice to the quantum theory. Although quantum theory is and must be centred on

---

die Fixierung des Ergebnisses in unserem Bewußtsein bewirken. Auf diesem ganzen langen Weg vom Vorgang bis zur Fixierung in unserem Bewußtsein müssen wir wissen, wie die Natur funktioniert, müssen wir die Naturgesetze wenigstens praktisch kennen, wenn wir behaupten wollen, daß wir etwas beobachtet haben. Nur die Theorie, das heißt die Kenntnis der Naturgesetze, erlaubt uns also, aus dem sinnlichen Eindruck auf den zugrunde liegenden Vorgang zu schließen. Wenn man behauptet, daß man etwas beobachten kann, so müßte man also eigentlich genauer so sagen: Obwohl wir uns anschicken, neue Naturgesetze zu formulieren, die nicht mit den bisherigen übereinstimmen, vermuten wir doch, daß die bisherigen Naturgesetze auf dem Weg vom zu beobachtenden Vorgang bis zu unserem Bewußtsein so genau funktionieren, daß wir uns auf sie verlassen und daher von Beobachtungen reden dürfen."

[40] Heisenberg (1958, p. 90).

observable quantities (otherwise we may deal with not well-defined concepts or even with invented ones), "things that are not directly observable may still exist and as such play a decisive role in the observable properties of an object. They may also help us to construct realistic models of the world."[41] When we cannot directly experience an object, we can nevertheless make real experience of *other* processes or objects that are somehow related to our target, which in this way is inferred. It is a kind of "contiguity of the experience" to ensure us about the existence of the target (recall the importance of continuity for mechanical causation). This is quite general and somehow concerns already observation, because, with the words of Max Born, the method of science "consists in finding correlations of one kind of subjective sense impressions with other kinds, using the one as indicators for the other, and in this way it establishes what is called a fact of observation".[42] In general, there are two ways in which objects or processes can be used to *infer* properties of other ones as marks of their reality:

- Either we transpose certain characters that other objects, which we consider to pertain to the same class of our targets, have to the latter, i.e. we proceed analogically or metaphorically or also taxonomically: in this case, we are in fact building a smaller class that includes properties of both the already known objects and our target; or
- We try to establish certain causal interdependencies between those other objects and processes and our targets, evaluating how the latter affect and are affected by them: in this case we proceed metonymically or also indexically.[43]

Both inferential processes need obviously to rely on some characters of the targets and the other involved objects. The reader may be horrified by the use of the word "metaphor" or "analogy" in a scientific context. In fact, modern science has rejected this kind of notions for the use that some scholasticism made of them. However, the authoritative American evolutionist Richard Lewontin recalls us that the language of science "is filled with metaphors" since "the entire body of modern science is an attempt at explaining phenomena that cannot be experienced directly by human beings".[44] It suffices to think of the role of the Correspondence principle to be easily convinced of this fact (Sect. 1.3.2). The relevance of analogy for extending scientific enquiry was explicitly stressed by Dirac.[45] This view finds today a formal support and a generalisation with Category theory, which allows us to affirm that "'comparison' is an important aspect of Mathematics, of Science, and, in fact, of general knowledge

[41] 'T Hooft (2016, p. 33).

[42] Born (1949, p. 33).

[43] On this subject the interested reader may have a look at Auletta (2011, Sect. 20.6), where also further literature is quoted. I recall here that Peirce clearly considered these two different kinds of association (Peirce 1898, pp. 234–36).

[44] Lewontin (2000).

[45] See e.g. Dirac (1945). On this problem see also Bokulich (2008, Sect. 3.2). The author stresses that Dirac had supported an inverse correspondence principle, according to which QM can lead analogically into new developments in CM.

acquisition", while analogy represents "the 'flip side' of comparison and is essential in the 'inductive' side of knowledge acquisition".[46]

Of course, the correct predictions that we are able to formulate, once having resisted experimental tests, will support our ontological ascription based on inferences developed through metaphoric or metonymic extension of characters and actions of other objects. So, most of what we call reality is caught through an opportune mix of experience and inference, as proposed by Wheeler, also in many cases in which we speak of experience as such or of objects that are undoubtedly real, and so far this accounts for the constructive nature of theories as well (as explained in the previous subsection). By now, keep in mind that "experience" and "observation" are not the same: the latter constitutes a part of the former.

Nevertheless, a problem still remains: there is in fact a difference here, because we can say that all of the objects previously quoted (electrons, black holes, etc.) are plus or minus localised (they are or happen in a limited portion of space-time), while features are by definition delocalised, being the interdependencies among different parts or components of the same system (or even among several systems). So, they are a very bizarre case and we need some further evidence for daring any ontological ascription to features.

In Which Sense Quantum Systems Are Real

In Sect. 2.4.3, I have reminded of an important experiment, the delayed choice. Let us consider its Mach–Zehnder version. In this experiment we are free up to some femtoseconds before detection to set the apparatus in such a way to either detect the path or the interference. So, what have we between the first beam splitter and the final detection (or, at least, the choice of the setup)? As a matter of fact, by tuning the laser source appropriately (and by previously testing it), we have pumped a single photon in the apparatus and we finally can detect again a photon (assuming that the detector works perfectly and there are no misses). Now, if sometime before detection we have not yet decided which of the alternative settings to put in place, it is clear that we cannot affirm that there is a corpuscular entity in between. Note also that no event happens before the final detection. However, the question is: is there an entity at all? In fact, we are not asking here of the properties or also the state of the system, but we are asking if there is *a physical system* before observation, independently from how we can describe it. If one follows the subjectivist interpretation of measurement or the epistemic version of the Copenhagen interpretation (Sects. 3.1.1 and 3.3.2), one is tempted to say that nothing exists in between. But this would have far-reaching consequences, much worse than to admit quantum features: in fact, in such a case, we would be forced to say that the incoming photon is somehow disappeared while

[46] Brown and Porter (2006). This enlightening paper is also a good conceptual introduction to category theory, of which a lot will be said below. I had only used the term *abductive* instead of *inductive* (see the end of Sect. 3.3.3). Among the first scientists to have acknowledged the role of analogy in science, and especially in mathematics, is H. Poincaré (1897). See also Poincaré (1902, pp. 158–59); Poincaré (1905, p. 38).

appearing eventually at detection (we cannot be agnostic on this crucial question!).[47] However, this would represent not only a violation of the principle of sufficient reason in its classical formulation (Sect. 2.4.1) but even of a more basic ontological requirement that is at the foundations not only of science but even of any form of rationality: no natural object can disappear into, or appear from, nothing (*ex nihilo nihil*). Any natural object can only be *transformed* in something else, although in ways that sometimes are not predictable or controllable.

Therefore, there must be something before the photon is detected and after it has entered the apparatus. What is the minimal reality that we are obliged to assume? Not a localised system and therefore not the single components or paths: in fact, we have not yet decided which setting to use, and only the choice of the path detection setup would make these components actionable or potentially acquirable (Sect. 3.2.3). Thus, in the absence of a specific choice of the apparatus and of the relative physical context, they only represent possibilities at this level, and although I am willing to attribute to possibilities some form of reality, what we need here is something more, that is, something capable to have *real* (observable) effects. Since the global state of the system can be conceived as consisting of those components and their interdependencies (which are always present independently of the choice of the settings), the only thing that remains are the *interdependencies themselves* among these components or paths (see also Sect. 2.1.2).[48] This is strong evidence that quantum features are real and are even the only entity to which we can attribute reality in the absence of any local interaction. I recall that there is an alternative option, consisting in assuming, following Deutsch (Sect. 3.1.2), that it is the correlation with the shadow photons (that in their respective worlds realise the different components of the superposition state of our photon) that is responsible for this phenomenon, what implies that we need to assume the reality of the shadow photons themselves. This solution, although viable, appears to be less economic than the proposed one. Moreover, can we really dismiss in this way the question about the reality of quantum correlations? What we need to show now is that quantum features can have influence on *single* experimental runs. Let us consider this problem.

A Proposed Experiment

We need an experiment that catches the spirit of the delayed-choice one and allows to show the influence of quantum features on each experimental run, solving both problems that have been raised. Let us briefly describe the ideal experiment proposed by some scholars.[49] A source laser pumps a photon in the state $|\gamma\rangle$ (Fig. 5.5). If we discard the cases in which both photons are detected by the same early detector and only consider the case in which, for each photon pair, one is detected by one early

[47] As pointed out in Auletta and Tarozzi (2004b).

[48] See also Auletta and Wang (2014, Sects. 4.8, 5.4–5.6).

[49] Auletta and Tarozzi (2004a). In this paper, due to some misprint, some displaced imaginary units appear in some calculations.

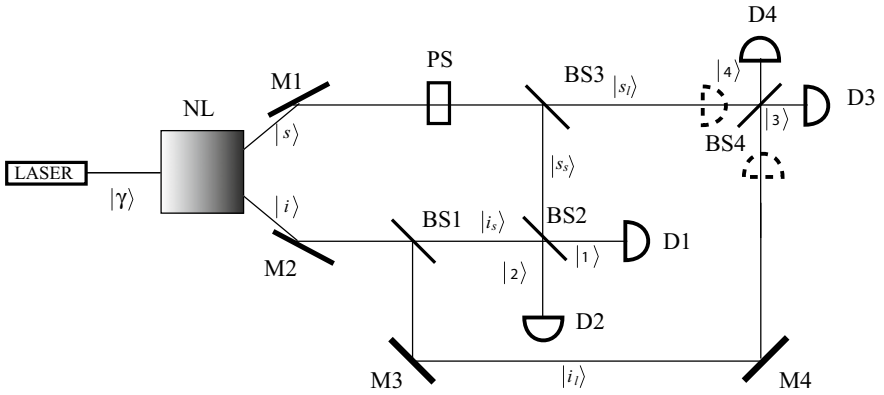


Fig. 5.5 Experiment proposed in Auletta and Tarozzi (2004a). A source laser pumps a photon in the state $|\gamma\rangle$. Successively, a parametric down-conversion allows the emission of two photons, the *idler photon* (in the state $|i\rangle$) and *signal photon* (in the state $|s\rangle$). If γ has frequency ν and energy $h\nu$, the two outgoing photons have smaller frequencies ν_i and ν_s (and energies $h\nu_i$ and $h\nu_s$), respectively, with $\nu = \nu_i + \nu_s$. The two beam splitters BS1 and BS3 split each photon into two components, the 'shorter' and 'longer' components ($|i_s\rangle$, $|i_l\rangle$ and $|s_s\rangle$, $|s_l\rangle$, respectively). The two shorter (longer) components, $|i_s\rangle$ and $|s_s\rangle$ ($|i_l\rangle$ and $|s_l\rangle$) are recombined at BS2 (BS4), giving rise to photons in states $|1\rangle$ and $|2\rangle$ ($|3\rangle$ and $|4\rangle$), falling to detectors D1 and D2 (D3 and D4). Eventually, detectors D3 and D4 are placed before BS4 when D1 or D2 has already clicked. Note that all beam splitters are here 50–50%

detector and the other by a late detector (by performing a so-called *post-selection*), we obtain that, when detector D1 clicks, detector D3 must click as well, and when detector D2 clicks, detector D4 must also click. Assuming that the initial state is a product state, after the nonlinear crystal and the first two beam splitters, the initial state evolves according to

$$|\gamma\rangle \overset{\text{NL,M1,M2}}{\longmapsto} |s\rangle\,|i\rangle \overset{\text{BS1,BS3}}{\longmapsto} \frac{1}{2}\left[(|s_s\rangle + |s_l\rangle)\,(|i_s\rangle + |i_l\rangle)\right] , \tag{5.4}$$

where the subscripts s and l indicate the shorter and the longer path, respectively. Considering the transformations of the kind (2.30):

$$|s_s\rangle \overset{\text{BS2}}{\longmapsto} \frac{1}{\sqrt{2}}(|1\rangle - |2\rangle)\,, \quad |s_l\rangle \overset{\text{BS4}}{\longmapsto} \frac{1}{\sqrt{2}}(|3\rangle + |4\rangle)\,, \tag{5.5a}$$

$$|i_s\rangle \overset{\text{BS2}}{\longmapsto} \frac{1}{\sqrt{2}}(|1\rangle + |2\rangle)\,, \quad |i_l\rangle \overset{\text{BS4}}{\longmapsto} \frac{1}{\sqrt{2}}(|3\rangle - |4\rangle)\,, \tag{5.5b}$$

we can obtain the final state

$$\overset{\text{BS2,BS4}}{\longmapsto} \frac{1}{2\sqrt{2}}\left[(|1\rangle + |2\rangle) + (|3\rangle + |4\rangle)\right]\left[(|1\rangle + |2\rangle) - (|3\rangle + |4\rangle)\right] . \tag{5.6}$$

When we discard double detections (i.e. all states of the form $|jj\rangle$, $j = 1, \ldots, 4$), it is easy to show that this state reduces to a special kind of EPR (entangled) state

$$|f\rangle = \frac{1}{\sqrt{2}} \left(|1\rangle\,|3\rangle + |2\rangle\,|4\rangle \right) , \tag{5.7}$$

although we deal here with a sequence of detections. It is interesting to stress that, because of this entanglement, we cannot know which photon has been detected at which of the two detectors D1 and D2, though, thanks to the entanglement, we can predict with certainty whether detector D3 or detector D4 will click *once* we know that either detector D1 or detector D2 has clicked: if D1 clicks then D3 will click and if D2 clicks then D4 will click. The point is that we obtain this information after an event has *already occurred* (detection by D1 or D2), unlike Wheeler's original proposal. This is precisely what makes the difference and allows for catching quantum features each experimental run.

Let us now consider what happens when detectors D3 and D4 are placed at a position before BS4 once a photon has already been detected by either D1 or D2: as said, it is a kind of delayed-choice experiment. In this alternative arrangement, we can know at the end of the experiment which photon has been detected by which detector and therefore the paths they followed, but we cannot foresee whether detector D3 or detector D4 will reveal the photon after either detector D1 or detector D2 has clicked. This is because, by displacing detectors D3 and D4, we no longer have the quantum interference terms

$$|1\rangle\langle 2| \otimes |3\rangle\langle 4| \text{ and } |2\rangle\langle 1| \otimes |4\rangle\langle 3| \tag{5.8}$$

between the two photons, which are similar to those displayed in Eq. (3.13). In other words, the quantum wave-like aspect of the photons has disappeared. Note that this is an effect due to the presence or absence of quantum features allowed by the experimental setup and not to some interaction. Of course, both classical and quantum components are present in the entangled state. Nevertheless, the typical quantum effects in play here require the contribution of the latter component.

In other words, we can make quantum features (as manifested in the entangled state (5.7)) and their effects appear and disappear according to both (i) our setup manipulations in due course and (ii) some detection event: their presence or absence will contribute to determine the experimental outcome. Summarising, we have obtained both a complementarity and a delayed-choice experiment that allows us to find evidence for interference (and therefore for features) *each single* experimental run.

Although to my knowledge no actual experiment has been performed on these outlines, these calculations appear quite eloquent. So, up to contrary evidence, we are authorised to assume that quantum features have in fact causal influence on single experimental runs. Clearly, this puts us again in a bizarre situation, since we are forced to admit something that can have influence on single happenings without possessing itself the nature of a single and localised reality (being by its own nature global or non-local). These problems will be further discussed in the following but,

by now, we can at least clarify the reason of the term *features*: they are characters of the state of quantum systems having influence on other elements of reality without being properties. In fact, properties are by definition locally ascribed and are, at least in principle, connected with direct experience, while features are not.

At the end of Sect. 3.4.2, we have introduced the idea that causal constraints can help us to overcome a possible anarchy that could be a consequence of the quantum of action if considered as opposite to any causal explanation. It is clear that quantum events (both as detections and as spontaneous physical happenings) have to do with the quantum of action. Quantum features can represent the balance relative to this aspect as far as they introduce pure quantum-mechanical regularities that can be real even in the absence of any event. In other words, quantum correlations keep QM as a deterministic theory when only probabilities of events can be calculated.

Is the Quantum State Real?

If the quantum state consists in general of quantum features as correlations among possible components, we can wonder whether the notion of state as such denotes some kind of physical reality or is purely epistemic (on the outline of Wigner and Bohr's epistemic interpretation: Sects. 3.1.1 and 3.3.2). A recent and extraordinary result proves that the epistemic interpretation is in fact untenable.[50] This is also known as the *PBR theorem* after the names of the three scholars (M. Pusey, J. Barrett and T. Rudolph) who have proved this result. Suppose two devices that can prepare two quantum systems in either of the two non-orthogonal states $|0\rangle$ and $|+\rangle$ (see Eq. (3.82a)). Suppose now that their probability distributions μ_0 and μ_+ overlap, so that there is a region in which this happens with a probability $\wp > 0$. This means that in such a region, the state is compatible with any of the following four possibilities: $|0\rangle \otimes |0\rangle$, $|0\rangle \otimes |+\rangle$, $|+\rangle \otimes |0\rangle$, $|+\rangle \otimes |+\rangle$. According to the epistemic interpretation, any of these states is possible. Now, consider a measurement of these two systems. Here, we shall not make use of the ordinary Bell states (3.170) but of an appropriate combination of them (at least $|0\rangle$ or $|+\rangle$ is present in any component of the basis vectors):

$$|0\rangle\,|+\rangle - |-\rangle\,|0\rangle = |-\rangle\,|1\rangle + |1\rangle\,|+\rangle = \frac{1}{\sqrt{2}}\,(|01\rangle + |10\rangle) = \left|\Psi^{+}\right\rangle, \tag{5.9a}$$

$$\frac{1}{\sqrt{2}}\,(|0\rangle\,|-\rangle + |1\rangle\,|+\rangle) = \frac{1}{2}\,(|00\rangle - |01\rangle + |10\rangle + |11\rangle) = \frac{1}{\sqrt{2}}\left(\left|\Phi^{+}\right\rangle - \left|\Psi^{-}\right\rangle\right), \tag{5.9b}$$

$$\frac{1}{\sqrt{2}}\,(|+\rangle\,|1\rangle + |-\rangle\,|0\rangle) = \frac{1}{2}\,(|00\rangle + |01\rangle - |10\rangle + |11\rangle) = \frac{1}{\sqrt{2}}\left(\left|\Phi^{+}\right\rangle + \left|\Psi^{-}\right\rangle\right), \tag{5.9c}$$

$$\frac{1}{\sqrt{2}}\,(|+\rangle\,|-\rangle + |-\rangle\,|+\rangle) = \frac{1}{\sqrt{2}}\,(|00\rangle - |11\rangle) = \left|\Phi^{-}\right\rangle, \tag{5.9d}$$

[50] Pusey et al. (2012).

where

$$|0\rangle\,|+\rangle = \frac{1}{\sqrt{2}}\left(|00\rangle + |01\rangle\right),\ |+\rangle\,|1\rangle = \frac{1}{\sqrt{2}}\left(|01\rangle + |11\rangle\right), \tag{5.10a}$$

$$|0\rangle\,|-\rangle = \frac{1}{\sqrt{2}}\left(|00\rangle - |01\rangle\right),\ |-\rangle\,|1\rangle = \frac{1}{\sqrt{2}}\left(|01\rangle - |11\rangle\right), \tag{5.10b}$$

$$|-\rangle\,|0\rangle = \frac{1}{\sqrt{2}}\left(|00\rangle - |10\rangle\right),\ |1\rangle\,|+\rangle = \frac{1}{\sqrt{2}}\left(|10\rangle + |11\rangle\right), \tag{5.10c}$$

$$|+\rangle\,|-\rangle = \frac{1}{2}\left(|00\rangle - |01\rangle + |10\rangle - |11\rangle\right),\ |-\rangle\,|+\rangle = \frac{1}{2}\left(|00\rangle + |01\rangle - |10\rangle - |11\rangle\right). \tag{5.10d}$$

The new orthonormal basis can be synthetically rewritten as

$$|\,\Xi_1\rangle := \left|\,\Psi^+\right\rangle, \tag{5.11a}$$

$$|\,\Xi_2\rangle := \frac{1}{\sqrt{2}}\left(\left|\,\Phi^+\right\rangle - \left|\,\Psi^-\right\rangle\right), \tag{5.11b}$$

$$|\,\Xi_3\rangle := \frac{1}{\sqrt{2}}\left(\left|\,\Phi^+\right\rangle + \left|\,\Psi^-\right\rangle\right). \tag{5.11c}$$

$$|\,\Xi_4\rangle := \left|\,\Phi^-\right\rangle. \tag{5.11d}$$

Let us call $\{|\,\Xi_1\rangle\,, |\,\Xi_2\rangle\,, |\,\Xi_3\rangle\,, |\,\Xi_4\rangle\}$ a PBR basis. Now, outcome $|\,\Xi_1\rangle$ is orthogonal to $|0\rangle \otimes |0\rangle$, hence quantum theory predicts that it has probability zero when the quantum state is $|0\rangle \otimes |0\rangle$. Similarly, outcome $|\,\Xi_2\rangle$ has probability zero if the state is $|0\rangle \otimes |+\rangle$, the same for $|\,\Xi_3\rangle$ if the state is $|+\rangle \otimes |0\rangle$, and finally $|\,\Xi_4\rangle$ has a zero probability if the state is $|+\rangle \otimes |+\rangle$. This leads immediately to the desired contradiction. At least $\wp$ of the time, "the measuring device is uncertain which of the four possible preparation methods was used, and on these occasions it runs the risk of giving an outcome that quantum theory predicts should occur with probability 0." As stressed by the authors, nothing has been said about the value of the probability $\wp$ per se to arrive at this contradiction. The conclusion drawn in this study is that "quantum state is a physical property of the system", although I prefer to formulate this result by saying that our description of the quantum state must have an ontological substrate that does not depend on the way in which we describe it. I stress that the quantum correlations again restrict the range of the possible outcomes, so that the 'ontological' nature of the quantum state goes together with the 'ontological' nature of quantum features. Since such an ontological substrate makes certain outcomes possible and excludes other ones, it could be considered as a collections of expectations.[51]

[51] This can be brought in harmony with the idea that the quantum state is prescriptive (Healey 2017, Sect. 4.5).

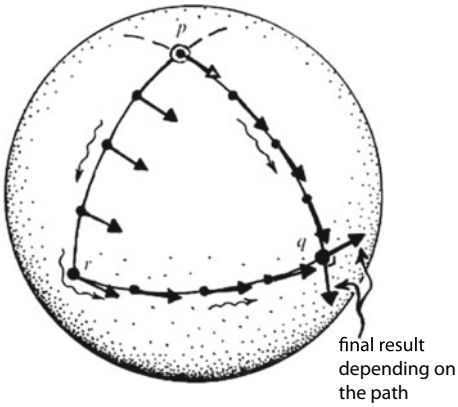


Fig. 5.6 Parallel transport. It is evident that, by transporting a vector from point p to point q along the shorter path gives a different result relative to the longer path going through r. Adapted from Penrose (2004, p. 297)

Geometric Phase

Quantum feature covers the whole range of the many quantum non-local interdependencies, like superposition, entanglement, or the Aharonov–Bohm effect.[52] About the latter effect, in Sect. 3.3.4 I have reminded of the circumstance that an isolated electromagnetic field can have still influence on the trajectory (and in general on the behaviour) of particles, where classically there is none. This shows how relevant is the phase when considering the evolution of quantum systems. This could be barely understood without assuming the reality of quantum features if we think that no physical potential is necessary here. In fact, I have reminded of the geometric phase as a perfect explanation of this. We can model the behaviour of the quantum particles by making use of the *parallel transport*. This consists in starting with a vector tangent to some point on a surface and transporting it along a path in such a way that it keeps the same orientation and direction. What happens is that, in curved space (as different from ordinary Euclidean flat space), the final orientation of the transported vector depends on the path followed (Fig. 5.6).[53] The concept of geometric phase (see Fig. 5.7) fits very well with the notion of quantum features, which as all correlations, display structural and formal aspects, as pointed out in Sect. 3.4.1. Moreover, as mentioned, this also shows quantum systems are sensitive to the geometry of their environment: it is an aspect of their sensitivity to physical conditions, which fits well with the idea that a quantum state can be expressed as a collection of expectations.

[52] Aharonov and Bohm (1959).

[53] Misner et al. (1970, Sect. 8.3).

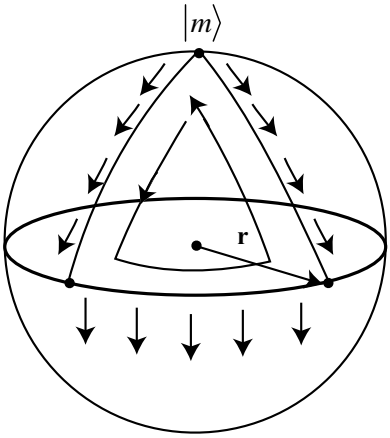


Fig. 5.7 The geometric phase takes its name from the fact that its formalism bears elements of differential geometry, in particular of the concept of measuring the curvature of a surface by parallel transport of a tangent vector. Here, we transport a vector **r** (represented by the small arrows) that is tangent to a curved surface along a path on this surface ($|m\rangle$ is an eigenstate of the energy as function of **r**). After a cycle on the closed path, the vector is not identical to the initial one: the two vectors have a non-vanishing angle between them, which is a measure for the curvature of the surface. Adapted from Auletta et al. (2009, p. 501)

5.1.3 Quantum Features and Discord

Quantum Discord

The difference between classical and quantum correlations has been quantified, so that the total information shared by two entangled systems (Sects. 3.2.2 and 3.4.1–3.4.2) can be divided into two parts: the classical part and the quantum part that is called *quantum discord*.[54] As mentioned, correlations can be expressed in terms of mutual information: Eq. (3.46) represents indeed a way to formalise this notion. I have not mentioned so far that, classically speaking, when considering sets of signals (or variables) X and Y, we have two equivalent formulations for mutual information:

$$I(X:Y) = S(X) + S(Y) - S(X,Y) \text{ and } I(X:Y) = S(Y) - S(Y|X), \quad (5.12)$$

where, as customary, $S(X)$ and $S(Y)$ are the entropies of X and Y, respectively, $S(X,Y)$ is the total joint entropy of the two systems, and (see also Eq. (3.155))

$$S(Y|X) = -\sum_{x\in X}\sum_{y\in Y} \wp(x,y)\lg\wp(y|x) \quad (5.13)$$

[54]The reference paper is Ollivier and Zurek (2001). See also Auletta and Wang (2014, Sect. 11.7).

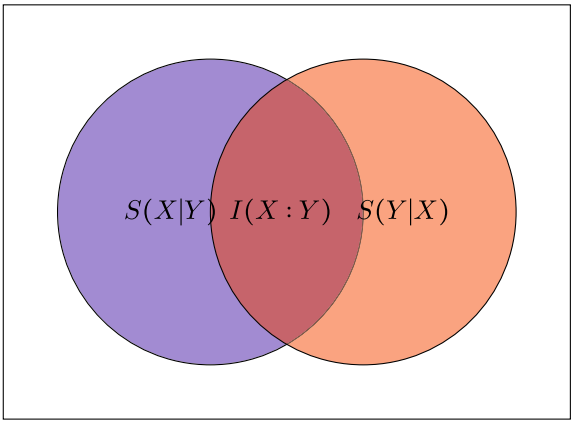


Fig. 5.8 Graphic representation of mutual information by means of the Venn diagrams (after the name of the British logician and philosopher John Venn (1834–1923)). The entropy $S(X)$ is the whole disc with subsets denoted by blue and purple colours, while $S(Y)$ is the whole disc with subsets denoted by red and purple colours. It is easy to verify that $I(X:Y) = S(X) - S(X|Y) = S(Y) - S(Y|X)$

is the conditional entropy of Y on X, representing how much the output Y is independent from the input X, being in this way a measure of the *equivocation* in the reception of the signal X (Fig. 5.8). Therefore, in accordance with the previous analysis, mutual information represents the amount of order between two coupled systems (when we have taken away from the entropy of a system the disorder represented by equivocation relative to a reference system, as displayed by the second of Eqs. (5.12)).

Now, quantum mechanically it turns out that the conditional entropy is not well defined when no potential outcomes are singled out that would allow us to specify the input. Recall, in fact, that quantum systems are always in some superposition state. Thus, the conditional von Neumann entropy $S_{\mathrm{VN}}(\mathcal{A}|\mathcal{S})$ of an apparatus $\mathcal{A}$ on a system $\mathcal{S}$ requires us to specify the state of $\mathcal{A}$ *given* the state of $\mathcal{S}$, but such a statement in quantum theory is ambiguous until the to-be-measured set of eigenstates of $\mathcal{S}$ is not selected. This is what makes coupling or the premeasurement step necessary, as already pointed out in Sects. 3.1.2–3.1.3, 3.2.3, and 3.3.2. Of course, this is ultimately due to the non-commutativity of quantum observables. Therefore, we need to substitute this quantity by a new kind of conditional entropy. To this purpose, let us consider the set of one-dimensional projectors $\{\hat{P}_j^S\}$ for the system, where the label j distinguishes different outcomes (eigenstates of some observable) of this measurement (and $\{|\,j\rangle\,\}$ is an orthonormal basis for the system). The state of $\mathcal{A}$ after the outcome corresponding to $\hat{P}_j^S$ occurs is given by (see also Eqs. (3.14) and (1.396))

$$\hat{\rho}_{A|\hat{P}_j^S} = \frac{\hat{P}_j^S \hat{\rho} \hat{P}_j^S}{\wp_j}, \tag{5.14}$$

where $\hat{\rho}$ is the density matrix of the composite system comprehending $\mathcal{S}$ and $\mathcal{A}$, while (Eq. (1.379))

$$\wp_j = \mathrm{Tr}(\hat{P}_j^S \hat{\rho}). \tag{5.15}$$

Then, instead of considering the conditional entropy $S_{\mathrm{VN}}(\hat{\rho}_{A|\hat{P}_j^S})$, we take into account its weighted average:

$$S_{\mathrm{VN}}(\mathcal{A}|\{\hat{P}_j^S\}) = \sum_j \wp_j S_{\mathrm{VN}}(\hat{\rho}_{A|\hat{P}_j^S}), \tag{5.16}$$

which means that $\mathcal{A}$ must be taken as conditional upon the whole set of projectors denoting all possible values (and therefore possible detection events) when measuring a certain observable of $\mathcal{S}$. For instance, in the case in which the whole is in a singlet state, $\mathcal{A}$ needs to be in the state "down" if $\mathcal{S}$ is "up" and vice versa, so that we have a set of joint events. This is another manifestation of the fact that quantum systems can pair code alphabets but do not share them (Sect. 3.3.7). This implies that the mutual information representing entanglement (as in Eq. (3.51)) is composed by the sum of a classical and quantum part (see again Fig. 5.3):

$$I(\mathcal{A}{:}\mathcal{S}) = \mathcal{C}(\mathcal{A}{:}\mathcal{S})_{\{\hat{P}_j^S\}} + \mathcal{Q}(\mathcal{A}{:}\mathcal{S})_{\{\hat{P}_j^S\}}, \tag{5.17}$$

where the classical part is defined as

$$\mathcal{C}(\mathcal{A}{:}\mathcal{S})_{\{\hat{P}_j^S\}} := S_{\mathrm{VN}}(\mathcal{A}) - S_{\mathrm{VN}}(\mathcal{A}|\{\hat{P}_j^S\}), \tag{5.18}$$

and the quantum discord as

$$\begin{aligned} \mathcal{Q}(\mathcal{A}{:}\mathcal{S})_{\{\hat{P}_j^S\}} &:= I(\mathcal{A}{:}\mathcal{S}) - \mathcal{C}(\mathcal{A}{:}\mathcal{S})_{\{\hat{P}_j^S\}} \\ &= S_{\mathrm{VN}}(\mathcal{S}) + S_{\mathrm{VN}}(\mathcal{A}|\{\hat{P}_j^S\}) - S_{\mathrm{VN}}(\mathcal{S},\mathcal{A}). \end{aligned} \tag{5.19}$$

Note that the classical part has the same form of the second expression (5.12) and, displaying antisymmetry, is in fact equal to the Holevo quantity (3.86), when considering that

$$S_{\mathrm{VN}}(\mathcal{A}) = S\left(\sum_j \hat{\rho}_{A|\hat{P}_j^S}\right). \tag{5.20}$$

However, when comparing the first expression (5.12) and Eq. (5.19), we see that we have essentially replaced the entropy of Y or $\mathcal{A}$ (as expressed by $S(Y)$ or $S(\mathcal{A})$) by the quantum conditional entropy of $\mathcal{A}$ on $\mathcal{S}$ as described by Eq. (5.16). When the quantum discord $\mathcal{Q}(\mathcal{A}{:}\mathcal{S})$ is not zero, it indicates the presence of quantum correlations or features. This may be also considered the quantumness of correlations, a term introduced by Zurek and Ollivier themselves. We may say that physical systems show the more pure quantumness, the more they are not separated.

A Terminological Clarification

A terminological clarification seems necessary here. Both terms, namely, quantum discord and quantumness, hint at the same issue as the notion of quantum features does. However, the discord is a *measure* of the quantum part of entanglement and

similarly the notion of quantumness. This means that the quantum discord describes quantum features in mathematical terms completely in that context. On the other hand, we can take quantum features as representing the content of the more vague notion of quantumness for *every* pure quantum state, and therefore it is the term that I prefer. In conclusion, we can consider the notion of quantum discord as the formal counterpart (in terms of information) of the ontological notion of quantum features in the case of entanglement, while the notion of features, in its generality, hints at what is *common and typical* quantum in all non-local and 'interference' quantum effects.

The Quantum Environment

The distinction between a classical and a quantum part of information (and correlations) allows us a deeper understanding of measurement and in particular of the role of the environment.[55] In fact, I have said that measuring a quantum system locally means to deal with a system that is simultaneously entangled with the environment (Sect. 3.1.3). We have also seen that orthogonality of the measured states is a necessary condition for getting one of the eigenvalues of the measured observable, although additional environmental conditions are also necessary. In fact, each of the possible expansions (code alphabets) of the system's state (representing the different observables) can be considered as entangled with a part of the environment, a so-called *fragment* of the whole universe. In other words, there is a quantum channel (entanglement) between each degree of freedom of the premeasured state and a fragment of the universe, what is clear evidence for the real effect of premeasurement, although being still an operation that is reversible. When measuring, a correlation is established between the measured system $\mathcal{S}$ and a fragment $\mathcal{F}$ of the environment $\mathcal{E}$, such that the environment's fragments acquire information about $\mathcal{S}$ (Fig. 5.9). What the fragment $\mathcal{F}$ and $\mathcal{S}$ know about each other is quantified by the mutual information

$$I(\mathcal{S} : \mathcal{F}) = S(\mathcal{S}) + S(\mathcal{F}) - S(\mathcal{S}, \mathcal{F}), \tag{5.21}$$

which measures the total correlations present. I have suppressed here the subscript VN for the sake of notation. One of the main quantities of interest is the typical fragment size needed by an observer to learn about the system. That is, a typical fragment size $\mathcal{F}_\delta$ contains $1 - \delta$ of the missing information $S(\mathcal{S})$ about the system:

$$I(\mathcal{S} : \mathcal{F}) \geq (1 - \delta)S(\mathcal{S}) \tag{5.22}$$

decreasing the remaining entropy to $\delta S(\mathcal{S})$. The parameter δ denotes here the information *deficit*, which quantifies the error tolerance of the observers. For this reason, there is also a minimal fragment size needed by an observer to learn about the system observed. Zwolak and Zurek show that the Holevo quantity (3.86) or (5.18) for

[55] I synthesise in the following the main arguments of Zwolak and Zurek (2013).

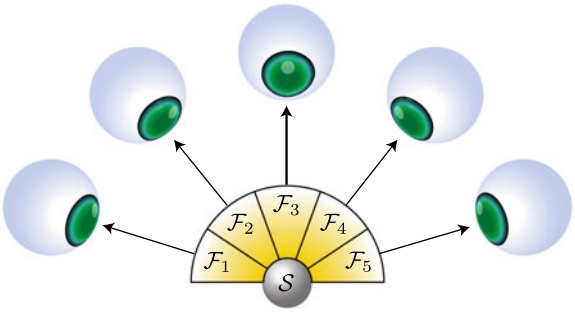


Fig. 5.9 A system, $\mathcal{S}$, interacts with an environment, $\mathcal{E}$, composed of many different fragments, $\mathcal{F}_k$. While the system decoheres (as indicated by the loss of phase coherence, which is characterised by the decay of the off-diagonal elements of the density matrix in the pointer basis), the environment fragments each acquire information about $\mathcal{S}$ that can then be transmitted to observers. To learn about the state of the system, each observer could intercept a different fragment. Adapted from Zwolak and Zurek (2013)

observables different from the locally selected one decreases depending on the degree of 'misalignment' between them and the eigenstates of the pointer observable. To show this, we rewrite Eq. (5.21) in accordance with Eq. (5.17) as

$$I(\mathcal{S}:\mathcal{F}) = \chi\left(\{\hat{P}_j^S\}:\mathcal{F}\right) + \mathcal{Q}\left(\{\hat{P}_j^S\}:\mathcal{F}\right), \tag{5.23}$$

where

$$\chi\left(\{\hat{P}_j^S\}:\mathcal{F}\right) = S\left(\sum_j \wp_j \hat{\rho}_{\mathcal{F}|j}\right) - \sum_j \wp_j S_{\mathcal{F}|j}. \tag{5.24}$$

Note that we have $S_{F|j} = S(\mathcal{F}|\hat{P}_j^S)$ and, being $\hat{\rho}$ is the initial density matrix of the compound system $\mathcal{S}$–$\mathcal{F}$, also

$$\hat{\rho}_{F|j} = \frac{\hat{P}_j^S \hat{\rho} \hat{P}_j^S}{\wp_j}, \tag{5.25}$$

in agreement with Eq. (5.14). Note that both the discord and the Holevo quantity in Eq. (5.23) represent a mutual information between the fragment $\mathcal{F}$ and the whole set of projectors denoting the possible eigenstates of a particular observable. This shows that whenever objective, classical information about a system is present, quantum information, as measured by the discord, about this system is out of reach for observers without access to nearly the whole environment and the system—a situation that can occur, at best, only in controlled laboratory experiments. In fact, the Holevo quantity and the quantum discord on complementary fragments of the environment (i.e. $\mathcal{F}$ and the rest of $\mathcal{E}$: $\mathcal{E}/\mathcal{F}$) are related by Eq. (5.19) or Eq. (5.23), replacing $\mathcal{F}$ by $\mathcal{E}/\mathcal{F}$:

$$\begin{aligned}\mathcal{Q}\left(\{\hat{P}_j^S\}:\mathcal{E}/\mathcal{F}\right) &= I(\mathcal{S}:\mathcal{E}/\mathcal{F}) - \chi\left(\{\hat{P}_j^S\}:\mathcal{E}/\mathcal{F}\right)\\ &= S(\mathcal{S}) - S(\mathcal{S}\mathcal{E}/\mathcal{F}) + S(\mathcal{E}/\mathcal{F}|\{\hat{P}_j^S\})\\ &= S(\mathcal{S}) - S(\mathcal{S}\mathcal{E}/\mathcal{F}) + \sum_j \wp_j S_{E/F|j}\\ &= S(\mathcal{S}) - S(\mathcal{F}) + \sum_j \wp_j S_{F|j}\\ &= S(\mathcal{S}) - \chi\left(\{\hat{P}_j^S\}:\mathcal{F}\right), \end{aligned} \tag{5.26}$$

where, since the conditional state of $\mathcal{E}$ is pure and $\mathcal{E}/\mathcal{F}$ and $\mathcal{F}$ are a bipartite split of such a state of $\mathcal{E}$, for the Araki–Lieb inequality (3.76) we must have $S_{E/F|j} = S_{F|j}$. The same for $S(\mathcal{S}\mathcal{E}/\mathcal{F}) = S(\mathcal{F})$. In other words, formally speaking, there is here a complementarity between quantum discord (globality) and Holevo quantity (locality), which represents the maximal amount of classical information that we can get locally (that is directly accessible) when measuring a quantum system: increasing the redundant (classical) information stored in a small fragment of the universe decreases the quantum information shared with the much larger fragment represented by the rest of the universe, as displayed in Fig. 5.10. Note that, if the fragment $\mathcal{F}$ is sufficiently small, it is in fact equivalent to the Holevo quantity. This dynamics is *global–local*. The complementarity displayed in Eq. (5.26) should then be interpreted as expressing not only a complementarity but also a fundamental *interconnection* between these two aspects.

Resuming, only those copies in the environment survive that are perfect. This happens when the scalar products of orthogonal components of the corresponding 'record states' vanish; and those copies turn out to be the measured observable's eigenstates (given the local experimental setup), as anticipated in Sect. 3.1.3. Thus, when we select information locally (through detection) we switch off all channels apart that between the selected observable and a specific fragment (Sect. 5.1.1). This fragment is the local accessible environment. The other ones represent inaccessible environment: this is the precise meaning of having downloaded information into the environment. Here, the so-called quantum Darwinism should be in action. With this term it is understood the (universal) spontaneous selection of observables during decoherence. In this way we have successfully embedded a local irreversible dynamics in a global reversible one, according to the results of Sect. 4.1.

Bohm's Implicate Order

I would like now to comment on some scholars who could give some support to the interpretation advanced here. In particular, Bohm's idea of *implicate order* (as expressing the universal network of quantum systems) that manifests itself locally as

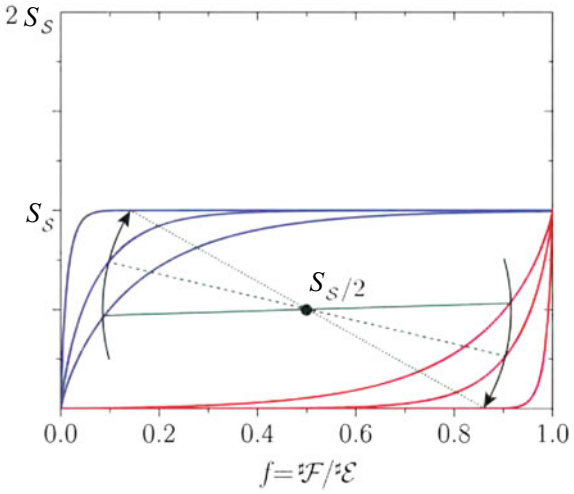


Fig. 5.10 As the Holevo quantity—quantum (represented by the blue curves) increases—i.e. as the classical information transmitted by the environment increases—the quantum discord (represented by the red curves) on the opposite side of the axis decreases. Since the discord is monotonically increasing with f, this means that the quantum information about $\mathcal{S}$ is pushed into correlations with the global environment as redundant (classical) information is increased. Thus, a state which has redundant information will have quantum information encoded in the environment. Adapted from Zwolak and Zurek (2013)

explicate (and partial) order is crucial.[56] Indeed, due to randomness of local quantum events we always catch, out of our local physical context, a partial aspect of the universal order (what is another way to say that we lose part of the information). Bohm made use of the term *metamorphosis* for denoting the particular dynamics that exists between implicate and explicate order, what is convergent with the dynamics global–local synthesised here.

5.1.4 Again on Complementarity

Entrenchment and Complementarity

The dynamical entrenchment global–local can be brought in harmony with the complementarity principle, that is, we expect that the same entrenchment will be displayed by the traditional complementarity wave–particle, and even that the former is in fact a more rigorous formulation of the latter. However, this demands a substantial reformulation of the complementarity principle. The traditional formulation of the complementarity principle, due to N. Bohr, dealing with a complementarity between

[56] Bohm (1980, pp. 147–71).

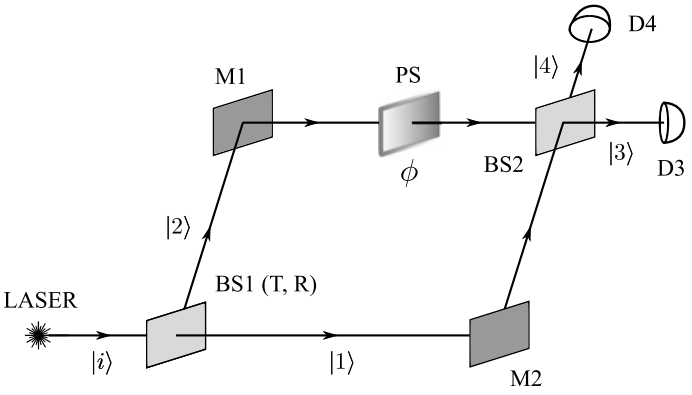


Fig. 5.11 Mach–Zehnder interferometer with a variable beam splitter BS1 (function of T and R), showing a smooth complementarity between wave-like and particle-like behaviour. Adapted from Auletta and Wang (2014, p. 96)

wave-like and corpuscular behaviours (Sect. 1.3.3), is in fact unsatisfactory for two different reasons[57]:

- It uses a terminology that is taken from classical physics but results inadequate for the quantum case (see the comments in Sect. 3.1.4). Actually, the wave-like nature of quantum entities is not of classical type. We might recall here in particular the phenomenon of self-interference (Sects. 1.2.2 and 5.1.2). Moreover, the corpuscular nature of quantum entities shows a kind of randomness and discontinuities that are unknown classically (Sect. 5.1.1).
- It assumes that these are two sharply alternative behaviours while, as a matter of fact, quantum systems display a continuous range of different behaviours between these two extremes, and therefore a basic entrenchment.

Interferometry with Variable Beam Splitter

This is shown by the so-called Greenberger–Yasin equality,[58] after the names of the American physicists Daniel Greenberger and Allaine Yasin. Let us consider the arrangement of the Mach–Zehnder interferometer displayed in Fig. 5.11. Compared with the previous experiments, the novelty is that now we have that the first beam splitter is no longer a 50–50 one. Instead, suppose that the beam splitter transmits and reflects the incoming photon with probabilities given by T^2 and R^2, respectively, where we shall assume that the respective transmission and reflection coefficients $T, R \in \mathbb{R}$ for the sake of simplicity (the essence of the following argument remains

[57] These arguments have been synthesised in Auletta and Torcal (2011). See also Auletta and Wang (2014, Sect. 5.5).

[58] Greenberger and Yasin (1988).

untouched by this assumption). I have also relabeled the states associated with the lower and upper paths as $|1\rangle$ and $|2\rangle$, respectively (and the states associated with the paths leading to the detectors as $|3\rangle$ and $|4\rangle$, respectively). In this case, for a photon initially in the state $|i\rangle$ $(= |1\rangle)$ the action of BS1 and PS can be written as

$$\begin{aligned} |i\rangle &\overset{\text{BS1}}{\longmapsto} \mathrm{T}\,|1\rangle + \mathrm{R}\,|2\rangle \\ &\overset{\text{PS}}{\longmapsto} \mathrm{T}\,|1\rangle + e^{i\phi}\mathrm{R}\,|2\rangle\,. \end{aligned} \tag{5.27}$$

Making use of the notorious expressions (2.30) for the action of BS2

$$|1\rangle \overset{\text{BS2}}{\longmapsto} \frac{1}{\sqrt{2}}\left(|3\rangle + |4\rangle\right) \quad \text{and} \quad |2\rangle \overset{\text{BS2}}{\longmapsto} \frac{1}{\sqrt{2}}\left(|3\rangle - |4\rangle\right), \tag{5.28}$$

we have

$$\mathrm{T}\,|1\rangle + e^{i\phi}\mathrm{R}\,|2\rangle \overset{\text{BS2}}{\longmapsto} \frac{1}{\sqrt{2}}\left[\mathrm{T}\left(|3\rangle + |4\rangle\right) + e^{i\phi}\mathrm{R}\left(|3\rangle - |4\rangle\right)\right]. \tag{5.29}$$

Upon collecting the terms, we find the final state of the photon after leaving BS2 is given by

$$|f\rangle = \frac{1}{\sqrt{2}}\left[\left(\mathrm{T} + e^{i\phi}\mathrm{R}\right)|3\rangle + \left(\mathrm{T} - e^{i\phi}\mathrm{R}\right)|4\rangle\right], \tag{5.30}$$

which can be considered a generalisation of Eq. (2.31). Hence, the final detection probabilities at D3 and D4 are, respectively, given by

$$\wp_3 = \frac{1}{2}\left(1 + 2\mathrm{TR}\cos\phi\right) \quad \text{and} \quad \wp_4 = \frac{1}{2}\left(1 - 2\mathrm{TR}\cos\phi\right). \tag{5.31}$$

In obtaining the above expressions, I have used the normalisation condition $\mathrm{T}^2 + \mathrm{R}^2 = 1$.

Path Predictability

We may choose various values for the coefficients T and R of BS1 as well as for the phase shift ϕ of PS, provided that the condition $\mathrm{T}^2 + \mathrm{R}^2 = 1$ is satisfied. Let us now introduce a simplification and consider the case in which the phase shift $\phi = 0$ (in radians). In the limiting case in which $\mathrm{T} = 0$ (and $\mathrm{R} = 1$), the photon is totally reflected and we know for sure that it has taken the upper path. In this case, the first beam splitter behaves as an ordinary mirror. In the opposite limiting case in which $\mathrm{R} = 0$ (and $\mathrm{T} = 1$), the photon is fully transmitted and therefore it takes the lower path with certainty. In this case it is as if we had taken the first beam splitter away. It is evident that in these limiting cases we have situations with full predictability of the path followed by the photon, which are analogous to the presence of an obstacle

discussed in Sect. 2.1.2. We can quantify *path predictability* by the absolute value of the difference between the probabilities for the photon to be reflected and transmitted, that is,

$$\hat{\mathcal{P}} = \left|\mathrm{T}^2 - \mathrm{R}^2\right|, \tag{5.32}$$

where the absolute value is justified by the fact that we do not know *a priori* which one of the probabilities (either T^2 or R^2) is larger. The reason for the above formula is that $\hat{\mathcal{P}}$ has the maximum value 1 when one of the two probabilities is 1 (and the other is 0), and the minimum value 0 when the two probabilities are equal, that is, $\mathrm{T}^2 = \mathrm{R}^2 = \frac{1}{2}$. Then, a pure particle-like (or local) behaviour has $\hat{\mathcal{P}}$ maximised, whereas a pure wave-like (or non-local) behaviour has $\hat{\mathcal{P}}$ minimised. In other words, for localising the photon we need to be able to *discriminate* between the two paths.

The previous equation can be easily derived as follows. Considering always the case where the phase shift $\phi = 0$, the probability amplitudes (up to a global phase factor) that the photon is detected at D3 and D4, according to Eq. (5.30), are, respectively, given by

$$\vartheta_3 = \sqrt{\wp_3} = \frac{1}{\sqrt{2}}|\mathrm{T} + \mathrm{R}|, \text{ and } \vartheta_4 = \sqrt{\wp_4} = \frac{1}{\sqrt{2}}|\mathrm{T} - \mathrm{R}|. \tag{5.33}$$

These expressions are allowed since we have assumed T and R to be real (in the general case a probability amplitude is not the square root of the relative probability). The product of the two probability amplitudes gives

$$\vartheta_3\vartheta_4 = \frac{1}{2}|\mathrm{T} + \mathrm{R}||\mathrm{T} - \mathrm{R}| = \frac{1}{2}\left|\mathrm{T}^2 - \mathrm{R}^2\right|, \tag{5.34}$$

so that $2\vartheta_3\vartheta_4$ is equal to the path predictability $\hat{\mathcal{P}}$ in Eq. (5.32). Since the latter expression is maximum when one of the two probabilities (either T^2 or R^2) is zero, so is the product of the two probability amplitudes $\vartheta_3\vartheta_4$. This is again a peculiarity of quantum probability (see also Sect. 2.1.2) since classically the product of two mutually exclusive probabilities is maximum when the two probabilities are equal.

Interference Visibility

On the other hand, we can also introduce a complementary parameter to quantify the visibility of interference fringes. Since the path predictability is maximum when one of the two coefficients (either T or R) is zero, the interference visibility should be proportional to the product of the two coefficients. In other words, interference fringes are visible only in the case in which the path predictability is *not maximum*. Conventionally, the *interference visibility* is defined by

$$\hat{\mathcal{V}} = 2\mathrm{TR}. \tag{5.35}$$

In the probabilities $\wp_3$ and $\wp_4$ in Eq. (5.31), the quantity $\hat{\mathcal{V}}$ represents precisely the quantum feature that is added to the purely classical part represented by the number $\frac{1}{2}$ (in the simplest case where $\phi = 0$). It is interesting to see that the interference visibility can be extracted from the difference between the probabilities $\wp_3$ and $\wp_4$ given by Eq. (5.31), that is,

$$\begin{aligned} \wp_3 - \wp_4 &= \frac{1}{2}\left[(\mathrm{T}+\mathrm{R})^2 - (\mathrm{T}-\mathrm{R})^2\right] \\ &= 2\mathrm{TR}, \end{aligned} \tag{5.36}$$

where I have assumed again that $\phi = 0$. This is understandable since the purely classical part cancels out in the difference of the two quantum probabilities, where it is helpful here to recall Eq. (3.229) and discussion.

Greenberger–Yasin Equality

From Eq. (5.35) we see that $\hat{\mathcal{V}}$ has the maximum value 1 when the two coefficients T and R are equal, that is, $\mathrm{T} = \mathrm{R} = \frac{1}{\sqrt{2}}$, and the minimum value 0 when one of the two coefficients is 1 (and the other is 0). This, together with the corresponding behaviours of $\hat{\mathcal{P}}$, allows us to conclude that the path predictability and the interference visibility are complementary quantities in the sense that the path predictability is gained at the expense of the interference visibility, and vice versa. In other words, any distinguishability between the paths of an interferometer is at the expenses of the visibility of the interference fringes. Indeed, it is also easy to verify that

$$\begin{aligned} \hat{\mathcal{P}}^2 + \hat{\mathcal{V}}^2 &= \mathrm{T}^4 + \mathrm{R}^4 - 2\mathrm{T}^2\mathrm{R}^2 + 4\mathrm{T}^2\mathrm{R}^2 \\ &= \left(\mathrm{T}^2 + \mathrm{R}^2\right)^2 \\ &= 1. \end{aligned} \tag{5.37}$$

This important equation is called the *Greenberger–Yasin equality*,[59] and tells us that the sum of these two behaviours is 1, so that any *intermediate* behaviour is possible as well. In such a context, Eddington's notion of wavicles makes sense provided that we do not understand such a notion in terms of a classical object (Sect. 1.1.3). Note that the principle is expressed here as a relation between two *particular* observables that represent a general undulatory and localised behaviour. Since there is a smooth complementarity between these two behaviours, I take this inequality to be the formal expression of the complementary interplay between features and localisation for a single system (while for a full dynamics we need the master equation). It is another piece of evidence that we need to assign an ontological status to quantum features since they are entrenched with corpuscular behaviours that are deemed to be real: if we have a continuum of intermediate mixed behaviours, why should we assign

[59] Greenberger and Yasin (1988). A similar result can be found in Englert (1996).

reality to only one of the poles? This is also supported by the mentioned fact that any detection event, according to decoherence, never fully destroys such features but simply washes them partially out in a local context.

It is worth mentioning that this understanding of complementarity completely transforms the status of the related principle. In fact, according to Bohr, it was rather the way in which we categorically deal with the quantum world, and so had an epistemic status telling us how to correctly approach to the knowledge of quantum phenomena (what are the epistemological restrictions on our understanding of the quantum world). At the opposite, here the principle is understood in basic ontological terms without excluding the epistemic aspects either.

5.1.5 A Participatory Universe: Relativity and Quantum Mechanics

Entanglement as Independent of Space-time

In Sect. 5.1.1, we have seen a potential conflict between QM and SR about the notion of event. In fact, the issue is bigger and concerns also quantum correlations. Indeed, quantum-mechanical features are independent of space and time (or of space-time). Due to the violation of separability, it is clear that quantum correlations are unaffected by spatial distance.[60] A performed delayed-choice experiment with space-like separation of entangled photons is a clear evidence for that.[61] A similar result is true for time.[62] In other words, while classical information can be copied (at some energy cost) but only be transmitted forward in time (to a receiver in the sender's forward light cone), quantum correlations (and therefore also entanglement) cannot be copied but connect any two arbitrary points in space-time.[63] In fact, the causal effects (in a formal sense: Sects. 3.4.1–3.4.2) that correlations induce on distant systems are instantaneous, so that, when we make conditional predictions on the basis of entanglement, these cannot be understood in temporal terms. For this reason, time reversal is a property of entangled states (Sect. 1.2.5).[64] This was proved by an interesting experiment; multisimultaneity is a causal model of relativistic physics that assigns a real-time ordering to any set of events, much in the spirit of the pilot-wave picture of de Broglie (Sect. 3.3.4). Contrary to standard QM, assumption of a real-time ordering predicts a disappearance of the correlations in a Bell-type experiment when both analysers are in relative motion such that each one, in its own inertial reference frame (Sect. 2.3.2), selects first the output of the photons' measurement.

[60] As pointed out in Maudlin (1994, p. 22).

[61] Ma et al. (2013). I shall come back on this experiment.

[62] Suarez (2001). This possibility was anticipated in Auletta (2000, Sect. 46.4). See also (Maudlin 1994, p. 23).

[63] As pointed out in Bennett and Shor (1998).

[64] As stated in Gottesman and Chuang (1999).

This prediction was tested but no disappearance of the correlations was observed, in agreement with quantum physics.[65] In fact, the sequence of two measurements on entangled systems can give the same result *both* locally and jointly (by comparing them) *whatever succession* of detections we assume to be true.

If this is true, it follows that the structure of space-time is less fundamental than quantum correlations and is likely to be an emergent result of quantum-mechanical dynamics, although in ways that need still to be explored. The above conclusion was drawn by J. Wheeler,[66] who also recalls that at Planck scale (that has dominated the first segment of the history of our universe) the concepts of *before* and *after* are deprived of any meaning, although we are confident that this stage of the universe is governed by quantum-mechanical laws. Thus, it is likely that also quantum events are somehow independent of the space-time network described by relativity. To appreciate these statements fully, it is important to recall that Wheeler was one of the most authoritative scholars in general relativity (e.g. the notion of black hole is due to him), and, as recalled by the German philosopher of science Adolf Grünbaum,[67] shifted from the initial position of considering space-time as the basic ontological stuff of our universe to regard them as relations extrapolated from our experience of quantum events.[68]

On Relativistic and Quantum Events Again

This suggests to deepen the examination started in Sect. 5.1.1 about the relation between relativistic and quantum events and focus on the issue of localisation. Let us consider proper time as a relativistic invariant. A system traces a world line in space-time, that is, a trajectory that describes its space-time evolution (Fig. 2.13, Sect. 2.3.2). I recall that its proper time is the time measured by a clock that is carried by that system, which makes appear the world line of that system straight if it is a free point particle. In other words, different observers should agree about the time elapsed between two events along that world line *on that* clock, even if they disagreed on the interval measured on *their own* clocks: only two clocks having identical world lines between two fixed events measure the same elapsed time between those events.[69] Obviously, these observers would agree on the interval as far as it can be observed, and whose results can be recorded and therefore made accessible to further observers; this, however, has no significance for any isolated and point-like event (not constituting an interval by definition).

Therefore, any "here–now" was coincident with the event itself that determines that localisation. However, dealing with the latter implies to deal with a space-time

[65] Stefanov et al. (2002).

[66] Wheeler (1990). See also Verlinde (2011) for a model, on which I shall come back.

[67] Grunbaum (1973, p. 729).

[68] The transactional interpretation of QM reaches a similar conclusion about the nature of events (Kastner 2013).

[69] Geroch (1978, p. 79).

interval and no longer with the original event itself.[70] In fact, any observation event, or observation of the original event, is *displaced* in space-time relative to the latter. I recall that the assumption that source events and their observation would coincide was rejected by Bohr in a quantum-mechanical context, as I have already mentioned in Sect. 3.3.2. Therefore, I think that we should carefully distinguish between two issues:

- The issue of *fixing* a proper space-time, where clearly observation and observed event need to be assumed to be coincident, as recalled in Sect. 5.1.1. As stated by P. Bridgman, we need to accept this as an ultimate simple thing.[71]
- The issue that the event that has been observed is exhausted by this act of observation. In fact, if quantum events are more fundamental than the structure of space-time, then they should be considered also as independent of the fixation of the local space-time.

If this is correct, then we can admit that the original (quantum) events have only *probability* to happen due to their randomness, but, once happened, the consequent effects on their environment, and thus their *local determination*, allow us to assign to them a 'point' in space-time. And we call the latter a *relativistic event*. I stress again that to get determined events is a process of *determination* (Sect. 3.2.3). In this way, any relativistic (or coincidence) event is a coarse graining of the point-like, totally fine-grained, quantum events. It is true that the notion of coincidence event would be not possible without an assumed continuity, so that such a coincidence is rather a limit (Sect. 2.2.1), while the quantum-mechanical event represents a basic discontinuity, so that, as it is often said, relativity and QM have different approaches to physics. Nevertheless, there is no conflict here, at the opposite of what it is also often assumed.

Now–Here

What appears incompatible with the sketched solution is the hypothesis of the block universe (Sect. 2.3.3).[72] It is in fact difficult to consider relativistic events as coarse-graining of quantum events when we assume that the predetermination of the former be not so compatible with the randomness of the latter. In fact, if we interpret the statement that something exists *now* (or now–here) as meaning an a-temporal mode of actual being, all relativistic events actually occurring would be predetermined, and this not as possible but as actually *occurring* events. In other words, to be temporally (or space-timely) determined should not imply being predetermined in any way (if not in the quantum sense of pure possibilities and according to probabilistic deter-

[70] See also Rovelli (2004, pp. 49–50); Grunbaum (1973, pp. 322–24).

[71] Bridgman (1927, p. 75).

[72] I summarise here the arguments exposed in Auletta (2011, Sect. 3.3.4).

minism, which is nonetheless independent of space-time)[73]: quantum mechanically any determination is always *ex post*, i.e. after the selection of a determined context or an occurred local interaction (Sect. 5.1.1).

Past and Future

There are also specific relativistic arguments against the idea that the whole universe is a single block of space-time relations, in which obviously there is no longer a time arrow. Even in a classical, Laplacian world, from a limited region of space we can deduce (retrodict) a great deal of the past history of the universe, while for performing similarly powerful predictions about future events we must consider initial conditions over a very wide region of space, what justifies the observed time asymmetry (Sect. 2.4.3).[74] The reason seems to be quite simple (Sect. 2.3.2): the events (as far as they happened) belonging to the absolute past of an event E could be (more or less) uniquely specified in records which are *part of that localised space-time region*, whereas events belonging to the absolute future of E are not (due to the spread of any signal), what does not exclude that they could be unpredictable events in the sense of QM. This suggests that it is the kind of dynamical effects that a certain event induces on its surroundings in the presence of correlations that determines the perceived time asymmetry and the related openness of dynamical processes.

A Participatory Universe

The mentioned considerations led Wheeler, in opposition to the idea of a block universe, to consider the universe as *participatory* and see the laws of nature as *emerging* from quantum events.[75] What does it mean that the universe is participatory? It means that everything that happens is essentially entrenched with the rest of the world. Specifically, any detection event is indissolubly related to a universal relationship, in terms of both causal constraints and causal effects. It is like the universe observing itself (Fig. 5.12). However, how can we interpret Wheeler's view? We face here the same dilemma of Sect. 5.1.1. In fact, we can deal with this in two different ways:

- We could consider this in Kantian or idealistic terms as the mind and only the observer's mind bringing to expression reality, or at least understand this connection in pure epistemic terms (Sect. 3.1.1). I agree that the mind can be considered like the universe taking awareness of itself, a view on which also the quantum objectivist Schrödinger would have agreed (Sect. 3.2.1). In fact, according to him,

[73] Consider in this context the difficulties with the notion of extended time (Bridgman 1927, pp. 76–77).

[74] This argument was presented by the British-Australian philosopher J. Smart (1920–2012) (Smart 1954). See also Lockwood (2005, Chap. 11).

[75] Wheeler (1983).

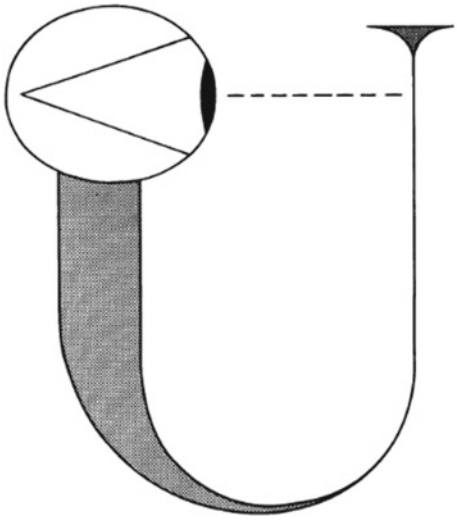


Fig. 5.12 Wheeler's idea of the participatory universe. The universe as a self-excited circuit of self-observation. Adapted from Wheeler (1983)

the barrier between mind and physical body "does not exist".[76] Nevertheless, the point here is whether the reality of the latter can emerge in the mind alone.

- The other possibility is to consider this participatory universe also in ontological terms, that is, as a dynamical connection between events and correlations that has to do with the very constitution of the universe, as stressed in the previous subsections. Schrödinger was certainly inclined to this second (ontological or holistically realist) interpretation. In this way, our knowledge of the world, the problem of intelligibility that worried so much Einstein (Sect. 3.1.1), could be tentatively solved, and in this sense again both Einstein and Schrödinger would share a common insight: that the world must be *knowable* in itself (that is, it must display objective characters that make improvement in knowledge possible), otherwise no physical theory would make sense. On these ideas also Bohm, who not by chance quotes Leibniz on this issue (Sect. 2.3.3), agreed.[77]

Wheeler seemed to finally be rather sympathetic with the epistemic or subjective understanding of this idea or at least appears agnostic on this point.[78] However, in this way that insight would lose much of its power: it would be more interesting to admit that the universe can come to observe itself precisely because there is (i) a fundamental entrenchment of everything, and (ii) spontaneous or "natural" events occur: the dynamical connection between these two aspects could make the emergence of a classical world included the human mind reasonable. For this reason, we can say that the theoretical effort of the present work goes into a direction that

[76] Schrödinger (1958, Chap. 3).

[77] Bohm (1980, pp. 196–213).

[78] Somewhere he also appears to follow the epistemic interpretation (Wheeler 1988).

somehow is located after or before the sharp dividing between subject and object. In other terms, we shall consider cognition (including observation) as a part of the ontology of the physical world.[79]

5.2 Three Subsequent Steps in the Measurement Process

5.2.1 General Considerations

The examination of the previous section allows us to set measurement in a correct framework. Summarising also what said in the previous two chapters, I recall that the whole measurement process can be divided into three steps. A first step in which we prepare the system, then a second step in which the premeasurement or coupling (entanglement) of the system and the apparatus is established, and finally the third step in which a selection is made by the detector. Hence, preparation, premeasurement and detection constitute the three fundamental controlled local operations that a system can undergo in QM.[80] We shall learn that this three-step process is in fact a model of how information is in general acquired in our universe.

Let us summarise what we have learnt so far about such operations. Following Heisenberg, such operations should be thought of as concrete interventions on the physical world that can somehow affect the system at hand. Obviously, such interventions are not necessarily associated with events or produce events (and this is why premeasurement is still a reversible operation), nor they imply a direct interaction with the target system, and therefore we need to purify this view from interactionism, as sketched in Sect. 3.4.3. Moreover, this operational view does not imply any subjectivism either, since analogues of the three operations can happen spontaneously in nature. For this reason, we need to purify this view also from any instrumentalist interpretation, as shown in the quoted subsection. Indeed, the basic difference between the operations that we perform in our labs and their spontaneous analogues is that the former, being instantiated in a controlled way, allow us to make an inference about the object system (and therefore making ascription of properties possible: see Sect. 3.1.4). As we shall see, it is only in this very specific sense that we can say that a quantum phenomenon is an observed phenomenon. Nevertheless, the possibility of a property ascription is not excluded for spontaneous processes either; it may only be very difficult and often not all of the required conditions are available to us.

The late Heisenberg, having overcome is juvenile standpoint, formulated a very good synthesis of the three steps of the measurement process in the operational sense supported here. He says[81]:

---

[79] This is the main goal of my book (Auletta 2011).

[80] See the summary in Auletta and Wang (2014, Sect. 12.3).

[81] Heisenberg (1958, pp. 46–47).

> Therefore, the theoretical interpretation of an experiment requires three distinct steps: (1) the translation of the initial experimental situation into a probability function; (2) the following up of this function in the course of time; (3) the statement of a new measurement to be made of the system, the result of which can then be calculated from the probability function. For the first step the fulfilment of the uncertainty relations is a necessary condition. The second step cannot be described in terms of the classical concepts; there is no description of what happens to the system between the initial observation and the next measurement. It is only in the third step that we change over again from the "possible" to the "actual."

It is clear that the prepared state has to be described in terms of probabilities. I also stress that the second point clearly points out the typical unitary evolution characterising premeasurement. Finally, Heisenberg understands detection as an actualisation out of possible outcomes. Let us call *quantum operationalism* this view purified from both instrumentalism and interactionism (although, as mentioned, it is likely that Heisenberg himself never fully abandoned the second assumption). I finally attract the reader's attention on the words "there is no description of what happens to the system between the initial observation and the next measurement": as we shall see, this will allow us to understand the basic quantum dynamics as an alternate 'telegraphic' sequence of events and multipath dynamics.

5.2.2 Preparation

Determinative Measurement

Let us consider the first step of the measurement process, namely, preparation. Any (binary) state of the form

$$|\psi\rangle = c_0 |0\rangle + c_1 |1\rangle \,, \tag{5.38}$$

where the (complex) coefficients need to be specified, can always be *prepared*. For instance, the state $|0\rangle$ may represent a photon's horizontal polarisation state, while the state $|1\rangle$ may represent the photon's vertical polarisation state. Then, the state $|\psi\rangle$ may represent a certain linear polarisation orientation, determined by the coefficients c_0 and c_1. A preparation can be understood as the *determination* or the selection of the state of a system (indeed, we might recall that it is also called *determinative measurement*). It is the procedure through which only systems in a certain (previously theoretically defined) state are selected and delivered for further procedures, that is, allowed to undergo subsequent operations (premeasurement and detection). According to previous analysis, we can say that we determine in this way the global piece of information (qubit) that can be delivered in some of the forms previously described (Sect. 3.3.6).

Independence of Preparation

Note that we have not yet specified the observable that we wish to measure. In fact, in most cases the observable that will be measured is one such that the prepared state

is a superposition of its eigenstates. The reason is simple: if the system is already in one of the eigenstates of the target observable, our measurement will simply get again the same state (since none of its orthogonal states can occur: Sect. 1.2.4), so that no new information is gained. It can be proved in fact that *any* procedure of preparation is independent of the choice of the observable and therefore from the physical steps of both premeasurement and detection.[82] As we shall see, this grounds what the Pavia school calls the principle of causality (to be kept distinct from the principle of information causality introduced in Sect. 3.3.7). By now, let us stress that it is precisely this circumstance to make of premeasurement a step that is not only a pure mathematical reformulation of the initial state. The conceptual independence of preparation relative to detection can be considered one of the main distinctive characters of QM relative to CM that, at the opposite, merges state, observable and property in a single entity (Sect. 2.2.1) with the consequence that any preparation is also the choice of a detection.

Superdense Coding

An exception to what said so far seems represented by superdense coding, in which Alice a Bob share an entangled pair, e.g. in state (3.170c).[83] Then, Alice makes an operation on her qubit and sends it to Bob: if she wishes to send 00, she makes nothing on such a qubit; if she wishes to send 01, she applies the X gate, if she wishes to send 10, she applies the Z gate, and if she wishes to send 11, she applies the iY transformation to it. The resulting possible four states represent the Bell basis, and Bob needs only to measure the two entangled qubits to get the desired information. It seems that we can send two bits of classical information, thanks to an ebit and Alice's local operation on her qubit, in disagreement with the principle of information causality formulated in Sect. 3.3.7. However, apart from the transmission of Alice's qubit, the classical code needs to be (classically) shared in advance, what presupposes classical communication between Alice and Bob.

Preparation is Not Unitary

Preparation is not an unitary transformation. In fact, some systems are discarded and only a part passes the test. Moreover, systems in different states can pass the test, what shows that there is not a single unitary transformation that rules this operation. These considerations are also true for the final step of detection. There is also a subtler issue. Suppose that we wish to prepare the system in the state described by the first of Eqs. (3.89)

[82] D'Ariano et al. (2017, Sects. 2.4 and 5.1).

[83] Reference papers are Bennett and Wiesner (1992); Mattle et al. (1996). See also Nielsen and Chuang (2000, Sect. 2.3).

$$|0\rangle = \frac{1}{\sqrt{2}}\left(|+\rangle + |-\rangle\right). \tag{5.39}$$

Obviously, we send several systems of the same kind (for instance, photons) that could also be in different states. We do not know this since we do not have a control on the systems before the preparation step. However, what happens for a single system? Suppose that the system is in a state like (3.82a) just before preparation. Then, it can be prepared in the desired state through a Hadamard transformation (4.6), so that we have

$$\mathrm{H}|+\rangle = |0\rangle\ . \tag{5.40}$$

This means that all states that are discarded can be also represented through a unitary operation (see the third row of Eqs. (4.11)):

$$\mathrm{XH}|+\rangle = |1\rangle\ , \tag{5.41}$$

since Eq. (4.15b) tells us that the product of any two unitary operations (like X and H) is itself unitary, and where I recall that $|1\rangle$ is given by the second (3.89). Note that the operator X in Eq. (5.41) makes the unitary operator XH a NOT-Hadamard (or rather Hadamard-NOT). This allows us to consider $|0\rangle$ and $|1\rangle$ as negations of each other: $|0\rangle = \left|1'\right\rangle, |1\rangle = \left|0'\right\rangle$, where the 0s and 1s of a columnar state vector need to be exchanged (without consideration of the normalisation constant). In other words, negation is strictly connected with orthogonality.

However, although the latter statement remains true, the above procedure making use of unitary operators is not generalisable to dimensions ≥ 3. For dimensions ≥ 3, such operations must be represented by projectors, which are not unitary. The fact that, in the 2D case, we can represent a preparation by means of unitary transformations is because X not only does select a subspace of the same space but also brings one component vector (e.g. $|0\rangle$) to the other one (e.g. $|1\rangle$). However, this depends on the contingent fact that we have here only two dimensions corresponding to the binary truth-value assignment. In fact, also the two transformations (5.40) and (5.41) are equivalent to projections.

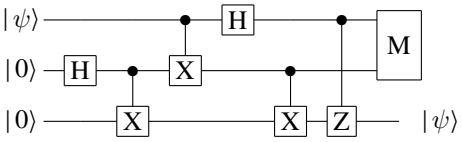
Outcome and Preparation

It is opportune to come back here to the issue of coarse-grained events and outcomes (Sects. 5.1.1 and 5.1.5). When we prepare a system, the physical outcome (the system whether passes the test or not and is absorbed) is represented by projections, as said. This is the quantum event. However, preparation is a coarse-grained event, because we are targeting a specific state. In fact, the pure physical action of a preparation filter is simply to block or not certain systems (quantum events). Such an operation in itself tells us nothing about preparation, and in fact it can occur in very different physical contexts and for very different purposes, e.g. this could also represent the final measurement outcome. It is only the insertion of this outcome in the whole

preparation procedure (a coarse-grained event) to make *a preparation* of this. This is evident if we consider that, as a part of the preparation, we could also envisage to repeat the test in order to be sure that the prepared systems are in fact all in the targeted state (and no fluctuation has left pass some systems in a state that have been not targeted). According to the number of Hilbert space dimensions of the system, preparation tells us what is the maximum number of (classical) bits that we can extract from it.

'Moving' the Classical Component

Note that, when we prepare a quantum system, we prepare it as an *information processor* that can deliver some answer to some question. In other words, with preparation we are targeting a *specific* result, and this is the main difference relatively to the evolution of a free system. In the case of measurement (but also in analogous spontaneous cases), we need some *classical component*. Often, it is said that we need a classical measure or communication for getting such an effect. This is true. The reason for this classical component is that quantum systems per se would only replicate their basic structure (through all the three steps of measurement the system, if not destroyed, can keep its fundamental quantum nature). Obviously, while in our experiments we make use of this classical component with purpose, in nature it is rather the *local result* of interaction processes. However, also in the light of the previous discussion, this classical component can be of different kinds and can also occur at different times along the whole measurement process. Such classical component can be Alice's classical communication to Bob in a teleportation protocol or the detection outcome as revealed by an apparatus in a measurement protocol or also a Bell state measurement in entanglement swapping. In fact, according to the examination of Sect. 4.4, in quantum computational circuits, the measurement can always be moved from an intermediate stage of the circuit to the end of the circuit (reciprocally, unterminated quantum wires may be assumed to undergo measurements)[84]:



(5.42)

After H_2, the first $CNOT_{23}$ gate, the $CNOT_{12}$ gate, and H_1 will bring the system to the corresponding output state of Eq. (4.87). Here, the teleportation example is shown: compare with circuit (4.92). Note that the final X and Z operations (and their combination) are now quantum controlled operations and no longer depend on classical communication (in other words, the final measurement replaces the classical communication). It is easy to see that the application of the second $CNOT_{23}$ gate followed by the controlled Z operation will allow us to get the state

[84] On this subject see Nielsen and Chuang (2000, Sect. 4.4).

$$|\Psi\rangle_{123} = \frac{1}{2}\left[|00\rangle + |01\rangle + |10\rangle + |11\rangle\right]_{12} |\psi\rangle_3 . \tag{5.43}$$

This can be seen as a direct quantum-mechanical 'instruction'.

States as Equivalence Classes

Since a state can be prepared using different procedures, we can operationally understand the state as an *equivalence class of preparations*. Indeed, different preparations can lead to the same state (they can be considered equivalent) and also different systems may be prepared in the same state. For instance, we can prepare photons in a state of (linear) vertical polarisation by

- Either inserting a vertical polarisation filter on the trajectory of photons that we know beforehand to be in an arbitrary polarisation state, provided that they are not all in a horizontal polarisation state (otherwise we got no output at all), or
- Mixing through a polarisation beam splitter two incoming beams, one prepared beforehand in polarisation at 45° and the other at 135° with opportune phase difference.

In other words, states as equivalence classes have the nature of types and, from a physical point of view, their ontological substrate is in the system prepared so and so, thus being a type–token (Sect. 3.4.2), displaying such character and specific sensitivity to physical contexts. We can see now why CM, treating physical systems as tokens, could erroneously assume that also the state of a physical system be a token (the already mentioned fallacy of the misplaced concreteness pointed out by Whitehead (Sect. 3.3.4)).

5.2.3 Premeasurement and Reversibility

Definition of Premeasurement

A *premeasurement* consists in an *interrogation* of a quantum system with respect to some *observable* or degree of freedom (such as position, momentum, energy, angular momentum and so on). As quantum features are an intrinsic characteristic of entanglement and premeasurement essentially consists in entangling the object system with the apparatus, they play an important role in premeasurement operations. Obviously, also classical correlations are important, otherwise we would not be able to pair apparatus and object system as explained in Sects. 5.1.2–5.1.3.

We might recall that QM seems to imply that the specific basis used for the expansion of the compound state of the apparatus and the system is irrelevant, and therefore that premeasurement is not about a specific observable (Sect. 3.1.2). As

mentioned, at a rather abstract level, several bases for the same compound state of the apparatus and the system are possible. This is a prerequisite of the equivalence of the Schrödinger and Heisenberg pictures in describing the evolution of the composite system (Sect. 2.1.1). However, we should not mix measurement procedures with algorithms. When we consider a specific physical situation (that is, once a particular setting is chosen), we introduce a further degree of determination or selection (relative to preparation) and are no longer authorised to treat different experimental contexts as equivalent, as recalled in Sect. 3.3.2. It would be highly unphysical to consider all observables as equivalent in a concrete experimental context, since changing the apparatus basis (i.e. the setting) means a concrete change of the apparatus as such, so that we may no longer assume to have the same or an equivalent measurement procedure. For instance, moving the detectors before or after the second beam splitter in the delayed-choice experiment described in Sect. 2.4.3 means to measure two different observables: the path observable and the interference visibility, as displayed in Sect. 5.1.4. In fact, preparation as such does not determine either possibilities. This is also evident by considering that the eigenstates of object system and apparatus displayed in Eqs. (3.25) are a superposition of those displayed in Eq. (3.18). Thus, choosing a certain premeasurement context univocally individuates a certain observable. Vice versa, if we wish to entangle two systems, we need to perform this operation through a particular observable. That a premeasurement context individuates a certain observable (or a certain degree of freedom) is also true, to a certain extent, from a classical point of view, since, as mentioned in Sect. 2.2.1, each apparatus is better suitable for measuring a certain observable and not others. In other words, premeasurement tells us how to acquire information from the system.

Moreover, since we deal here with code–alphabet coupling (Sect. 3.3.7), I recall that:

- The states of the apparatus can only be coupled with that of the object system, which means that the association is *arbitrary*, because
- We cannot manipulate codification non-locally since, as stressed, codification is a local operation. In other words, this association is arbitrary since is the *result of a local choice*.

This clarifies the relation between the postulated reversibility of the premeasurement step and the fact that we cannot arbitrarily change the chosen setting, two apparently contradictory statements. In fact, when we codify information we could always (counterfactually) have chosen another form of codification (the reversibility aspect) in a sequential process: nevertheless, our local choice (independently of the time in which it happens) was to codify information *in this and not in that way* (this implies an irreducibility of the choice) and nobody can manipulate this local act of selection from the exterior or non-locally. Let us consider these issues by means of concrete (and sophisticated) experiments. We shall also positively prove the reversibility of this step.

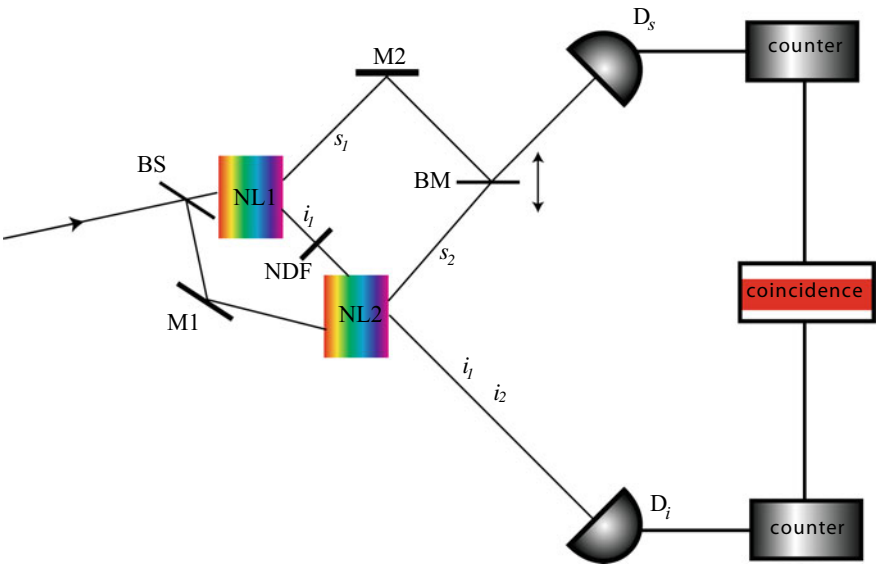


Fig. 5.13 Schematic representation of Mandel et al.'s experiment. With some slight differences, we have here the same experimental arrangement of Fig. 3.8, Sect. 3.3.5. An initial photon beam is split by a beam splitter (BS) and the resulting beams, after being deflected by the mirror M1, travel toward two nonlinear crystals of $LiIO_3$ (NL1 and NL2). From NL1 a signal photon (s_1) and an idler photon (i_1) emerge: The i-photon passes through NL2 and will be aligned with the second idler (i_2), which is emitted by NL2 together with the second signal photon (s_2). The two s-photons are combined by the mirror M2 and the beam merger (BM), and the outgoing beam falls on detector D_s, whereas the two idler photons fall on detector D_i. BM may be vertically displaced (in order to allow different interference fringes at Ds). An additional difference with the previous experiment is that a neutral density filter (NDF) is inserted between NL1 and NL2. In the case where the transmissivity of the NDF is 100%, when examining the coincidences we cannot distinguish between the two (idler and signal) photon pairs from either NL1 or NL2: this is the so-called fourth-order interference. When the transmissivity of the NDF is 0, then i_1 is blocked and a coincidence can only result from the signal and idler photon pair emitted by NL2. Here, there is no ambiguity and no interference. For values of transmissivity between 0 and 1, we have intermediate possibilities. Adapted from Auletta et al. (2009, p. 309)

Mandel et al.'s Experiment

Let us introduce the first one, performed by the experimental team of L. Mandel.[85] As displayed in Fig. 5.13, we have essentially two alternative setups.

(i) When the transmissivity of the NDF is 100%, one cannot distinguish between pairs $|s_1\rangle$, $|i_1\rangle$ and $|s_2\rangle$, $|i_2\rangle$: this reflects itself into the presence of interference fringes in the coincidence rate of detection at D_i and D_s, obtained for different vertical displacements of BM.
(ii) On the contrary, when the transmissivity is zero, no interference is visible.

[85] See the original study (Wang et al. 1991). For the remnant of this subsection see Auletta et al. (2009, Sect. 9.5).

In other words, we can at any time change the transmissivity of the NDF switching from setting (i) to setting (ii), and in this way erase the interference, or alternatively switching from setting (ii) to setting (i) and reconstructing it. We can even alternate these procedures. Now, in the context of setup (i), it is particularly interesting to note that the counting rate registered by D_s alone suffices to exhibit interference, because D_s cannot distinguish if the s-photon comes from NL1 or from NL2. It is also remarkable that interference is present not only when the transmissivity is 100% (the ideal case) but also whenever it is >0, in agreement with our analysis of the smooth complementarity (Sect. 5.1.4). Regarding setup (ii), it is not so obvious why blocking the photon i_1 washes out the interference: since the down-conversions in both NL1 and NL2 are spontaneous, detector D_s should not be able to distinguish whether the s-photon comes from NL1 or NL2 (Fig. 5.14). Why, then, do the results

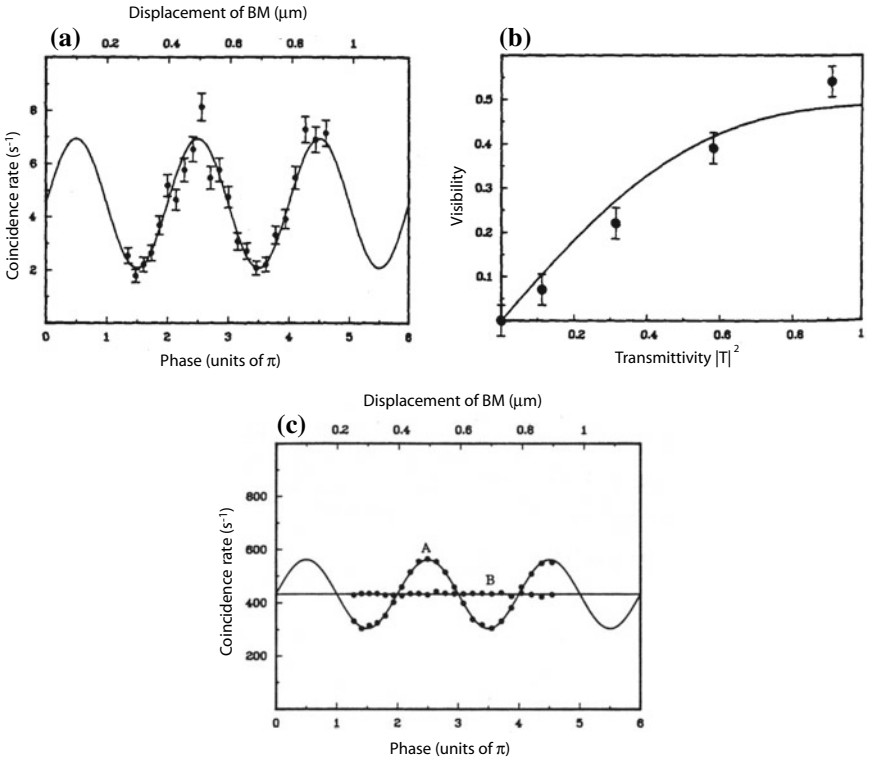


Fig. 5.14 **a** The results of measuring the coincidence rate of detectors D_s and D_i for various displacements of BM are here shown. The solid line represents the best-fitting sinusoidal function of the expected periodicity. Error bars show the statistical uncertainty. **b** Measured visibility $\mathcal{V}$ of the interference (Sect. 5.1.4) registered by the coincidence counting rate of detectors D_s and D_i for various filter transmissivities $|T|^2$. **c** Measured signal photon counting rate of the sole detector D_s as a function of BM displacement. Curve A: filter transmissivity $|T|^2 = 0.91$; curve B: $|T|^2 = 0$. Adapted from Wang et al. (1991)

show no interference for transmissivity = 0? It is clear that this phenomenon cannot be caused by some disturbance of the signal photons. Then, the conclusion is that the sole *in principle* distinguishability of the two 'paths' giving rise to the interference (signal photon coming from NL1 or from NL2) suffices to destroy the latter. As a matter of fact, if i_1 is blocked, we *could* use the information from detector D_i to establish whether the s-photon comes from NL1 (no coincidence since for s_1 the path is longer than for i_2) or from NL2 (we have coincidence since the paths for s_2 and i_2 are equal). We find here a generalisation of the result found with the interaction-free measurement (Sect. 2.1.2). I do not need to stress how these general results agree with the idea that experimental conditions have a real influence in themselves (without any dynamical action) on both the state of the systems and the possible further steps in measurement, for instance, producing localisation of a system. What is more remarkable is that the sole possibility to acquire information (a potentiality, in fact, representing the latter a reservoir of information: Sects. 3.2.3 and 3.3.6) alters the probabilities of events. Such a state of affairs needs to be explained out of the dynamics of quantum systems. In fact, these conclusions are not understandable if probabilities were not objective (Sect. 5.1.1) and information did not play a crucial role as again an objective although potential quantity (Sect. 3.4.3). As a final comment, I stress again that quantum systems are sensitive to the geometry of their environment, as anticipated in Sects. 3.3.4 and 5.1.3.

Quantum Eraser

The second experiment on reversibility of premeasurement displays a complementarity between wave-like and corpuscular behaviours. This experiment was first proposed by the physicists M. O. Scully, B.–G. Englert and H. Walther,[86] and, helpful also for further considerations, is known as the *quantum eraser*. In the original proposal, we have an atomic beam which goes through two slits of wall I (Fig. 5.15); behind this wall there is a further series of slits which are used as collimators to define two atomic beams that reach the narrow slits of wall II where the interference originates. Between wall I and wall II and after the collimators, the atomic beams are orthogonally intersected by an intense source of light, for instance, a laser beam, which brings the internal state $|\iota\rangle$ of the two-level atoms from an unexcited (ground) state $|g\rangle$ into an excited state $|e\rangle$. Thereafter, each of the two beams passes through a microcavity (a device for trapping radiation). Finally, they fall on a screen. The atomic source is adjusted in such a way that there is at most one atom at time in the apparatus.

Let us consider the formal details which are very instructive. In the interference region, the 3D wave function describing the center-of-mass motion of the atoms (where $\mathbf{r}$ indicates the center-of-mass coordinate) is the superposition of the two terms referring to slit 1 and slit 2, so that the total (center-of-mass plus internal) wave function is given by (see also Eq. (3.41))

[86] Scully et al. (1991).

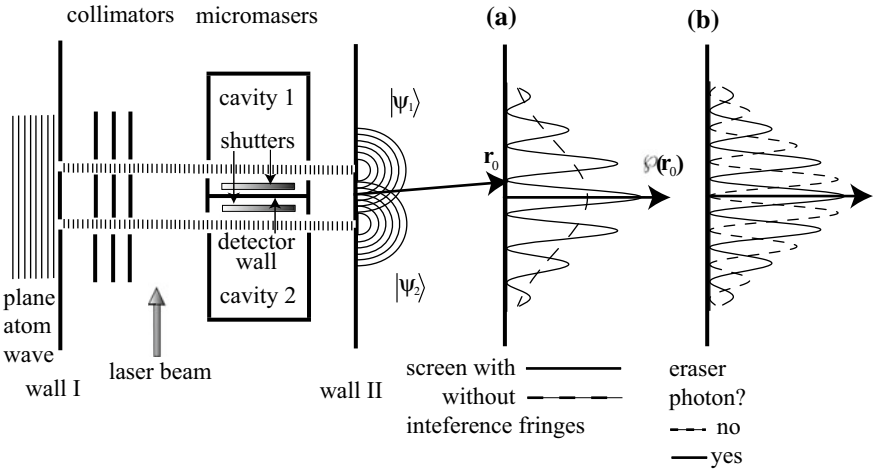


Fig. 5.15 Scully, Englert and Walther's proposed experiment (see also Fig. 1.3, Sect. 1.1.3). A set of slits (after the wall I) collimates two atomic beams that pass through the narrow slits (wall II) where the interference pattern originates. This setup is supplemented by two microcavities and a laser beam to provide which-path information. **a** Quantum-erasure configuration in which electro-optic shutters separate microwave photons in the two cavities from the thin-film semiconductor (the central detector wall) which absorbs microwave photons and acts as a photodetector. In the absence of the laser beam, there is no possibility to obtain which-path information and we have interference on the screen (solid line). Instead, introducing the laser beam, we may acquire which-path information and the interference is destroyed (dashed line). **b** The shutters are open. The probability density of the particles on the screen depends upon whether a photocount is observed in the detector wall ('yes') or not ('no'), demonstrating that the correlation between the event on the screen and the eraser photocount is necessary to retrieve the interference pattern. Adapted from Auletta et al. (2009, p. 312)

$$\Psi(\mathbf{r}) = \frac{1}{\sqrt{2}} \left[\psi_1(\mathbf{r}) + \psi_2(\mathbf{r})\right] |\iota\rangle \ , \tag{5.44}$$

and the probability density of particles falling on the screen at the point $\mathbf{r} = \mathbf{r}_0$ will be given by

$$\begin{aligned} \wp(\mathbf{r}_0) &= \Psi(\mathbf{r}_0)\Psi^*(\mathbf{r}_0) \\ &= \frac{1}{2}\left[|\psi_1(\mathbf{r}_0)|^2 + |\psi_2(\mathbf{r}_0)|^2 + \psi_1^*(\mathbf{r}_0)\psi_2(\mathbf{r}_0) + \psi_2^*(\mathbf{r}_0)\psi_1(\mathbf{r}_0)\right] \langle\iota|\iota\rangle \ . \end{aligned} \tag{5.45}$$

Before coming to the interference region, the atoms pass through the microcavities. The cavity frequency is tuned in resonance with the energy difference between the excited and ground states of the atoms. The velocity of the atoms may be selected in such a way that, after being beforehand prepared in an excited state by the laser beam, on passing through either one of the cavities each atom will emit a microwave photon (which remains trapped in the cavity) and leave which-path information. After the atom has passed through the cavity it is again a free particle and its momentum keeps the initial value.

When passing through the cavities and making the transition from $|e\rangle$ to $|g\rangle$ by releasing a photon, the state of the global system (atomic beam plus cavity) is given by

$$\Psi(\mathbf{r}) = \frac{1}{\sqrt{2}}\left[\psi_1(\mathbf{r})|1_1 0_2\rangle + \psi_2(\mathbf{r})|0_1 1_2\rangle\right]|g\rangle \ , \tag{5.46}$$

where $|1_1 0_2\rangle$ denotes the state in which there is one photon in cavity 1 and none in cavity 2. In contrast to Eq. (5.45), the probability density on the screen is now

$$\begin{aligned}\wp(\mathbf{r}_0) = \frac{1}{2}\big[&|\psi_1(\mathbf{r}_0)|^2 + |\psi_2(\mathbf{r}_0)|^2 + \psi_1^*(\mathbf{r}_0)\psi_2(\mathbf{r}_0)\langle 1_1 0_2|0_1 1_2\rangle \\ &+\psi_2^*(\mathbf{r}_0)\psi_1(\mathbf{r}_0)\langle 0_1 1_2|1_1 0_2\rangle\big]\langle g|g\rangle \ .\end{aligned} \tag{5.47}$$

Since the two cavity state vectors $|1_1 0_2\rangle$, $|0_1 1_2\rangle$ are orthogonal to each other, the interference terms vanish in Eq. (5.47) and diffraction fringes are washed out, so that Eq. (5.47) reduces to

$$\wp(\mathbf{r}_0) = \frac{1}{2}\left[|\psi_1(\mathbf{r}_0)|^2 + |\psi_2(\mathbf{r}_0)|^2\right]\langle g|g\rangle \ , \tag{5.48}$$

which represent the classical information about the path followed. Let us now separate the detectors in the cavity by a shutter–detector combination, so that, when the shutters are closed, the photons are forced to remain in either the upper or the lower cavity. However, if the shutters are opened, light will be allowed to interact with the photodetector wall and in this way the radiation will be absorbed and the memory of the passage erased (such an operation is called *quantum erasure*). After the erasure, will we again obtain the interference fringes which were eliminated before? The answer is yes, so that interference effects can be restored by manipulating the *which-path* detectors long after the atoms have passed and before reaching the final (detection) wall, a kind of delayed choice (Sect. 2.4.3).

This result can be formally expressed as follows. Let us include the photodetector walls into the description. These are initially in the ground state $|g\rangle_{\mathrm{D}}$, so that Eq. (5.46) modifies to

$$\Psi(\mathbf{r}) = \frac{1}{\sqrt{2}}\left[\psi_1(\mathbf{r})|1_1 0_2\rangle + \psi_2(\mathbf{r})|0_1 1_2\rangle\right]|g\rangle_{\mathrm{A}}|g\rangle_{\mathrm{D}} \ , \tag{5.49}$$

where the subscript A denotes now the atom explicitly for avoiding confusion. After absorbing the photon, the photodetector passes to an excited state $|e\rangle_{\mathrm{D}}$. If we introduce symmetric and antisymmetric atomic states (see also Sect. 1.3.1)

$$\psi_{\pm}(\mathbf{r}) = \frac{1}{\sqrt{2}}\left[\psi_1(\mathbf{r}) \pm \psi_2(\mathbf{r})\right] \ , \tag{5.50}$$

together with symmetric and antisymmetric states of the radiation fields contained in the cavities

$$|\Psi^{\pm}\rangle = \frac{1}{\sqrt{2}}\left[|1_1 0_2\rangle \pm |0_1 1_2\rangle\right] , \tag{5.51}$$

we can rewrite Eq. (5.49) as

$$\Psi(\mathbf{r}) = \frac{1}{\sqrt{2}}\left[\psi_+(\mathbf{r})|\Psi^+\rangle + \psi_-(\mathbf{r})|\Psi^-\rangle\right]|g\rangle_{\rm A}|g\rangle_{\rm D} . \tag{5.52}$$

The action of the quantum eraser on the system is to change Eq. (5.52) into

$$\Psi'(\mathbf{r}) = \frac{1}{\sqrt{2}}\left[\psi_+(\mathbf{r})|0_1 0_2\rangle|e\rangle_{\rm D} + \psi_-(\mathbf{r})|\Psi^-\rangle|g\rangle_{\rm D}\right]|g\rangle_{\rm A} . \tag{5.53}$$

The reason for the difference between the two components is that the interaction Hamiltonian between radiation and photodetectors only depends on symmetric combinations of radiation variables so that the antisymmetric state remains unchanged.

Now, as long as the final state of the photodetector is unknown, the atomic probability density at the screen is

$$\begin{aligned}\wp(\mathbf{r}_0) &= {\rm Tr}_{A,F,D}\left[\Psi'^*(\mathbf{r}_0)\Psi'(\mathbf{r}_0)\right] \\ &= \frac{1}{2}\left[\psi_+^*(\mathbf{r}_0)\psi_+(\mathbf{r}_0) + \psi_-^*(\mathbf{r}_0)\psi_-(\mathbf{r}_0)\right] \\ &= \frac{1}{2}\left[\psi_1^*(\mathbf{r}_0)\psi_1(\mathbf{r}_0) + \psi_2^*(\mathbf{r}_0)\psi_2(\mathbf{r}_0)\right] ,\end{aligned} \tag{5.54}$$

where the trace has been performed on the atomic, field and detector degrees of freedom. Clearly, Eq. (5.54) does not show any interference terms, in agreement with Eq. (5.48). However, if we compute the probability density for finding both the photodetector excited (that is, we perform a projection on that state) and the atom at $\mathbf{r}_0$ on the screen, we have

$$\begin{aligned}\wp_{e_{\rm D}}(\mathbf{r}_0) &= {\rm Tr}_{A,F}\left[|\,e\rangle_{\rm D}\,\langle e\,|\,\Psi'^*(\mathbf{r}_0)\Psi'(\mathbf{r}_0)\right] \\ &= |\psi_+(\mathbf{r}_0)|^2 \\ &= \frac{1}{2}\left[|\psi_1(\mathbf{r}_0)|^2 + |\psi_2(\mathbf{r}_0)|^2\right] + \Re\left[\psi_1^*(\mathbf{r}_0)\psi_2(\mathbf{r}_0)\right],\end{aligned} \tag{5.55}$$

which exhibits the same interference term as Eq. (5.45). I recall that $\Re$ denotes the real part of a complex number (first of Eqs. (1.9)). In similar way, the probability of finding both the photodetector in ground state and the atom at $\mathbf{r}_0$ on the screen is

$$\begin{aligned}\wp_{g_{\rm D}}(\mathbf{r}_0) &= {\rm Tr}_{A,F}\left[|\,g\rangle_{\rm D}\,\langle g\,|\,\Psi'^*(\mathbf{r}_0)\Psi'(\mathbf{r}_0)\right] \\ &= |\psi_-(\mathbf{r}_0)|^2 \\ &= \frac{1}{2}\left[|\psi_1(\mathbf{r}_0)|^2 + |\psi_2(\mathbf{r}_0)|^2\right] - \Re\left[\psi_1^*(\mathbf{r}_0)\psi_2(\mathbf{r}_0)\right],\end{aligned} \tag{5.56}$$

giving rise to the dashed antifringes indicated in Fig. 5.15b. In practice, it is again expression of an interference but totally out of phase relative to that expressed by the previous equation. Fringes and antifringes can be related to two sets of orthogonal states on the plane shown in Fig. 4.5, Sect. 4.2.3 like $\{|+\rangle\ ,|-\rangle\}$ and $\{|y_+\rangle\ ,|y_-\rangle\}$.

It is important not to forget that the photon is a quantum system: when the shutters are opened, the two cavities become a single larger one. Now, the photon's wave is a combination of the two partial waves, such that either the two waves reinforce each other (constructive interference) and the photosensor detects the photon, or they mutually extinguish each other, with the consequence that the photosensor detects no photon (destructive interference). In other words, there is only a 50% probability of detecting the photon.

Discussion

The summary of the above conceptual experiment naturally leads us to formulate the following question:

- Do we want to know whether we registered a 'slit 1'-atom or a 'slit 2'-atom,
- Or are we interested in ascertaining one of the two situations, i.e. having the microwave-photon sensor either excited ($|e\rangle_{\mathrm{D}}$) or not ($|g\rangle_{\mathrm{D}}$)?

We cannot answer both questions at the same time. In other words, we *either* know the *atom path* without using the eraser and hence without knowing anything about the photodetector and its state (first alternative), *or* we desire to know the latter (second alternative) and we must reproduce the *interference* by losing all information about the particles' path. In the former case, the absence of interference (Eqs. (5.48) and (5.54)) is due to the sum of probabilities (5.55) and (5.56).[87]

The assumption that *which-path* detectors are recoil-free has been questioned by the physicist P. Storey and co-workers[88]: they pointed out that the which-path determination would not be possible without a double momentum transfer between the cavity system constituted by detector–photon (when a photon is emitted and then reabsorbed from the opposite direction) and the atom, whose magnitude is in the limits of the uncertainty relation between position and momentum. However, Scully and co-workers showed the correctness of their results that have also been experimentally confirmed.[89] Therefore, there is no significant change in the spatial wave function of the atoms. It is only the correlation between the center-of-mass wave function and the photon's degrees of freedom in the cavities that are responsible for the loss of interference.

[87] In a further experiment, the erasure is provided by one of two entangled photons (Kim et al. 2000).

[88] Storey et al. (1994).

[89] Englert et al. (1995).

Lessons

The experiment confirms some lessons that we have already drawn from the previous experiment:

(1) The washing out of the interference is reversible as long as we do not actually acquire and/or store information.
(2) When measuring, one cannot consider the object system as separated from the premeasurement context of the experiment being performed.
(3) It is not necessary a detection (of the photon) for erasing which-path information: indeed, we have interference also for the case described by Eq. (5.56). This lesson can be generalised as follows: although in some situations may play a role, local interactions are not necessary for correlating quantum systems, neither for coupling detector and object system. This means that the interaction Hamiltonians introduced in Sect. 3.1 are in fact not interaction Hamiltonian in the usual sense of the word. The alternative is to enlarge our understanding of interaction and including in such a category both local interactions (with exchange of energy nor momentum) and non-local quantum correlations, in agreement with our analysis of the causal role of correlations (Sects. 3.4.2 and 5.1.2–5.1.4). There are also further lessons to be drawn:
(4) The second experiment supports the statement that the complementarity principle is genuinely a fundamental principle of QM, although, as seen, the way in which we interpret this is different from Bohr's original formulation (Sect. 5.1.4).
(5) The experiment shows that we have a swap of information (see Eqs. (4.28a)–(4.28b)): the presence of the photon in a cavity and its absence (it is being detected) can be understood as a qubit that goes from the atom to one of the cavities and then to the detector, when the information is erased: from $|1\rangle_A \otimes |0\rangle_j$ to $|0\rangle_A \otimes |1\rangle_j$ and then to $|0\rangle_j |1\rangle_D$ (where $j = 1, 2$). Such a swap can be also reversible, for instance, consider Fig. 2.4a, Sect. 2.1.3. The photon is with high probability in the upper path after N steps (with N large). However, if after this procedure we let the photon pass another series of N beams splitters but with probability $\cos^2(\pi/2N)$ to be transmitted, the photon is now in the lower path with high probability. We can also reiterate the game. Thus, we swap a qubit from $|1\rangle_u |0\rangle_l$ to $|0\rangle_u |1\rangle_l$ and vice versa.

Observables as Equivalence Classes

Premeasurement is selecting an observable, and, as we shall see, allows the correct sampling of the event space. An observable is a physical magnitude, namely, a collection of possible eigenvalues (each of them can be represented by, or associated with, a projector). It represents a 'physical dimension' of the system for describing the latter. Moreover, since when measuring a certain observable a number of slightly different concrete physical contexts (settings) could be equally good, from an operational point of view an observable is an *equivalence class of premeasurements*.

For instance, we can perform the same kind of complementarity or delayed-choice experiment by

- Either choosing an interferometer or
- Performing a so-called double-slit experiment.

Since observables are classes and therefore 'universals' (types), the laws of Nature, being relations among observables (Sect. 2.4.2) express relations among universals.[90] This sheds light on our previous examination. The fact that an observable is an equivalence class of operations clearly confirms that it is an element of reality (as wished by EPR: see Sect. 3.3.2), although with two provisos: the notion of observable makes sense only (i) in the context of potential conditions (i.e. not yet activated) out of which we can get one of its eigenvalues, and (ii) in the context of a theory. However, since it is interpreted and is an equivalence class, it cannot be taken as a primary ontological substrate. In fact, this physical–ontological substrate of observables is represented by the coupling interaction that involves entanglement (the physical correlations including quantum features) that we build between the object system and the apparatus. In fact, when several systems become interconnected, quantum features spread in such a way to involve also classical correlations among the values of the observables pertaining to different systems (Sect. 5.1.3): thus, I stress again that pure quantum correlations (features) are types while entanglement, when involving also physical contexts, is a type–token. But this is precisely what singles out certain observables in the coupled systems. This seems to be like the chicken and the egg problem: what comes first? Locally speaking, the notion of observable, being an equivalence class of premeasurements that are performed by entangling systems, is conceptually dependent on correlations (including quantum features) among those systems.

5.2.4 Detection

Selection Device

Detection (or measurement in the strict sense of the word) is an *answer* to our interrogation. When we establish an entanglement with an apparatus, we are actually also entangling the object system with some detector. Although a detector is in general considered to be a part of an apparatus, these two devices have two conceptually very different functions. A measurement apparatus, in the strict sense of the word, is a *coupling* device for selecting an observable, while a detector is an *event selection* device; thus, we can still use the term "apparatus" to cover both functions but we should avoid any confusion on this point. This justifies the fact that, properly speaking, the apparatus is an interface between the detector and the object system. When a suitable selection is made, the detection apparatus is in one of its basis states

[90] Armstrong (1983, Chap. 6).

and, through the coupling established by the apparatus with the object system, it tells information about the latter, and therefore allows us to ascribe a *property* (the value of an observable) to the object system. In other words, this connection allows for a certain random outcome telling us something about the input state.[91] Note that, as for preparation, also observation or the attribution of a property is the result of a coarse-graining, while what is primarily and not coarse-grained is the detection outcome as such.

In this subsection we shall deal with detection as a local event embedded in a generalised dynamics that takes both reversible and irreversible aspects into account.

Amplitude Operators

As mentioned in Sect. 2.1.3, many scholars use the term *effect* instead of "property". To the former notion is connected a powerful formalism known as positive operator value measure (POVM).[92] Suppose that the initial state of the apparatus is some ready or initial state $|\,a_0\rangle$ while the state of the object system is some superposition state $|\,\psi\rangle$. Then, the premeasurement step takes the form

$$|\,\psi\rangle\,|\,a_0\rangle \overset{\hat{U}_t}{\longmapsto} \sum_j c_j\,|\,j\rangle\,\left|\,a_j\right\rangle, \tag{5.57}$$

where $\{|\,j\rangle\,\}$ is an orthogonal basis for the system representing the eigenstates of some system observable to be measured, $\{\left|\,a_j\right\rangle\}$ is an orthogonal basis for the apparatus, and $\hat{U}_t$ is the time evolution operator whose form depends on the coupling of the system and the apparatus. As usually, we assume that the entanglement between the object system and the apparatus created during the premeasurement step is the result of a unitary transformation. Indeed, I remind the reader that only the first (preparation) and the final (detection) steps of selection are not unitary, in agreement with the results of the two previous subsections. In the density matrix formalism, the initial state of the object system and the apparatus appearing on the LHS of the transformation (5.57) may be described by the (factorised) density matrix $\hat{\rho}_S \otimes \hat{\rho}_A$, where

$$\hat{\rho}_S = |\,\psi\rangle\,\langle\psi\,| \quad \text{and} \quad \hat{\rho}_A = |\,a_0\rangle\,\langle a_0\,|\,. \tag{5.58}$$

Therefore, in agreement with Eq. (2.18), we have the following unitary transformation:

$$\hat{\rho}_S \otimes \hat{\rho}_A \longmapsto \hat{U}_t(\hat{\rho}_S \otimes \hat{\rho}_A)\hat{U}_t^{\dagger}. \tag{5.59}$$

[91] On this subject see Auletta (2005).

[92] Fundamental and now classical textbooks on the subject are Davies (1976); Kraus (1983). See also Busch et al. (1995); de Muynck (2002). Fort a summary, see Auletta et al. (2009, Sect. 9.10); Auletta (2014b).

Any time that a system and the relative apparatus interact (in the broad sense of the word, explored in the previous subsection), we may write the transformation $\hat{\mathcal{T}}$ on the system as a result of the interaction, as the partial trace

$$\hat{\mathcal{T}}\hat{\rho}^{\mathcal{S}} = \mathrm{Tr}_{\mathcal{A}}\left[\hat{U}_t\left(\hat{\rho}^S \otimes \hat{\rho}^A\right)\hat{U}_t^\dagger\right] . \tag{5.60}$$

The transformation $\hat{\mathcal{T}}$ is properly speaking a superoperator (an operator acting on operators) and should be discriminated by a double hat. However, nothing is essentially improved by introducing this complication of the notation. Equation (5.60) may be rewritten as

$$\begin{aligned}\hat{\mathcal{T}}\ \hat{\rho}^S &= \sum_j \left\langle a_j \left| \hat{U}_t\left(\hat{\rho}^S \otimes |\, a_0\rangle\,\langle a_0\,|\right)\hat{U}_t^\dagger \right| a_j \right\rangle \\ &= \sum_j \hat{\vartheta}_j\left(\hat{\rho}^S \otimes \hat{I}_A\right)\hat{\vartheta}_j^\dagger\ ,\end{aligned} \tag{5.61}$$

with $j \in \mathbb{N}$ ($\mathbb{N}$ being the set of natural numbers), where $\hat{I}_A$ is the identity operator (in the apparatus Hilbert space $\mathcal{H}_A$), and

$$\hat{\vartheta}_k = \left\langle a_k \left| \hat{U}_t \right| a_0 \right\rangle \text{ and } \hat{\vartheta}_k^\dagger = \left\langle a_0 \left| \hat{U}_t^\dagger \right| a_k \right\rangle \tag{5.62}$$

are called the *amplitude operators*. Note that these amplitudes are operators since the unitary operator $\hat{U}_t$ acts on *both* system and apparatus (while the involved states are of the latter only). Thus, the operation $\hat{\mathcal{T}}$ acting on the component $\hat{\rho}^S$ describing the initial state of the system, turning out to be a tracing out of the apparatus on the evolved state of system apparatus, is equivalent to the action of a set of amplitude operators on the initial state of the compound system.

I also note that, in general, any transition from some initial state $|\, i\rangle$ to a final state $|\, f\rangle$ needs to obey *Fermi's golden rule* (after the name of E. Fermi), according to which the probability for such transformation is

$$\wp(i \mapsto f) = \left|\left\langle f \left| \hat{O} \right| i \right\rangle\right|^2 \delta(E_f - E_i), \tag{5.63}$$

where $\hat{O}$ is some operator (e.g. the interaction Hamiltonian between two particles) and the delta function is for energy conservation (E_i, E_f being initial and final values of energy, respectively), when it is relevant to the problem.

Effect Operators

On this basis, we can introduce the *effect operator*

$$\hat{E}_j := \hat{\vartheta}_j^\dagger \hat{\vartheta}_j, \tag{5.64}$$

which is a projection-like operator that plays an important role in the theory of generalised measurement. Definition (5.64) shows that effects are Hermitian operators since they satisfy Property (1.82f). I remark that any density matrix $\hat{\rho}$ can be written as $\hat{\rho} = \hat{E}\hat{U}$ for some effect operator $\hat{E}$ and unitary operator $\hat{U}$. This is called *polar decomposition* and corresponds to the homologous for complex numbers $z = |z|e^{i\phi}$. From the definition of effect operator introduced here, it follows that,

$$\begin{aligned}\sum_j \hat{E}_j &= \sum_j \hat{\vartheta}_j^\dagger \hat{\vartheta}_j = \sum_j \left\langle a_0 \left| \hat{U}_t^\dagger \right| a_j \right\rangle \left\langle a_j \left| \hat{U}_t \right| a_0 \right\rangle \\ &= \sum_j |\, a_j \rangle\langle a_j \,| \left\langle a_0 \left| \hat{U}_t^\dagger \hat{U}_t \right| a_0 \right\rangle \\ &= \hat{I}_A, \end{aligned} \tag{5.65}$$

where use has been made of the completeness relation for the apparatus states

$$\sum_j \hat{P}_{a_j} = \sum_j |\, a_j \rangle\langle a_j \,| = \hat{I}_A. \tag{5.66}$$

Note that in such a case, the transformation $\hat{\mathcal{T}}$ defined by the effect operators (5.65) is quantum-deterministic (because it sums to identity). There are, however, cases of probabilistic transformations so that in general we have

$$\sum_j \hat{E}_j \leq \hat{I}. \tag{5.67}$$

In other words, quantum-deterministic transformations (and projectors) are only a special (and limiting) case of transformations (effects). This implies that effects are bounded from above by the identity operator, i.e. $\forall j$, $\hat{E}_j \leq \hat{I}$. Moreover, it can be shown that any $\hat{E}_j$ is positive semidefinite, i.e.[93]

$$\left\langle \phi_A \left| \hat{E}_j \right| \phi_A \right\rangle \geq 0 \quad \forall\, |\, \phi_A \rangle \in \mathcal{H}_A. \tag{5.68}$$

This explains the term POVM for denoting effect operators. Note that effects like density matrices (and projectors) are convex (Sect. 1.4.1). However, unlike the projectors, the effect operators in general does not satisfy the requirement (1.90) of orthogonality, namely,

$$\hat{E}_j \hat{E}_k \neq \delta_{jk} \hat{E}_k. \tag{5.69}$$

Compare Eqs. (5.61) and (3.12). It is such a property that allows a generalised theory of measurement, since it permits intermediate situations between two canonical observables, the so-called *unsharp measurement*.

[93] Kraus (1983, p. 10).

It is important to understand that effects separate for states, where *separating* means here that the same effect occurs with different joint probabilities over two different states: $\forall \hat{\rho} \neq \hat{\rho}'$, there is an effect $\hat{E}$ such that

$$\mathrm{Tr}\left(\hat{E}\hat{\rho}\right) \neq \mathrm{Tr}\left(\hat{E}\hat{\rho}'\right). \tag{5.70a}$$

Note that, according to Eq. (5.64),

$$\mathrm{Tr}\left(\hat{E}\hat{\rho}\right) = \mathrm{Tr}\left(\hat{\vartheta}^{\dagger}\hat{\vartheta}\hat{\rho}\right) = \mathrm{Tr}\left(\hat{\vartheta}\hat{\rho}\hat{\vartheta}^{\dagger}\right), \tag{5.70b}$$

in agreement with Eq. (5.61), where I have used the cyclic property (1.371c) of trace. Symmetrically, states separates for effects when the state as a functional over effects have two different values over two different effects[94]: $\forall \hat{E} \neq \hat{E}'$, there is a density matrix $\hat{\rho}$ such that

$$\mathrm{Tr}\left(\hat{E}\hat{\rho}\right) \neq \mathrm{Tr}\left(\hat{E}'\hat{\rho}\right). \tag{5.70c}$$

This is obviously true for projectors as a special case. Note that since non-orthogonal effects are able to discriminate a set of discriminable quantum states, discriminability and orthogonality are different notions, and since the latter is important for repeatable measurements in finite dimensions, discriminability is also independent of repeatability.[95]

Amplitude Operators Describe the Whole Measurement Process

It is not difficult to build amplitude operators for the system in analogy with Eqs. (5.62):

$$\hat{\vartheta}_m = \left\langle o_m \middle| \hat{U}_t \middle| \psi \right\rangle, \tag{5.71}$$

and similarly for its conjugate, where $\hat{U}_t$ is again an unitary operator for both system and apparatus, $|o_m\rangle$ is the resulting eigenstate of the measured observable $\hat{O}$ of the system, and $|\psi\rangle$ the initial state of the latter.

From its definition, it is clear that the amplitude operator $\hat{\vartheta}_m$ describes the three steps of the measurement of a given observable:

(i) Preparation of the initial state of the system (i.e. the input $|\psi\rangle$ that in Eqs. (5.62) is vicariously represented by the initial state of the apparatus),
(ii) Unitary time evolution (i.e. coupling or premeasurement) that entangles the system with the apparatus and allows us to select an observable (i.e. the processing represented by $\hat{U}_t$), and

[94] D'Ariano et al. (2017, Sect. 2.4).
[95] D'Ariano et al. (2017, Sects. 2.8.5 and 3.6–3.7).

(iii) Detection outcome $|o_m\rangle$, where the apparatus output $|a_m\rangle$ is again vicarious in Eqs. (5.62) and in fact allows us to assign a property to the system.

Note that the amplitude operator is the operatorial counterpart of the parameter θ on the Poincaré sphere, while the unitary time operator $\hat{U}_t$ includes both shifts in amplitudes and shifts in relative phase.

This formalism also allows an understanding of quantum-mechanical formalism that fits very well with our treatment of quantum systems in terms of information.[96] We may take any ket $|\psi\rangle$ as the input state, while any bra $\langle\varphi|$ as the output state. In this way, any projector $|\psi\rangle\langle\psi|$ represents a selection act while any scalar product $\langle\varphi|\psi\rangle$ means a possible transition, i.e. a transition amplitude. In the case in which it is equal to zero, it mean a forbidden transition (due to the orthogonality of the two states). Any operator that is in between an input and an output (like the above amplitude operators or, more generally, $\left\langle\varphi\left|\hat{O}\right|\psi\right\rangle$), as in Eq. (5.63), is the operation that bridges between input and output. According to the examination of Sect. 3.2.4, it can be understood as a channel that connects output with input. When input and output coincide, i.e. $\left\langle\psi\left|\hat{O}\right|\psi\right\rangle$, we have the probability amplitude that the transformation $\hat{O}$ does not change the state $|\psi\rangle$ that corresponds to the expectation value of $\hat{O}$ in that state.

The Environment

For the sake of later examination, let us widen our perspective and consider the environment instead of the apparatus (in fact we should consider both but to take into account only two systems introduces a helpful simplification).[97] The transformation $\hat{\mathcal{T}}$ is essentially of the form (5.61) with environment's states at the place of apparatus's states:

$$\begin{aligned}\hat{\mathcal{T}}\,\hat{\rho}^S &= \sum_j \left\langle e_j\left|\hat{U}_t\left(\hat{\rho}^S\otimes|e_0\rangle\langle e_0|\right)\hat{U}_t^\dagger\right|e_j\right\rangle \\ &= \sum_j \hat{\vartheta}_j\left(\hat{\rho}^S\otimes\hat{I}_E\right)\hat{\vartheta}_j^\dagger\,,\end{aligned} \tag{5.72}$$

where $\{|e_k\rangle\}$ is an orthonormal basis for the environment spanning its Hilbert space $\mathcal{H}_\mathcal{E}$ and the amplitude operators are given by

$$\hat{\vartheta}_k := \left\langle e_k\left|\hat{U}_t\right|e_0\right\rangle. \tag{5.73}$$

[96] On this subject see Finkelstein (1996).

[97] On this helpful formalism see Nielsen and Chuang (2000, pp. 356–73, 386–89); Auletta et al. (2009, Sect. 14.3).

Assume that we have a measurement outcome k. The probability that the system be in the final state

$$\begin{aligned}\hat{\varrho}_k^S &= \hat{P}_k \left(\hat{\mathcal{T}}\ \hat{\rho}^S\right) \hat{P}_k \\ &\propto \mathrm{Tr}_{\mathcal{E}} \left[\hat{P}_k \hat{U}_t \left(\hat{\rho}^S \otimes |\, e_0\rangle\, \langle e_0\,|\right) \hat{U}_t^\dagger \hat{P}_k\right] \\ &= \hat{\vartheta}_k \hat{\rho}^S \hat{\vartheta}_k^\dagger\,, \end{aligned} \tag{5.74}$$

where $\propto$ means "proportional to" and $\hat{P}_k = |\, e_k\rangle\, \langle e_k\,|$, can be written as

$$\wp_k = \mathrm{Tr}\left(\hat{\vartheta}_k \hat{\rho}^S \hat{\vartheta}_k^\dagger\right)\ . \tag{5.75}$$

We can rewrite such a probability as

$$\wp_k = \mathrm{Tr}\left(\hat{E}_k \hat{\rho}^S\right), \tag{5.76}$$

where I have used the cyclic property of the trace. Recalling the von Neumann's formula (3.14), let us normalise the state $\hat{\varrho}_k^S$ as

$$\hat{\varrho}_k^S = \frac{\hat{\vartheta}_k \hat{\rho}^S \hat{\vartheta}_k^\dagger}{\mathrm{Tr}_{\mathcal{S}}\left(\hat{E}_k \hat{\rho}^S\right)}\,, \tag{5.77}$$

and write the transformation (5.72) as (see Eq. (1.365))

$$\hat{\mathcal{T}}\,\hat{\rho}^S = \sum_j \wp_j \hat{\varrho}_j^S\ . \tag{5.78}$$

Let us now consider the special case in which the measurement projects the environment from an initial superposition of states $|\, e_k\rangle$'s onto the state $|\, e_m\rangle$. In this case, the final state of the compound system is given by (see again Eq. (3.14))

$$\hat{\rho}_f^{SE} = \frac{\hat{P}_m \hat{U}_t \left(\hat{\rho}^S \otimes \hat{\rho}^E\right) \hat{U}_t^\dagger \hat{P}_m}{\mathrm{Tr}\left[\hat{P}_m \hat{U}_t \left(\hat{\rho}^S \otimes \hat{\rho}^E\right) \hat{U}_t^\dagger\right]}\,, \tag{5.79}$$

where $\hat{P}_m = |\, e_m\rangle\, \langle e_m\,|$ and

$$\hat{\rho}^E = \sum_j w_j\, |\, e_j\rangle\langle e_j\,|\ , \tag{5.80}$$

the w_j being weights, i.e. real and positive numbers. On the outline of Eq. (5.74), the net transformation acting on the state of the object system may then be written as

$$\begin{aligned}\hat{\mathcal{T}}_m\hat{\rho}^S &= \mathrm{Tr}_{\mathcal{E}}\left[\hat{P}_m\hat{U}_t\left(\hat{\rho}^S\otimes\hat{\rho}^E\right)\hat{U}_t^\dagger\hat{P}_m\right]\\ &= \sum_{jk} w_j \mathrm{Tr}_{\mathcal{E}}\left[\hat{P}_k\hat{P}_m\hat{U}_t\left(\hat{\rho}^S\otimes|e_j\rangle\langle e_j|\right)\hat{U}_t^\dagger\hat{P}_m\hat{P}_k\right]\\ &= \sum_{jk} w_j \left\langle e_k\left|\left(|e_k\rangle\langle e_k|\,\hat{P}_m\hat{U}_t\left(\hat{\rho}^S\otimes|e_j\rangle\langle e_j|\right)\hat{U}_t^\dagger\hat{P}_m\,|e_k\rangle\langle e_k|\right)\right|e_k\right\rangle\\ &= \sum_{jk}\hat{\vartheta}_{jk}\hat{\rho}^S\hat{\vartheta}_{jk}^\dagger\,, \end{aligned} \tag{5.81}$$

where

$$\hat{\vartheta}_{jk} = \sqrt{w_j}\left\langle e_k\left|\hat{P}_m\hat{U}_t\right|e_j\right\rangle \text{ and } \hat{\vartheta}_{jk}^\dagger = \sqrt{w_j}\left\langle e_j\left|\hat{P}_m\hat{U}_t^\dagger\right|e_k\right\rangle\,. \tag{5.82}$$

Note that this corresponds to the polar decomposition of the amplitude operators, namely, in a projection-like part and a unitary part. These formulae are more general than the previous ones because they allow for the possibility of diverse codifications of the apparatus or the environment, as pointed out in Sect. 3.3.7.

Lindblad's Master Equation

In order tho deal with measurement problems and especially the relation between a system and a lager environment, a formal tool like the master equation is very helpful. Here, I consider in particular the *Lindblad's master equation*,[98] after the name of the Swedish mathematical physicist Göran Lindblad, which, in the Schrödinger picture for density matrices, can be written in the form

$$\frac{d}{dt}\hat{\rho}(t) = \hat{\mathcal{L}}\hat{\rho}(t)\,, \tag{5.83}$$

where $\hat{\mathcal{L}}$ is the *Lindblad operator*[99] for the quantum Liouville tranformation (Eqs. (2.19)–(2.20)) given by

$$\hat{\mathcal{L}}\hat{\rho} = \left(\hat{\mathcal{L}}_{\mathrm{d}} + \hat{\mathcal{L}}_{\mathrm{nd}}\right)\hat{\rho}\,, \tag{5.84}$$

where the non-dissipative term

$$\hat{\mathcal{L}}_{\mathrm{nd}}\hat{\rho} = -\frac{i}{\hbar}\left[\hat{H},\hat{\rho}\right] \tag{5.85a}$$

[98] See Lindblad (1983). For derivation of this equation see Auletta et al. (2009, Sect. 14.2). On Master equation and decoherence see Schlosshauer (2007, Chap. 4).

[99] Which is again a superoperator, a formal aspect that does not need to be considered here.

represents the unitary evolution given by the von Neumann Eq. (2.17), while the term $\hat{\mathcal{L}}_d$ represents the dissipative part given by

$$\hat{\mathcal{L}}_{\mathrm{d}}\hat{\rho} = \frac{1}{2}\sum_j \left(\left[\hat{\Lambda}_j\hat{\rho}, \hat{\Lambda}_j^\dagger\right] + \left[\hat{\Lambda}_j, \hat{\rho}\hat{\Lambda}_j^\dagger\right]\right) , \tag{5.85b}$$

where, in this context, the set of operators $\{\hat{\Lambda}_j\}$ could represent the action of external harmonic oscillators on a system in a bath of such oscillators. In this case, they could stand for annihilation operators on photons (absorption of photons) of the system.

A Two-Level Atom Coupled with Vacuum

Let us consider, for the sake of simplicity, the specific example of a two-level atom coupled to the vacuum electromagnetic field and undergoing spontaneous emission.[100] Therefore, the field is assumed to be in an initial state $\hat{\rho}_E = |0\rangle\langle 0|$ with no photon present. The coherent part of the atom's evolution is described by the Hamiltonian

$$\hat{H} = \frac{1}{2}\hbar\omega\hat{\sigma}_z , \tag{5.86}$$

where $\hbar\omega$ is the energy difference of the atomic levels, $\hat{\sigma}_z$ is the Pauli Z matrix, and

$$\omega = 2\pi\nu \tag{5.87}$$

is the angular frequency, involved in oscillations and rotations, e.g. a rigid rotator, whether classical or quantum-mechanical, with momentum of inertia I, its angular momentum (1.217) is $\mathbf{L} = \mathrm{I}\omega$. Let us consider the most basic model with a single photon that can be absorbed or emitted. This is particularly instructive since it connects immediately this formalism with that of quantum information. The emission of a photon, bringing the atom from the excited level (state) $|1\rangle$ (a photon has been gained before) to the ground state $|0\rangle$ (where the photon is emitted), is described by the operator

$$\hat{\Lambda}_- = \sqrt{2\gamma}\hat{\sigma}_- , \tag{5.88}$$

that is present in the dissipative part (5.85b) of the master equation, where

$$\hat{\sigma}_- = \hat{\sigma}_x - i\hat{\sigma}_y = |0\rangle\langle 1| = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \tag{5.89a}$$

is the so-called lowering operator (bringing the system from the excited to the ground state: see also Eqs. (1.241)), $\hat{\sigma}_x$, $\hat{\sigma}_y$ are the usual Pauli matrices, and $\gamma \in \mathbb{R}$ is the rate of spontaneous emission (loss of the photon), depending on the physical context. The

[100] I follow here (Nielsen and Chuang 1997, Sect. 8.3.5; Auletta et al. 2009, Sect. 14.3).

other possible process, namely, the absorption of a photon, which would be described by the following raising operator (bringing the system from the ground to the excited state):

$$\hat{\sigma}_+ = \hat{\sigma}_x + i\hat{\sigma}_y = |1\rangle\langle 0| = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \tag{5.89b}$$

is not allowed at the start since initially no photons are present in the radiation field (and the atom is supposed to be in the excited state). From Eqs. (5.84)–(5.85) and making use of Eq. (5.88), we immediately obtain the master equation ruling this process as

$$\begin{aligned} \hat{\mathcal{L}}\hat{\rho} = \frac{d\hat{\rho}}{dt} &= -\frac{i}{\hbar}\left[\hat{H}, \hat{\rho}\right] + \frac{1}{2}\left(\left[\hat{\Lambda}_-\hat{\rho}, \hat{\Lambda}_+\right] + \left[\hat{\Lambda}_-, \hat{\rho}\hat{\Lambda}_+\right]\right) \\ &= -\frac{i}{\hbar}\left[\hat{H}, \hat{\rho}\right] + \gamma\left(2\hat{\sigma}_-\hat{\rho}\hat{\sigma}_+ - \hat{\sigma}_+\hat{\sigma}_-\hat{\rho} - \hat{\rho}\hat{\sigma}_+\hat{\sigma}_-\right), \end{aligned} \tag{5.90}$$

where $\hat{\rho}$ is the density matrix of the two-level atom and $\hat{\Lambda}_+ = \sqrt{2\gamma}\hat{\sigma}_+$. Now, we shall see that this master equation can be reduced to Eq. (5.72) by replacing $\hat{\rho}$ by its unitarily evolved state $\hat{\rho}'$:

$$\hat{\rho}'(t) = \hat{U}^\dagger\hat{\rho}\hat{U} = e^{\frac{i}{\hbar}\hat{H}t}\hat{\rho}e^{-\frac{i}{\hbar}\hat{H}t}. \tag{5.91}$$

This allows us to drop the non-dissipative term (5.85a), and, since

$$\hat{\sigma}'_- = e^{\frac{i}{\hbar}\hat{H}t}\hat{\sigma}_-e^{-\frac{i}{\hbar}\hat{H}t} \quad \text{and} \quad \hat{\sigma}'_+ = e^{\frac{i}{\hbar}\hat{H}t}\hat{\sigma}_+e^{-\frac{i}{\hbar}\hat{H}t}, \tag{5.92}$$

the master equation for $\hat{\rho}'(t)$ becomes

$$\hat{\mathcal{L}}\hat{\rho}' = \gamma\left(2\hat{\sigma}'_-\hat{\rho}\hat{\sigma}'_+ - \hat{\sigma}'_+\hat{\sigma}'_-\hat{\rho} - \hat{\rho}\hat{\sigma}'_+\hat{\sigma}'_-\right). \tag{5.93}$$

Thanks to the form of energy (5.86), the Hamiltonian and the spin matrices commute, and this can be rewritten as

$$\hat{\mathcal{L}}\hat{\rho}' = \gamma\left(2\hat{\sigma}_-\hat{\rho}'\hat{\sigma}_+ - \hat{\sigma}_+\hat{\sigma}_-\hat{\rho}' - \hat{\rho}'\hat{\sigma}_+\hat{\sigma}_-\right). \tag{5.94}$$

Since we have succeeded in using only the density matrix $\hat{\rho}'$ on both the LHS and the RHS, we can drop the prime for simplifying the notation.

Bloch Vector Representation

In order to solve Eq. (5.94), it is suitable to use the Bloch vector representation for $\hat{\rho}$, after the name of the Swiss physicist Felix Bloch (1905-1983). It is interesting to note that there is a formal similarity between a two-level atom interacting with the

electromagnetic field and a spin–$\frac{1}{2}$ magnetic dipole precessing in a magnetic field. In order to bring out this analogy in the most clear form, let us write the density matrix of the atom in the form

$$\hat{\rho} = \begin{bmatrix} \rho_{00} & \rho_{01} \\ \rho_{10} & \rho_{11} \end{bmatrix} , \tag{5.95}$$

where I have used the basis (4.4) and $\rho_{jk} = \langle j \,|\hat{\rho}|\, k \rangle$, with $\{|\, j\rangle \,, |\, k\rangle\} = \{|\, 0\rangle \,, |\, 1\rangle\}$.

Following expansion (1.326), any density matrix (pure or mixed) $\hat{\rho}$ of a two-dimensional system, using an arbitrary vector $\mathbf{s}$, may be written as

$$\hat{\rho} = \frac{1}{2}\left(\hat{I} + \mathbf{s}\cdot\hat{\boldsymbol{\sigma}}\right) , \tag{5.96}$$

where $\hat{\boldsymbol{\sigma}}$ is the two-dimensional 'spin' operator (Eqs. (1.324)) and

$$\mathbf{s}\cdot\hat{\boldsymbol{\sigma}} = \sum_j s_j \hat{\sigma}_j , \tag{5.97}$$

where $j \in \{x, y, z\}$. The coefficients used here are the Cartesian components of the *Bloch vector*

$$\mathbf{s} \equiv \left(s_x, s_y, s_z\right) , \tag{5.98}$$

defined by its (real) components as

$$s_x = \rho_{01} + \rho_{10} , \tag{5.99a}$$
$$s_y = i\,(\rho_{01} - \rho_{10}) , \tag{5.99b}$$
$$s_z = \rho_{00} - \rho_{11} . \tag{5.99c}$$

In explicit matricial form Eq. (5.96) may be written as

$$\hat{\rho} = \frac{1}{2}\begin{bmatrix} 1 + s_z & s_x - i s_y \\ s_x + i s_y & 1 - s_z \end{bmatrix} . \tag{5.100}$$

It is easy to verify that, by substituting expressions (5.99) into the matrix (5.100), we obtain the density matrix (5.95), where we need to consider that $\rho_{00} = 1 - \rho_{11}$.

The first two Cartesian components of the Bloch vector $\mathbf{s}$ are then linked to the coherences (the off-diagonal terms) of the density matrix, while the reverse of the third component (i.e. $\rho_{11} - \rho_{00}$) is the so-called *population inversion* of the atom. The Bloch vector is a 3D vector contained within a sphere with radius 1. The Bloch sphere representation is then a density matrices' geometric representation (compare with Fig. 1.28, Sect. 1.4.2, and Fig. 3.6, Sect. 3.2.4), and any pure state can be described as a vector pointing to the surface of the Bloch sphere (Fig. 5.16).

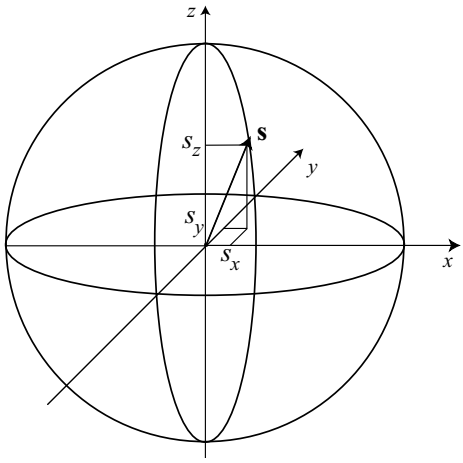


Fig. 5.16 Bloch sphere representation of states of a two-level quantum system. Adapted from Auletta et al. (2009, p. 527)

Equations of Motion

It is instructive to show that calculations made on the two expressions (5.95) and (5.100) lead to the same result. In order to illustrate the formalism of the Bloch vector, let us go back to the master equation (5.94) for a spontaneously emitting two-level atom, with $\hat{\rho}$ replacing $\hat{\rho}'$, as before. We need now to derive the equation of motion for the three components of the Bloch vector. To this purpose, first I recall that

$$\hat{\sigma}_- |0\rangle = 0, \quad \hat{\sigma}_- |1\rangle = |0\rangle \ , \tag{5.101a}$$

$$\hat{\sigma}_+ |0\rangle = |1\rangle \ , \quad \hat{\sigma}_+ |1\rangle = 0 \ , \tag{5.101b}$$

or also, making use of Eqs. (5.89), we have

$$\hat{\sigma}_+ \hat{\sigma}_- = |1\rangle \langle 0|0\rangle \langle 1| = |1\rangle \langle 1| = \hat{P}_1, \tag{5.102a}$$

and

$$\hat{\sigma}_- \hat{\sigma}_+ = |0\rangle \langle 1|1\rangle \langle 0| = |0\rangle \langle 0| = \hat{P}_0. \tag{5.102b}$$

Then, taking into account Eq. (5.94), we may calculate the time derivatives of the elements of $\hat{\rho}$:

$$\begin{aligned} \dot{\rho}_{11} &= \left\langle 1 \left| \dot{\hat{\rho}} \right| 1 \right\rangle = \gamma \left(2 \langle 1 | \hat{\sigma}_- \hat{\rho} \hat{\sigma}_+ | 1 \rangle - \langle 1 | \hat{\sigma}_+ \hat{\sigma}_- \hat{\rho} | 1 \rangle - \langle 1 | \hat{\rho} \hat{\sigma}_+ \hat{\sigma}_- | 1 \rangle \right) \\ &= \gamma \left(2 \langle 1 | 0 \rangle \langle 1 | \hat{\rho} | 1 \rangle \langle 0 | 1 \rangle - \langle 1 | 1 \rangle \langle 1 | \hat{\rho} | 1 \rangle - \langle 1 | \hat{\rho} | 1 \rangle \langle 1 | 1 \rangle \right) \\ &= -2\gamma \langle 1 | \hat{\rho} | 1 \rangle = -2\gamma \rho_{11}. \end{aligned} \tag{5.103a}$$

Analogously, we have for $\dot{\rho}_{00}$

$$\begin{aligned}\dot{\rho}_{00} =&\left\langle 0\left|\dot{\hat{\rho}}\right|0\right\rangle = \gamma\left(2\left\langle 0\left|\hat{\sigma}_-\hat{\rho}\hat{\sigma}_+\right|0\right\rangle - \left\langle 0\left|\hat{\sigma}_+\hat{\sigma}_-\hat{\rho}\right|0\right\rangle - \left\langle 0\left|\hat{\rho}\hat{\sigma}_+\hat{\sigma}_-\right|0\right\rangle\right)\\ =&\gamma\left(2\left\langle 0\mid 0\right\rangle\left\langle 1\left|\hat{\rho}\right|1\right\rangle\left\langle 0\mid 0\right\rangle - \left\langle 0\mid 1\right\rangle\left\langle 1\left|\hat{\rho}\right|0\right\rangle - \left\langle 0\left|\hat{\rho}\right|1\right\rangle\left\langle 1\mid 0\right\rangle\right)\\ =&2\gamma\left\langle 1\left|\hat{\rho}\right|1\right\rangle = 2\gamma\rho_{11}. \end{aligned} \tag{5.103b}$$

Finally, the time derivative of the matrix elements ρ_{01}, ρ_{10} are

$$\begin{aligned}\dot{\rho}_{01} =&\left\langle 0\left|\dot{\hat{\rho}}\right|1\right\rangle = \gamma\left(2\left\langle 0\left|\hat{\sigma}_-\hat{\rho}\hat{\sigma}_+\right|1\right\rangle - \left\langle 0\left|\hat{\sigma}_+\hat{\sigma}_-\hat{\rho}\right|1\right\rangle - \left\langle 0\left|\hat{\rho}\hat{\sigma}_+\hat{\sigma}_-\right|1\right\rangle\right)\\ =&\gamma\left(2\left\langle 0\mid 0\right\rangle\left\langle 1\left|\hat{\rho}\right|1\right\rangle\left\langle 0\mid 1\right\rangle - \left\langle 0\mid 1\right\rangle\left\langle 1\left|\hat{\rho}\right|1\right\rangle - \left\langle 0\left|\hat{\rho}\right|1\right\rangle\left\langle 1\mid 1\right\rangle\right)\\ =&-\gamma\left\langle 0\left|\hat{\rho}\right|1\right\rangle = -\gamma\rho_{01}, \end{aligned} \tag{5.103c}$$

$$\begin{aligned}\dot{\rho}_{10} =&\left\langle 1\left|\dot{\hat{\rho}}\right|0\right\rangle = \gamma\left(2\left\langle 1\left|\hat{\sigma}_-\hat{\rho}\hat{\sigma}_+\right|0\right\rangle - \left\langle 1\left|\hat{\sigma}_+\hat{\sigma}_-\hat{\rho}\right|0\right\rangle - \left\langle 1\left|\hat{\rho}\hat{\sigma}_+\hat{\sigma}_-\right|0\right\rangle\right)\\ =&\gamma\left(2\left\langle 1\mid 0\right\rangle\left\langle 1\left|\hat{\rho}\right|1\right\rangle\left\langle 0\mid 0\right\rangle - \left\langle 1\mid 1\right\rangle\left\langle 1\left|\hat{\rho}\right|0\right\rangle - \left\langle 1\left|\hat{\rho}\right|1\right\rangle\left\langle 1\mid 0\right\rangle\right)\\ =&-\gamma\left\langle 1\left|\hat{\rho}\right|0\right\rangle = -\gamma\rho_{10}, \end{aligned} \tag{5.103d}$$

respectively. The solutions of these differential equations are then given by

$$\rho_{11}(t) = e^{-2\gamma t}\rho_{11}(t_0), \tag{5.104a}$$

$$\rho_{00}(t) = 1 + e^{-2\gamma t}\left[\rho_{00}(t_0) - 1\right], \tag{5.104b}$$

$$\rho_{01}(t) = e^{-\gamma t}\rho_{01}(t_0), \tag{5.104c}$$

$$\rho_{10}(t) = e^{-\gamma t}\rho_{10}(t_0), \tag{5.104d}$$

where, for deriving the second solution, I have used following procedure:

$$\rho_{00}(t) = 1 - \rho_{11}(t) = 1 - e^{-2\gamma t}\rho_{11}(t_0) \tag{5.105a}$$

$$= 1 - e^{-2\gamma t}\left[1 - \rho_{00}(t_0)\right]\ . \tag{5.105b}$$

From these results and making use of expressions (5.99), we can finally obtain the time derivatives of the Bloch vector's components:

$$\dot{s}_x = \dot{\rho}_{01} + \dot{\rho}_{10} = -\gamma\left(\rho_{01} + \rho_{10}\right) = -\gamma s_x\ , \tag{5.106a}$$

$$\dot{s}_y = i\left(\dot{\rho}_{01} - \dot{\rho}_{10}\right) = -i\gamma\left(\rho_{01} - \rho_{10}\right) = -\gamma s_y\ , \tag{5.106b}$$

$$\dot{s}_z = \dot{\rho}_{00} - \dot{\rho}_{11} = 2\gamma\rho_{11} + 2\gamma\rho_{11} = 4\gamma\rho_{11} = 2\gamma\left(1 + s_z\right)\ . \tag{5.106c}$$

The solutions of these differential equations are finally

$$s_x(t) = e^{-\gamma t}s_x(t_0)\ , \tag{5.107a}$$

$$s_y(t) = ie^{-\gamma t}\left[\rho_{01}(t_0) - \rho_{10}(t_0)\right] = e^{-\gamma t}s_y(t_0)\ , \tag{5.107b}$$

$$\begin{aligned} s_z(t) =& \rho_{00}(t) - \rho_{11}(t) = 1 + e^{-2\gamma t}[\rho_{00}(t_0) - \rho_{11}(t_0) - 1] \\ =& 1 + e^{-2\gamma t} s_z(t_0) - e^{-2\gamma t} \, . \end{aligned} \tag{5.107c}$$

Taking into account results (5.104) (or also (5.107)), it is helpful to write here the evolved density matrix of the atom:

$$\hat{\rho}(t) = \begin{bmatrix} 1 + e^{-2\gamma t}(\rho_{00} - 1) & e^{-\gamma t}\rho_{01} \\ e^{-\gamma t}\rho_{10} & e^{-2\gamma t}\rho_{11} \end{bmatrix} . \tag{5.108}$$

Back to Amplitude Operators

Let us now come back to the formalism of amplitude operators. This evolution, taking into account the transformation (5.72), is equivalent to

$$\hat{\rho}(t) = \hat{\mathcal{T}}\hat{\rho}(t_0) = \hat{\vartheta}_0\hat{\rho}(t_0)\hat{\vartheta}_0^\dagger + \hat{\vartheta}_1\hat{\rho}(t_0)\hat{\vartheta}_1^\dagger \, , \tag{5.109}$$

where $\hat{\rho}(t_0)$ is the matrix (5.95)—or (5.100) if we use results (5.107) instead of (5.104)—and

$$\hat{\vartheta}_0 = \begin{bmatrix} 1 & 0 \\ 0 & e^{-\gamma t} \end{bmatrix} \, , \quad \hat{\vartheta}_1 = \begin{bmatrix} 0 & \sqrt{1 - e^{-2\gamma t}} \\ 0 & 0 \end{bmatrix} \tag{5.110}$$

are amplitude operators, where γ is a real number. Note that we also have

$$\hat{\vartheta}_0^\dagger = \hat{\vartheta}_0 \text{ and } \hat{\vartheta}_1^\dagger = \begin{bmatrix} 0 & 0 \\ \sqrt{1 - e^{-2\gamma t}} & 0 \end{bmatrix} . \tag{5.111}$$

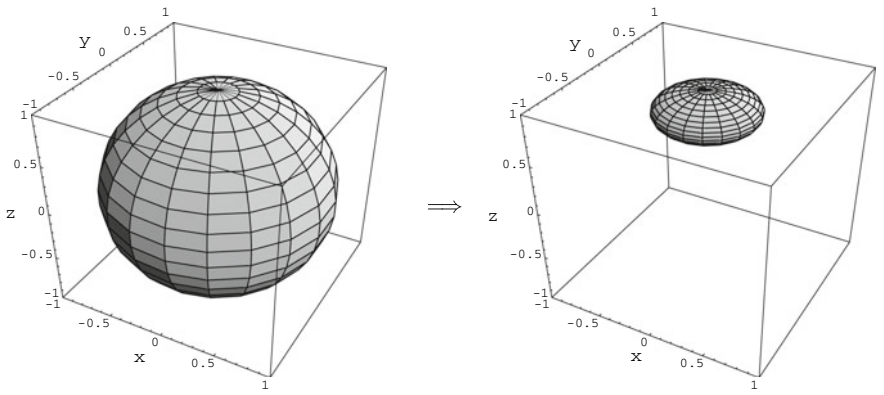


Fig. 5.17 Amplitude damping as a contraction of the Bloch sphere (where the probability of contraction is $\wp_1 = 0.8$). Adapted from Nielsen and Chuang (2000, p. 383)

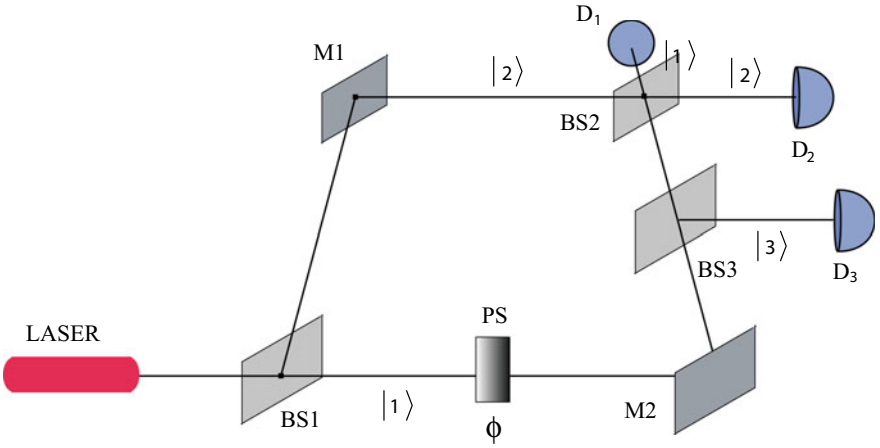


Fig. 5.18 Example of POVMs by means of an interferometry experiment. A single photon state enters from the left Mach–Zehnder interferometer with phase shift ϕ. In the right arm of the interferometer an additional beam splitter BS3 is inserted, with transmission parameter $\sqrt{\eta}$. Three detectors are placed at three outputs of the interferometer. Adapted from Auletta et al. (2009, p. 334)

By inserting these expressions into Eq. (5.109), we obtain in fact the matrix (5.108). Let us make the substitution $\gamma' = 1 - e^{-2\gamma t}$, so that we have

$$e^{-\gamma t} = \sqrt{1 - \gamma'} \text{ and } \sqrt{1 - e^{-2\gamma t}} = \sqrt{\gamma'}, \tag{5.112}$$

I note that, in the case we are considering, the process (5.109) describes *amplitude damping* and $\gamma' = \sin^2 \theta$, according to the model, represents the probability of spontaneous emission, which tends to 1 as t goes to infinity. Note also that such an amplitude damping induces a shrink of the Bloch sphere as shown in Fig. 5.17.

It is interesting to remark that the amplitude operator formalism is more general than that of the master equation. In fact, the amplitude operator formalism may describe state changes without the assumption of a continuous time evolution. Nevertheless, both these formal tools allow us to express a dynamics that is globally reversible but locally irreversible according to the previous analysis of the entrenchment of global and local aspects (Sect. 5.1.3).

Interferometry with Unsharp Observables

I have said that the POVM allows to measure intermediate cases between conjugate observables (the so-called *unsharp observables*), which represents an instance of the analysis of complementarity developed in Sect. 5.1.4. Let us consider a simple example[101]: the experimental arrangement shown in Fig. 5.18. While BS1 and BS2

[101] I follow here (Martens and de Muynck 1990; de Muynck et al. 1991). See also Auletta et al. (2009, Sect. 9.10.3).

are 50–50, the beam splitter BS3 has a transmission parameter $\sqrt{\eta} \in \mathbb{R}$. The initial state of the photon is

$$|\psi\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \tag{5.113}$$

which is parallel to path 1. The first beam splitter, the mirror M1, and the phase shifter induce the following transformation on the initial state $|1\rangle$:

$$|\psi\rangle = |1\rangle \overset{\text{BS1,M1,PS}}{\longmapsto} \frac{1}{\sqrt{2}} \left(e^{i\phi} |1\rangle + |2\rangle \right) , \tag{5.114}$$

which, after the second mirror and BS3, becomes

$$\overset{\text{M2,BS3}}{\longmapsto} \frac{1}{\sqrt{2}} \left[e^{i\phi} \left(\sqrt{\eta} |1\rangle + \sqrt{1-\eta} |3\rangle \right) + |2\rangle \right] . \tag{5.115}$$

The final state, after BS2, making use always of the notorious expressions (2.30), can be written as

$$\begin{aligned} |f\rangle &= \frac{1}{\sqrt{2}} \left[\frac{1}{\sqrt{2}} e^{i\phi} \sqrt{\eta} \left(|1\rangle + |2\rangle \right) + e^{i\phi} \sqrt{1-\eta} |3\rangle + \frac{1}{\sqrt{2}} \left(|1\rangle - |2\rangle \right) \right] \\ &= \frac{1}{2} \left[\left(e^{i\phi} \sqrt{\eta} + 1 \right) |1\rangle + \left(e^{i\phi} \sqrt{\eta} - 1 \right) |2\rangle + e^{i\phi} \sqrt{2(1-\eta)} |3\rangle \right] , \end{aligned} \tag{5.116}$$

where in the last line I have grouped the coefficients for the three components. The detection probabilities can then be computed as follows:

$$\wp_1 = \frac{1}{4} \left(e^{i\phi} \sqrt{\eta} + 1 \right) \left(e^{-i\phi} \sqrt{\eta} + 1 \right) = \frac{1}{4} (\eta + 1) + \frac{\sqrt{\eta} \cos\phi}{2} , \tag{5.117a}$$

$$\wp_2 = \frac{1}{4} (\eta + 1) - \frac{\sqrt{\eta} \cos\phi}{2} , \tag{5.117b}$$

$$\wp_3 = \frac{1-\eta}{2} , \tag{5.117c}$$

where I have made use of Eqs. (2.34). Note that the last term in both $\wp_1$ and $\wp_2$ expresses the interference. We can distinguish between three possible cases:

Interference Observable

- $\eta = 1$. In this case, in agreement with the results of Sects. 2.1.1 and 5.1.4, the outcome probabilities are

$$\wp_1 = \frac{1}{2} (1 + \cos\phi) = \cos^2 \frac{\phi}{2} , \tag{5.118a}$$

$$\wp_2 = \frac{1}{2}\,(1-\cos\phi) = \sin^2\frac{\phi}{2}\,, \tag{5.118b}$$

$$\wp_3 = 0\,, \tag{5.118c}$$

since, $\forall x$, we have the trigonometric identities

$$\cos(x/2) = \pm\sqrt{\frac{(1+\cos x)}{2}} \text{ and } \sin(x/2) = \pm\sqrt{\frac{(1-\cos x)}{2}}. \tag{5.119}$$

In fact, no photon takes the path $|3\rangle$. Note that we can make use of an analogue of Eq. (1.379) for the first two probabilities:

$$\wp_k = \wp\,(D_k) = \mathrm{Tr}\left(\hat{P}_k^{\mathcal{V}}\,\hat{\rho}\right) = \left\langle \hat{P}_k^{\mathcal{V}} \right\rangle_{\psi}\,, \quad \text{for} \quad k = 1, 2, \tag{5.120}$$

with initial density matrix $\hat{\rho} = |\,\psi\rangle\,\langle\psi\,|$. Projectors $\hat{P}_k^{\mathcal{V}}$, $k = 1, 2$, need to express all kinds of (maximal) interferences. Considering Fig. 4.5, Sect. 4.3, we see that these states must lie on that plane or any parallel plane along the orthogonal z-axis. The more general form to express all possible interferences is to make use of the Euler formulas (1.181) and write the transformed state (5.114) as the generalisation

$$\left|\,\psi'\right\rangle = \cos\frac{\phi}{2}\,|\,1\rangle + ie^{\theta}\sin\frac{\phi}{2}\,|\,2\rangle\,, \tag{5.121}$$

where I have used a 'reversed transformation' relative to the usual one on the Poincaré sphere shown in Fig. 3.6, Sect. 3.2.4 (we focus here on parallels instead of meridians). Since the angle θ is irrelevant here, it can be set =0. Thus, making use of $\hat{P}_1^{\mathcal{V}} = \left|\,\psi'\right\rangle\left\langle\psi'\,\right|$, we have for the two interference visibility detections

$$\hat{P}_1^{\mathcal{V}} = \left(\cos\tfrac{\phi}{2}\;i\sin\tfrac{\phi}{2}\right)\left(\cos\tfrac{\phi}{2}\;-i\sin\tfrac{\phi}{2}\right) = \begin{bmatrix} \cos^2\frac{\phi}{2} & -\frac{1}{2}\sin\phi \\ \frac{1}{2}\sin\phi & \sin^2\frac{\phi}{2} \end{bmatrix}, \tag{5.122a}$$

$$\hat{P}_2^{\mathcal{V}} = \hat{I} - \hat{P}_1^{\mathcal{V}} = \begin{bmatrix} \sin^2\frac{\phi}{2} & \frac{1}{2}\sin\phi \\ -\frac{1}{2}\sin\phi & \cos^2\frac{\phi}{2} \end{bmatrix}, \tag{5.122b}$$

because for projectors we have the property (1.94) and I have used the trigonometric property $\cos\frac{x}{2}\sin\frac{x}{2} = \frac{1}{2}\sin x$. Note that the above forms of projectors can always be transformed in their usual diagonalized ones, e.g. (1.84). To reduce projectors to their standard form, it suffices to consider the eigenstates of the interference observable for a given ϕ. It is easy to verify that the above projectors allow us to derive probabilities (5.118), thanks to the mean values (5.120) and the vectorial expression (5.113). In fact,

$$\left\langle\psi\left|\,\hat{P}_1^{\mathcal{V}}\right|\psi\right\rangle = \begin{pmatrix}1 & 0\end{pmatrix}\begin{bmatrix} \cos^2\frac{\phi}{2} & -\frac{i}{2}\sin\phi \\ \frac{i}{2}\sin\phi & \sin^2\frac{\phi}{2} \end{bmatrix}\begin{pmatrix}1\\0\end{pmatrix} = \cos^2\frac{\phi}{2}, \tag{5.123a}$$

$$\left\langle \psi \left| \hat{P}_2^{\mathcal{V}} \right| \psi \right\rangle = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{bmatrix} \sin^2 \frac{\phi}{2} & \frac{i}{2} \sin \phi \\ -\frac{i}{2} \sin \phi & \cos^2 \frac{\phi}{2} \end{bmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \sin^2 \frac{\phi}{2}. \quad (5.123b)$$

Since the projector (5.122) projects into the eigenstates of the interference or superposition observable (see also Eq. (5.35)), in such context the latter can be written as

$$\hat{\mathcal{V}} = \hat{P}_2^{\mathcal{V}} - \hat{P}_1^{\mathcal{V}} , \quad (5.124)$$

with eigenvalues $+1, -1$, respectively.

Which-Path Observable

- $\eta = 0$. In this case, we can clearly distinguish an upper and lower path: component $|2\rangle$ taking the upper path, and split into components $|1\rangle$ and $|2\rangle$ at BS2, detected at either D_1 or D_2, while component $|1\rangle$ takes the lower path and reflected into $|3\rangle$ by BS3 is detected at D_3. Thus, the outcome probabilities are purely classical:

$$\wp_1 = \wp_2 = \frac{1}{4} , \text{ and } \wp_3 = \frac{1}{2} . \quad (5.125)$$

Again, we can make use of projectors for up and down paths

$$\wp_u = \wp_1 + \wp_2 = \left\langle \hat{P}_u^{\mathcal{P}} \right\rangle_\psi , \quad \wp_d = \wp_3 = \left\langle \hat{P}_d^{\mathcal{P}} \right\rangle_\psi , \quad (5.126)$$

where

$$\hat{P}_u^{\mathcal{P}} = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \end{pmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad (5.127a)$$

$$\hat{P}_d^{\mathcal{P}} = \hat{I} - \hat{P}_u^{\mathcal{P}} = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \quad (5.127b)$$

since we have considered the states for which the phase is 0 (i.e. $e^{i\phi} = 1$). Obviously, also these two projectors can be rewritten in the standard form. Then, these projectors represent a resolution of the path observable (see also Eq. (5.32))

$$\hat{\mathcal{P}} = \hat{P}_d^{\mathcal{P}} - \hat{P}_u^{\mathcal{P}} . \quad (5.128)$$

Here, we obviously have a typical classical probability according to a projection-like (or von Neumann's) reduction (Sect. 2.2.3).

POVM

- In all cases where $0 < \eta < 1$ we have a POVM. In this case, we make use of the general probabilities (5.117) which can be computed as

$$\wp_j = \mathrm{Tr}\left(\hat{E}_j \hat{\rho}\right) = \left\langle \hat{E}_j \right\rangle_\psi \,, \quad \text{for} \quad j = 1, 2, 3 \,, \tag{5.129}$$

thanks to the effects operators or POVMs

$$\begin{aligned} \hat{E}_1 =& \frac{1}{2}\left[\left(\hat{P}_u^{\mathcal{P}} + \eta \hat{P}_d^{\mathcal{P}}\right) - \sqrt{\eta}\left(\hat{P}_2^{\mathcal{V}} - \hat{P}_1^{\mathcal{V}}\right)\right] \\ =& \frac{1}{4}\begin{bmatrix} (1+\eta) + 2\sqrt{\eta}\cos\phi & (1-\eta) - 2\sqrt{\eta}\sin\phi \\ (1-\eta) + 2\sqrt{\eta}\sin\phi & (1+\eta) - 2\sqrt{\eta}\cos\phi \end{bmatrix}, \end{aligned} \tag{5.130a}$$

$$\begin{aligned} \hat{E}_2 =& \frac{1}{2}\left[\left(\hat{P}_u^{\mathcal{P}} + \eta \hat{P}_d^{\mathcal{P}}\right) + \sqrt{\eta}\left(\hat{P}_2^{\mathcal{V}} - \hat{P}_1^{\mathcal{V}}\right)\right] \\ =& \frac{1}{4}\begin{bmatrix} (1+\eta) - 2\sqrt{\eta}\cos\phi & (1-\eta) + 2\sqrt{\eta}\sin\phi \\ (1-\eta) - 2\sqrt{\eta}\sin\phi & (1+\eta) + 2\sqrt{\eta}\cos\phi \end{bmatrix}, \end{aligned} \tag{5.130b}$$

$$\hat{E}_3 = (1-\eta)\,\hat{P}_d^{\mathcal{P}} = \frac{1}{2}(1-\eta)\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \tag{5.130c}$$

where I have used the expressions (5.122) and (5.127) as well as the property $\cos^2 x - \sin^2 x = \cos 2x$. The three expectations for the three detectors can be computed as follows (where the means are calculated on the initial state $|\psi\rangle = |1\rangle$):

$$\begin{aligned} \wp_1 &= \begin{pmatrix} 1 & 0 \end{pmatrix} \frac{1}{4}\begin{bmatrix} (1+\eta) + 2\sqrt{\eta}\cos\phi & (1-\eta) - 2\sqrt{\eta}\sin\phi \\ (1-\eta) + 2\sqrt{\eta}\sin\phi & (1+\eta) - 2\sqrt{\eta}\cos\phi \end{bmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= \frac{1}{4}(\eta+1) + \frac{\sqrt{\eta}\cos\phi}{2} \,, \end{aligned} \tag{5.131a}$$

$$\begin{aligned} \wp_2 &= \begin{pmatrix} 1 & 0 \end{pmatrix} \frac{1}{4}\begin{bmatrix} (1+\eta) - 2\sqrt{\eta}\cos\phi & (1-\eta) + 2\sqrt{\eta}\sin\phi \\ (1-\eta) - 2\sqrt{\eta}\sin\phi & (1+\eta) + 2\sqrt{\eta}\cos\phi \end{bmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= \frac{1}{4}(\eta+1) - \frac{\sqrt{\eta}\cos\phi}{2} \,, \end{aligned} \tag{5.131b}$$

and

$$\begin{aligned} \wp_3 &= \begin{pmatrix} 1 & 0 \end{pmatrix} \frac{1-\eta}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= \frac{1-\eta}{2} \,, \end{aligned} \tag{5.131c}$$

in agreement with probabilities (5.117). Therefore, the POVM observable (intermediate between which-path and interference visibility), in analogy with expression (1.96), may be written as

$$\hat{O}_{\mathrm{POV}} = o_1\hat{E}_1 + o_2\hat{E}_2 + o_3\hat{E}_3 \ . \tag{5.132}$$

We see that the fact that effect operators are not necessarily orthogonal implies that their number in a complete set can be larger than the dimensions of the Hilbert space under consideration. Nevertheless, in this particular case, these three effects sum to identity:

$$\begin{aligned}\hat{E}_1 + \hat{E}_2 + \hat{E}_3 &= \frac{1}{2}\left[\hat{P}_u^{\mathcal{P}} - \eta\hat{P}_d^{\mathcal{P}} + \sqrt{\eta}\left(\hat{P}_2^{\mathcal{V}} - \hat{P}_1^{\mathcal{V}}\right)\right] \\ &\quad + \frac{1}{2}\left[\hat{P}_u^{\mathcal{P}} + \eta\hat{P}_d^{\mathcal{P}} + \sqrt{\eta}\left(\hat{P}_2^{\mathcal{V}} - \hat{P}_1^{\mathcal{V}}\right)\right] + (1-\eta)\,\hat{P}_d^{\mathcal{P}} \\ &= \hat{P}_u^{\mathcal{P}} + \eta\hat{P}_d^{\mathcal{P}} + \hat{P}_d^{\mathcal{P}} - \eta\hat{P}_d^{\mathcal{P}} \\ &= \hat{I} \ . \end{aligned} \tag{5.133}$$

It is interesting to note that when we deal with conjugate pairs of observables, like position and momentum, the uncertainty relations determine the amount of smearing for their joint measurability.[102] In the next chapter we shall consider this problem.

The use of POVMs and relative operations allows us to treat measurement avoiding von Neumann's projection postulate (Sects. 2.2.3 and 3.1.1).[103] This shows that the latter is unnecessary. This is also why we can always displace von Neumann's cut,[104] that is, the distinction between systems that we include and those that we exclude froim our description.

Properties as Equivalence Classes

Detection is the operation through which a particular outcome is selected. This outcome allows property ascription. On the other hand, properties may be operationally understood as *equivalence classes of detection events* (see Sect. 5.2.4). In fact, many different detectors and even detection contexts work in the same way. This also shows that properties cannot be identified with quantum events themselves (which are tokens), neither with any form of reality that could be directly experienced, but, as anticipated, are rather inferred given a certain coupling and a certain event. Therefore, they are conceptually dependent on the latter. It is true that when a system is in a given eigenstate of an observable, we often assume that it has the property associated with this state. However, in so doing, we are mixing two different issues[105]:

[102] See Busch et al. (1995, Sect. I.1.2).

[103] D'Ariano et al. (2017, Sect. 2.12).

[104] D'Ariano et al. (2017, Sect. 7.10).

[105] It may be noted that in Ludwig (1983, I, p. 7) preparation–registration procedures are part of the fundamental (operational–ontological) domain of QM but not the assigned properties. For this reason, he stresses that such selection procedures or operations should not be mixed up with formal tools like density matrices or projectors (Ludwig 1983, I, Chap. 3).

1. One issue is that the state can be thought of as the complete catalogue of all probabilities to obtain certain outcomes (the algorithm aspect of the problem).
2. Another issue is whether or not this state is subjected to detection in order to ascertain the property that is associated with one of the outcomes that in fact occurs (the operational aspect).

This distinction is very important and we have already met it in Sects. 3.1.2 and 5.2.3 when speaking of premeasurement. Indeed, any eigenstate of a selected observable, from a pure formal point of view, can be also considered as a superposition (Sect. 3.2.3). In other words, if the system would undergo an experimental procedure selecting an appropriate complementary observable, this possible output eigenstate would no longer instantiate the property we assumed to be real. However, it would be a weird situation if properties did appear and disappear depending on the (mathematical) expansion we are considering.[106]

Thus, properties are not intrinsic to the system. This is the reason why properties are never acquired but inferred given certain detection events.[107] In other words, given a local shift in order and relative changes in information accessibility as witnessed by an irreversible detection event, we are authorised to infer that the system is now in such a state that it is meaningful to ascribe a certain property as a general character (a type–token) that also equivalent events would counterfactually testify to. Thus, a too strict connection between the notion of property and that of eigenvalue associated to the mathematical operator that represents an observable does not help to correctly frame the issue of the nature of properties.

It may be noted that while states and observables are equivalence classes of controlled operations, properties are equivalence classes of the uncontrolled results of operations: they are what several events and systems share. To detect is still a controlled operation, but what we control here are the conditions in which a detection event occurs and the fact that it eventually already occurred, but not its occurring. This is why a detector's clicking/not clicking codifies information. Properties can also be ascribed in the preparation or premeasurement step, but only in a conditional and probabilistic sense when taking into account the whole measurement procedure that the system is submitted to. In fact, in assigning properties without considering any particular procedure, we may incur in situations that are classically impossible, as the Kochen–Specker theorem displays (Sect. 3.3.4). This obviously does not mean that the state is a pure formal entity either (Sect. 5.1.2). However, it needs to be ultimately understood in the context of the dynamic process happening among open systems.

The already quoted work of G. Ludwig has strongly contributed to the notion of states[108] and properties as equivalence classes.[109] About properties, it is worth

[106] Recent attempts at using the so-called weak measurement for showing a supposed dissociation between system and properties (Denkmayr et al. 2014) have subsequently been explained in terms of ordinary quantum interference effects (Corrêa et al. 2015).

[107] See also Rovelli (2005).

[108] Ludwig (1983, I, p. 42).

[109] Ludwig (1983). See also Kraus (1983), Busch et al. (1995).

mentioning that Ludwig preferred the notion of *effect*.[110] He had two reasons for that:

- The first is that after von Neumann the notion of property was identified with a projector, which obviously is a formal–descriptive tool and not an operational one. In other words, a projector does not denote an equivalence class.
- The second was that he likes to introduce a more general notion of property that has subsequently become the POVM theory, and is become crucial for a generalised notion of operation.

Notwithstanding the important generalisation provided by the POVM, the notion of effect is now mostly used by the scientific community in the same way as previously the notion of property (for denoting tokens), the only difference being that the latter should denote for most physicists projectors, and the former effects operators like those described by Eq. (5.64).[111] In order to maintain a fundamental distinction in the spirit (if not in the word) of Ludwig, I prefer to use the term *effect* for denoting effect operators as a generalisation of projectors, and the term *property* for denoting both sharp and unsharp (in the sense of the POVM) properties understood as equivalence classes of detection events that could be represented by the action of projectors or effect operators.

5.2.5 *Generalised Operations*

The Concept of Operation

We may recall that the von Neumann's projection postulate requires orthogonality (Sect. 2.2.3). The enlarged vision of measurement allowed by effect operators helps us to circumvent the conflict between unitary and non-unitary transformations. What is crucial is that renouncing orthogonality does not impair the discriminability of observation (Sect. 5.2.4). Thus, our examination so far suggests that we introduce a general notion of physical transformation that can cover both reversible (unitary) ones and local irreversible ones: a useful generalisation of the class of quantum transformations. This also implies a generalisation of the notion of measurement (that we in fact already introduced but not in a formal way). This generalisation is provided by the concept of *operation*. An *operation* is defined as follows.[112] An operation $\hat{\mathcal{T}}$ is a positive linear mapping from a state space into another (possibly itself), which satisfies the following requirement (known as *Kraus condition*,[113] after the name of the German physicist Karl Kraus (1938–1988)):

[110] Ludwig (1983, I, pp. 31, 43–47).

[111] It might be noted that already Kraus, although still maintaining the notion of equivalence class, prefers to speak of equivalence classes of instruments and not of operations (Kraus 1983, p. 6).

[112] See Davies (1976, pp. 17–18). For what follows see also Auletta et al. (2009, Sect. 9.10).

[113] See Kraus (1983).

$$0 \leq \mathrm{Tr}(\hat{\mathcal{T}}\hat{\rho}) \leq \mathrm{Tr}(\hat{\rho}) \tag{5.134}$$

and the norm $\| \mathcal{T} \|$ is defined as

$$\| \mathcal{T} \| = \mathrm{Sup}\left\{\mathrm{Tr}\left(\hat{\mathcal{T}}\hat{\rho}\right)\right\} \tag{5.135}$$

$\forall\hat{\rho} \in \mathcal{H}$, where $\mathcal{H}$ is the Hilbert space of the system. In other words, the operation $\hat{\mathcal{T}}$ does not necessarily preserve the trace. When the operation is trace-preserving, we have quantum-deterministic (unitary) dynamics Eq. (1.372). It is important to remark that the generalisation provided by the notion of operation makes transformations like (5.61) or (5.81) particular cases when a local (partial) irreversible dynamics is described. Note that operations, states and effects build all convex sets (Sects. 1.4.1 and 5.2.4).

Previously, I have mentioned that $\hat{\mathcal{T}}$ (as a general symbol of a wide range of operations that will be, at least in part explored) is a superoperator. Now we can understand such a notion in more formal terms: a superoperator is a second-level operator (often being a combination of first-level operators) acting (from both the left and the right) on (other) first-level operators. For distinguishing first-level and second-level operators in the following, I shall always denote the latter with calligraphic letters like $\mathcal{T}$.[114] Since the operation $\hat{\mathcal{T}}$ is represented by a superoperator acting on the density operator $\hat{\rho}$, it is opportune to introduce a specific formalism allowing to threat the density operator as a ket.

D-Shaped Kets and Bras

Generally speaking, due to the resolution of identity (1.94) of any projector operator, any operator $\hat{O}$ may be represented in the following form:

$$\hat{O} = \sum_j \sum_k |\, j\rangle \langle j \,|\, \hat{O} \,|\, k\rangle \langle k \,| = \sum_{j,k} |\, j\rangle \langle j \,|\, \hat{O} \,|\, k\rangle \langle k \,| \ , \tag{5.136}$$

where $\{|\, j\rangle\}, \{|\, k\rangle\}$ are arbitrary alternative orthonormal bases on the underlying Hilbert space $\mathcal{H}$: this can be considered as a generalisation of the first of Eqs. (1.95). In this way, operators may be considered as—not necessarily normalised—'vectors' in a super Hilbert space[115], which is the direct product of the original Hilbert space $\mathcal{H}$ and its dual $\mathcal{H}^*$ (Sect. 1.2.1). In fact, we can rewrite the previous equation as

$$\hat{O} = \sum_{j,k} o_{jk} \,|\, j\rangle \langle k \,| \ , \tag{5.137}$$

[114] Properly speaking, also effect operators are superoperator, but for them I follow the general convention.

[115] See Royer (1989).

where

$$o_{jk} = \left\langle j \left| \hat{O} \right| k \right\rangle \tag{5.138}$$

are the matrix elements of $\hat{O}$. Note that the 'cross' operators $|\, j\rangle\, \langle k\,|$ are not projectors. Therefore, we may associate to any operator $\hat{O}$ a D-shaped ket $\left| \hat{O} \right)$ and a D-shaped bra $\left(\hat{O} \right|$, defined by the superpositions

$$\left| \hat{O} \right) := \sum_{j,k} o_{jk} \,|\, j, k) \quad , \tag{5.139a}$$

$$\left(\hat{O} \right| := \sum_{j,k} o^*_{jk} \, (j, k\,| \quad , \tag{5.139b}$$

where

$$|\, j, k) \; := |\, j\rangle\, \langle k\,| \, , \tag{5.140a}$$

$$(j, k\,| := |\, k\rangle\, \langle j\,| \tag{5.140b}$$

represent the basis in which the D-shaped ket (and the D-shaped bra) is expanded. Their scalar product may be represented as

$$(l, m\,|\, j, k\,) = \langle l\,|\; j\rangle\, \langle k\,|\; m\rangle = \delta_{l,j}\delta_{k,m}, \tag{5.141}$$

from which, together with Eqs. (5.140), the generalised scalar product follows

$$\begin{aligned} \left(\hat{O} \left| \hat{O}' \right. \right) &= \sum_{l,m,j,k} o^*_{lm} o'_{jk} \, (l, m\,|\, j, k\,) \\ &= o^*_{lm} o'_{jk} \delta_{l,j} \delta_{k,m} \\ &= \sum_{l,m} o^*_{lm} o'_{lm} \\ &= \sum_{lm} o^\dagger_{ml} o'_{lm} \\ &= \mathrm{Tr} \left(\hat{O}^\dagger \hat{O}' \right) \, , \end{aligned} \tag{5.142}$$

where the trace is over l, and $o^\dagger_{ml}$ is a matrix element of the matrix $\hat{O}^\dagger$ that is the transposed conjugate of $\hat{O}$, and is the counterpart of o_{lm}. From Eq. (5.142) it follows that we may reformulate Eq. (5.138) as

$$\left\langle j \left| \hat{O} \right| k \right\rangle = \sum_{l,m} o_{lm} \delta_{l,j} \delta_{k,m}$$
$$= \left(j, k \left| \sum_{l,m} o_{lm} \right| l, m \right)$$
$$= \left(j, k \left| \hat{O} \right. \right) . \tag{5.143}$$

Probability of Transmission

The previous formalism allows us to symbolise the operation $\hat{\mathcal{T}}$ on $\hat{\rho}$ by

$$\hat{\mathcal{T}} \left| \hat{\rho} \right) . \tag{5.144}$$

We may think $\hat{\mathcal{T}}$ to be a 'test' which is undergone by the system in the state $\hat{\rho}$. Then, using Eq. (5.142), the probability of the transmission of a state $\hat{\rho}$ by an operation $\hat{\mathcal{T}}$ (i.e. the probability that the system in that state passes the test) is

$$\mathrm{Tr} \left(\hat{\mathcal{T}} \hat{\rho} \right) = \left(\hat{\mathcal{T}} \left| \hat{\rho} \right. \right) . \tag{5.145}$$

When the probability is =1, we have deterministic unitary mapping from pure states to pure states in particular, while probabilistic but not deterministic operations are not trace-preserving.

In order to normalise the output or final state $\hat{\rho}_f$, upon transmission, we need to write

$$\hat{\rho}_f = \frac{\hat{\mathcal{T}} \left| \hat{\rho}_i \right)}{\left(\hat{\mathcal{T}} \left| \hat{\rho}_i \right. \right)} , \tag{5.146}$$

where $\hat{\rho}_i$ is some initial state. The previous equation can be considered as a generalisation of Eq. (3.14). This formalism allows us to write

$$\left(\hat{\rho}_f \left| \hat{\mathcal{T}} \right| \hat{\rho}_i \right) \tag{5.147}$$

for expressing the operation bringing a system from the initial state $\left| \hat{\rho}_i \right)$ to a final state $\left| \hat{\rho}_f \right)$.

Deterministic Transformations

For an effect operator $\hat{E}_j$,when having an initial input state $\hat{\rho}_i$, in agreement with Eq. (5.70b), we can write

$$\mathrm{Tr}\left(\hat{E}_j\hat{\rho}_i\right) = \wp\left(j|i\right) = \left(\hat{E}_j\,|\hat{\rho}_i\right). \tag{5.148}$$

This generalisation of Eq. (5.145) allows us to redefine property (1.373b) for any system $\mathcal{S}$ as follows:

$$(e_S\,| = \mathrm{Tr}_{\mathcal{S}}, \tag{5.149}$$

which represents the deterministic transformation

$$\hat{\mathcal{T}} = \sum_j \hat{E}_j = \hat{I}, \tag{5.150}$$

when we have equality sign in Eq. (5.67), and where e is the identity as a pure number. This is obviously satisfied when these effects are in fact projections operators. Thus, when we have an outcome with probability =1 (like the case of a measurement of an observable on a system's state that is already an eigenstate of that observable), this deterministic transformation corresponds to apply the identity operator on the initial state. In other words, for all quantum systems $\mathcal{A}$, $\mathcal{B}$, we represent the tracing out of system $\mathcal{A}$ and the marginal state of system $\mathcal{B}$, i.e. $\mathrm{Tr}_{\mathcal{A}}\,\hat{\rho}_{AB}$ as follows:

$$\mathrm{Tr}_{\mathcal{A}}\left(\hat{\rho}_{AB}\right) = \sum_j \hat{P}_j\hat{\rho}_{AB}\hat{P}_j = \left(e_A\,|\hat{\rho}_{AB}\right) = |\,\hat{\varrho}_B\big)\,, \tag{5.151}$$

for projectors $\hat{P}_j = |\,j\rangle_{\,A}\langle j\,|$ acting on the Hilbert space of the system $\mathcal{A}$ and where $|\,j\rangle_{\,A}$ is an orthonormal basis for that subspace.

Operations on States and Observables

Now, there must exist some sort of equivalence between the application of an operation to the system's state and of a related operation to a given observable of the system. This equivalence is formally expressed by the following theorem[116]: for an operation $\hat{\mathcal{T}}$ there exist amplitude operators $\hat{\vartheta}_k$ (where $k \in \mathbb{N}$ pertains to a finite or infinite set K) on the Hilbert space satisfying, in agreement with Eq. (5.67),

$$\sum_{k\in K}\hat{\vartheta}_k^{\dagger}\hat{\vartheta}_k \leq \hat{I}\,, \tag{5.152}$$

such that, for a given system's observable $\hat{O}$ and an arbitrary state $\hat{\rho}$, the operations $\hat{\mathcal{T}}$ and $\mathcal{T}^*$ are given by (see Eq. (5.61) and also Eq. (2.47))

$$\hat{\mathcal{T}}\,|\hat{\rho}) = \sum_{k\in K}\hat{\vartheta}_k\hat{\rho}\hat{\vartheta}_k^{\dagger}\,, \tag{5.153a}$$

[116] See Kraus (1983, p. 42).

$$\hat{\mathcal{T}}^* \Big| \hat{O}\Big) = \sum_{k\in K} \hat{\vartheta}_k^\dagger \hat{O} \hat{\vartheta}_k \,, \tag{5.153b}$$

respectively, so that the probabilities computed through these two kinds of operations be equal:

$$\Big(\hat{\mathcal{T}}\hat{\rho}\,\Big|\hat{O}\Big) = \Big(\hat{\rho}\,\Big|\hat{\mathcal{T}}^*\hat{O}\Big)\,. \tag{5.154}$$

These transformations are positivity-preserving mappings (they preserve the positivity of the involved operators). Equations (5.153a) and (5.153b), which are known as the *Kraus theorem*, after the name of Karl Kraus, describe the same dynamics from two different perspectives. Therefore, they can be understood as generalised Schrödinger (active) and Heisenberg (passive) transformations, respectively (Eqs. (2.8a) and (2.18)).

Some Considerations

The notion of operation allows a new view of QM: in the words of D'Ariano and co-workers, is "a reversible picture of an irreversible world".[117] What is beautiful, as stressed by the same authors, is that *all maps* that satisfy the formal requirements of the theory are allowed. In other words, any hypothetical transformation that is compatible with the set of states that is allowed in quantum theory must be a physical transformation allowed by the theory.[118]

Since we are considering measurement as a particular and controlled process of a larger class of spontaneous interactions, we support a weaker form of operationalism that focusses on objective processes (ruled ultimately by master equations) than on observation, although our controlled operations still represent the main guiding line when dealing with these problems. A weaker operationalism of this kind has been also called *process operationalism*.[119]

5.2.6 Some Lessons

Summarising, as anticipated in Sect. 3.2.3, each step of measurement (from preparation through premeasurement up to detection) can be understood as a further degree of determination through appropriate selections: of a state, of an observable, and of an outcome. In other words, the whole sequence can be seen as a dynamical process through which, starting from some initial possibilities and thanks to a suitable potential context, the actual reality (the event) is activated and information is acquired.[120]

[117] D'Ariano et al. (2017, Sect. 7.9).

[118] D'Ariano et al. (2017, Sect. 8.12).

[119] Coecke and Lal (2012).

[120] As pointed out in Auletta and Torcal (2011).

Thus, the view of QM supported here is relational–interactional: instead of trying to define systems as they are, we look rather at the way in which they *interact, or are related*, with other systems.[121] This is particularly true for the notion of observable that is rooted in a coupling operation. Let us examine these aspects in details.

According to the objective interpretation of the formalism of QM (that we share so far), possibilities represent something that precedes any specification of physical context. In fact, any context can at most drop some of those possibilities but not generate new ones. Thus, considered from the point of view suggested by quantum physics, possibilities are not evanescent forms of reality or embryonic, as it is often assumed in philosophy[122]: at the opposite, they are *fully predetermined*.[123] Nothing that could happen somehow or somewhere (in whatever physical universe) is out of this combinatorics of possibilities at all levels. For this reason, it makes full sense, as the supporters of the MWI say, to affirm that any possibility has its alternatives. But, precisely because everything is contained in such a combinatorics, nothing is particularly determined in one sense or the other, and therefore possibilities cannot determine anything else. Thus, we have a predetermined but not particularly determined reality that does not determine further happenings by itself either. This is the global aspect of QM.

At the opposite, potential factors (e.g. the experimental conditions in which certain observables are selected and the relative entanglement) are *post-determined* as far as they can bring to determination only in a delayed form and thanks to additional dynamical factors that activate them. As pointed out in Sect. 3.4.3, to say "potential" means exactly this: to contribute to further determination. How can potential realities do that? Precisely because they do not represent an exhaustive catalogue of everything that is possible but both (i) represent a restriction of this space of possibilities and (ii) for this reason they constitute a complex of contingent physical conditions that allow for the further determination of the possible events selected in this way. Note that they constitute these physical conditions because they combine quantum and classical correlations. So, we seem to be justified in affirming that it is the involvement of classical correlations among *different systems* what makes here the difference between possibilities and potentialities and making of interacting quantum systems type–tokens. In other words, a plurality of systems is a necessary condition for potentiality.

This brings us back the problem of the local accessibility to information. Depending on the different local contexts (and their causal constraints), quantum systems, thanks to the presence of the classical part of correlations, become sensitive to different aspects of those local environments (Sects. 5.1.3 and 5.2.3). For instance, in an interferometry experiment, when there is an object the system becomes sensitive to its presence; otherwise, it is sensitive to the setup with two open paths. As a consequence, specific information becomes potentially available, for instance, the

[121] Rovelli (1996, 2005). See also Epperson and Zafiris (2013).

[122] For instance by Peirce (1903, 1.25).

[123] For reasons that are partly different from my approach, the so-called transactional interpretation of QM has stressed the 'reality' of possibility (Kastner 2013, Chap. 4).

information about the presence of the object. Such information about environmental details or aspects of the surroundings is potential information, i.e. information that could be extracted from the system at a later moment (Sect. 3.2.3). Such an information, being contextual (depending on, but also being about, contexts and their details), can be made accessible or not always depending on the changing contexts (typically, during premeasurement). Thus, a quantum system can lose information accessibility to certain local details and might have new access to other local details depending on the change of physical context. Therefore, we need to put together different issues (Sect. 3.4.2):

- On the one hand, the infinite amount of information that is contained in every quantum systems allows a possible sensitivity to any kind of environment and every of its details, but as far as none of them is actual, such amount is *ineffective regarding the physical consequences*: possibilities are types.
- On the other hand, when there are specific physical contexts, the system can become sensitive to certain specific local details and displays information that in fact can be acquired, that is, potential information: the system is a type–token.
- Finally, when some detection events (tokens) happen in that context or even outside but involving systems that can somehow influence our reference system, they will determine a local loss of potential information that cannot be made fully reversible at that local level.

As anticipated, these three degrees of determination are in fact *three kinds of selection*: selection of an in input state, selection of a context and therefore of a code alphabet, and selection of an outcome or information acquisition.

Since quantum systems are sensitive to the details of the environment, including the presence or absence of other quanta, the best hypothesis is that they *map the environmental information in themselves*. In other words, they modify and limit the infinite combinatorics of possibilities that they represent when are free in accordance with what is their environment: whether they are trapped in a potential well or a microcavity or not, whether a certain path is blocked or not, whether particles with which they are entangled become entangled with other quanta, whether there is an isolated magnetic or electric field and so on. All these environmental details will determine the dynamics of the system. This is not far away from Bohm's intuition, although it is to certain extent the reversal of de Broglie–Bohm's pilot wave or active information (Sects. 3.3.4, 3.3.6, and 3.4.1): it is not the wave to pilot the dynamics but is rather the wave that is *piloted* by the environment. It is also true, as mentioned in the quoted subsections, that followers of Bohm seem to move in such a direction. Then, the more the system is confined and limited, the more we have a dominance of *local* correlations on the global ones, according to our analysis in Sect. 5.1.3. Thus, we can conceive an event as the *result of a local cross between the dynamics of two or more quanta* that have maps of their local environment that are partly *incongruent*, as the example of the Aharonov–Bohm effect shows. This means that at least for a time interval they become a single local system (like a photon absorbed by an electron) that need to select that submap that the two quanta shared. Now, the consequence is the collapse into one of the possible components of their compound state. And

this is the event: the two interacting systems need to agree on the same fragment of the *local* environment in order to generate an event. In other words, an event is the final result of the growing dominance of local correlations (especially of their classical components) over the global correlations. Clearly, such a collapse cannot be predetermined: we should be the two or more quanta themselves or know all details of their environmental maps in order to be able to make such predictions. However, although not predictable, events became in this way fully explainable.

On the basis of this inquiry, we can refine the notion of system. In its generality, a *physical system* can be broadly defined as an ordered complex of interrelated elements, rather than "interacting," as originally defined by the Austrian biologist and father of system theory Ludwig von Bertalanffy (1901–1972).[124] The reason is that, according to our previous analysis, these parts can be very well simply correlated, so that local interaction is not necessary to define a system. When dealing with systems, three features must be considered: the number of components, their species and their relations (structure: Sect. 3.4.1). The advantage of this notion is that we are free to extrapolate out of a compound system and consider the part a system on its own as well as we can always consider different systems as a single compounded one provided that they are at least correlated. Thus, a photon is a system, an electromagnetic field is a system and an atom is a system but also its protons and electrons. There is a further advantage: this notion covers both controlled and uncontrolled situations, so that, operationally speaking, we can say that a *system* is whatever we can, at least in principle, submit to one of the local operations discussed in this section. Finally, being a system a type–token, we avoid the problems connected with the classical assumption of the system as a token (an object) (Sects. 3.3.4 and 3.4.3).

5.3 Four Convergent Interpretations

In the previous examination, we have proved four *interpretational no-go theorems*:

1st: *No classical realism* (in the sense of Einstein). Any form of realism need to take into account that there is a basic indetermination and a not perfect localisation of quantum systems, as displayed by the Kochen–Specker theorem and Bell inequalities. Moreover, notions like state, observable and property are equivalence classes and not tokens. The main attack to classical realism comes from Bohr's epistemic phenomenism (with his insistence on the quantum of action, strictly connected with the notion of event happening randomly) and Heisenberg's juvenile instrumentalism.

2nd: *No classical interactionism* (in the sense of both Bohr's phenomenism and especially the young Heisenberg's standpoint). Indirect and interaction-free measurements show that this standpoint needs to be corrected. The main attack to interactionism comes from Schrödinger's objectivism.

[124] von Bertalanffy (1950). See also Auletta (2011, Sect. 2.4.4). Science as knowledge of systems and in particular of relations is the object of Carnap (1928).

3rd: *No epistemic status for quantum states and observables* (as supported by Bohr's phenomenism). Although these notions are types, their ontological substrates are processes that need to have a token component in terms of the real physical systems involved, as displayed by both the PBR and Auletta–Tarozzi (proposed) experiments. The main attack to Bohr's phenomenism comes from both Einstein's realism and Schrödinger's objectivism.

4th: *No physics without specific physical contexts* (physics cannot be reduced to involved correlations, which is the standpoint towards which Schrödinger's objectivism is inclined). Here, the reported experiments of Mandel et al. as well as Scully et al. show how important contexts are. The main attack to Schrödinger's objectivism comes from both Heisenberg's interactionism and Bohr's epistemic phenomenism.

Thus, it appears that the four standpoints of these giants are incompatible. However, we have also remarked some interesting convergences:

- Heisenberg is convergent with Einstein on dynamical realism: interactions are what really counts in physics. And they could also have converged on a dispositional or potential understanding of observables.
- Heisenberg and Schrödinger converge on both (i) the relevance of interference terms for quantum systems and (ii) a kind of rough idea of decoherence: for Heisenberg, the focus is more on the role of the environment, and for Schrödinger the focus is on the notion of local extrapolation. Note that Wheeler's notion of participatory universe is kind of compromise between Schrödinger's standpoint and that of the Copenhagen school.
- Einstein and Schrödinger converge on the fundamental epistemological requirement that the world must be objective and knowable; otherwise, physics loses its main sense, with the consequence that it shows a rational order that does not depend on our description of it.
- Bohr and Einstein could converge on the relevance of the theory also for what we observe and call reality. The main point here is that only a theory (and related laws of nature) can coherently frame our experience.
- Bohr and Heisenberg clearly converge on the relevance of physical contexts. This is perhaps the most important contribution of the Copenhagen school, and it is a pity that Bohr seemed to forget it in his discussion with EPR, especially because Einstein could have agreed as the discussion on the reality of observables shows.

Thus, how can we evaluate such differences and convergences? The divergences were certainly real and are part of the history of physics. However, they also reflected a still immature stage of the theory and especially the lack of many experimental evidences that have been accumulated along the years. Thus, we can guess that more knowledge had persuaded these actors to reinforce their convergence also by fine-tuning their own standpoints.

Thus the question is: is it possible to integrate the different points of view of the major actors of this drama? If we start from *dynamical realism* (since I follow Einstein in ascribing to dynamics a kind of ontological supremacy), but involve

in the concept of locality that of *quantum event* (or quantum of action) and add a correction to our ontology that also comprehends *global quantum features*, then we could integrate (i) Einstein's dynamical realism (local particles although not perfectly localised together with local dynamical interactions) with both (ii) Schrödinger's quantum objectivism (non-local correlation are also real and provide a bridge to the mind) but purified from his denial of local physical contexts, and (iii) Bohr's phenomenism (the quantum of action and therefore (at least observed) events are real) when purified from its epistemic form: a quantum phenomenism. This compromise could be fully accomplished if merge this stuff with (iv) the notion of potentiality (Sect. 3.4.3), that is, with Heisenberg's late *operational realism* but purifying it from its interactionism.

5.4 Summary

- There is a crucial difference between quantum and relativistic events. This conflict could be overcome if we consider coincidence events as an approximation and therefore as a the result of a coarse graining.
- Events, once happening, are fully determined and can give rise to further processes.
- We have defined an event as a singular happening. It has no alternatives, since when it happens the access to any of them is blocked, being the actual event conditionally necessary.
- Probabilities are irreducible and objective.
- Quantum features can only be inferred; we cannot make direct experience of them. One of the major problems is that it is believed that we can only reconstruct them over many experimental runs, which would suggest that quantum features are kind of mathematical or statistical entity. However, a proposed experiment shows that we can find evidence for them in each experimental run.
- The notion of state has an ontological substrate.
- In entanglement, we distinguish a classical and a quantum component (quantum discord). Entanglement with fragments of the universe shows the progressive local dominance of classical correlations during measurement or local interactions.
- We can assume that there is a dynamical entrenchment of global and local aspects. This suggests a reformulation of the complementarity principle in terms of a smooth complementarity, whose the Greenberger–Yasin inequality is the formal expression. Thus, the complementarity principle is disentangled from its epistemic versions.
- We have dealt with the important notion of a participatory universe for accounting for the emergence of the mind. However, events occurred long before the appearance of *homo sapiens*, so that, again, we cannot support a pure epistemic understanding of the Copenhagen interpretation. The idea of the block universe seems to deny the Nowness of any event. It has been shown the impossibility of strict localisation, which would imply a major conflict between relativity and QM.

- Overcoming operational epistemology (consisting in interactionism plus instrumentalism), we need to enlarge the concept of operation. In particular, preparation can be seen as determinative measurement, while premeasurement as an interrogation of a system relative to a degree of freedom. Finally, detection is an answer to our interrogation and allows us to ascribe a property to the object system.
- States are equivalence classes of preparations.
- Premeasurement is a reversible operation. Mandel's experiment shows the connection between reversibility of premeasurement and change of setup: any change of the context (or of the target observable) means a change of the experimental setup. Quantum-erasure experiments show that in fact we flip the information from one system to another. Observables are equivalence classes of premeasurements.
- Detection (and measurement in general) is included in the generalised dynamics of master equation. Effects operators allow the definition of POVM, a new class of observables. Properties are equivalence classes of detections.
- The formalism of operations allows to describe any kind of dynamics of both states and observables. Von Neumann's cut becomes unnecessary.
- Quantum systems map environmental information in themselves.
- The interpretation proposed in the present book is a mix of Einstein's realism, Bohr's phenomenism, Schrödinger's objectivism and Heisenberg's operationalism. All their ideas are very helpful for understanding QM, provided that we avoid here any unilaterally.

References

Aharonov, Y., and D. Bohm. 1959. Significance of Electromagnetic Potentials in the Quantum Theory. *Physical Review* 115: 485–91; Report in [Shapere and Wilczek 1989, 104–110].

Aristotle. 2019. *Liber de Interpretatione*. In [Aristotle 1949]; Also in [Aristotle (1831), 16–24].

Aristotle. 1949. *Categoriae et Liber de Interpretatione* (ed. L. Minio-Paluello). Oxford.

Aristotle. 1831. *Opera* (ed. Immanuel Bekker). Berlin: G. Reimerum.

Armstrong, D.M. 1983. *What is a Law of Nature*. Cambridge: Cambridge University Press.

Atkins, P.W., and J. De Paula. 2006. *Physical Chemistry*, 2nd ed. Oxford: Oxford University Press.

Atkins, P.W., and R. Friedman. 2005. *Molecular Quantum Mechanics*, 4th ed. Oxford: Oxford University Press.

Auletta, G., and (in collaboration with I. Colagè, P. D'ambrosio, and L. Torcal). 2011. *Integrated Cognitive Strategies in a Changing World*. Rome: G and B Press.

Auletta, G., and G. Tarozzi. 2004a. Wavelike Correlations versus Path Detection: Another Form of Complementarity. *Foundations of Physics Letters* 17: 889–95.

Auletta, G., and G. Tarozzi. 2004b. On the Physical Reality of Quantum Waves. *Foundations of Physics* 34: 1675–94.

Auletta, G., and L. Torcal. 2011. From Wave–Particle to Features–Event Complementarity. *International Journal of Theoretical Physics* 50: 3654–3668. https://doi.org/10.1007/s10773-011-0833-8.

Auletta, G., and S.-Y. Wang. 2014. *Quantum Mechanics for Thinkers*. Singapore: Pan Stanford Publication. https://doi.org/10.4032/9789814411721.

Auletta, G., M. Fortunato and G. Parisi. 2009. *Quantum Mechanics*. Cambridge: Cambridge University Press; Revised paperback ed. 2013.

Auletta, G. 2000. *Foundations and Interpretation of Quantum Mechanics: In the Light of a Critical-Historical Analysis of the Problems and of a Synthesis of the Results*. Singapore: World Scientific; Revised ed. 2001.
Auletta, G. 2005. Quantum Information and Inferential Reasoning. *Foundations of Physics* 35: 155–69.
Auletta, G. 2011. *Cognitive Biology: Dealing with Information from Bacteria to Minds*. Oxford: Oxford University Press.
Auletta, G. 2012. Causation Upside Down? *Revista Portuguesa de Filosofia* 68 (1–2): 9–32.
Auletta, G. 2014a. The Quest for Reality. in [Fano 2017, 65–84].
Auletta, G. 2014b. Protective and State Measurement: A Review. in [Gao 2014, 39–62].
Ballentine, L.E. 1970. The Statistical Interpretation of Quantum Mechanics. *Review of Modern Physics* 42: 358–81.
Bennett, C.H., and P.W. Shor. 1998. Quantum Information Theory. *IEEE Transactions on Information Theory* 44: 2724–42.
Bennett, C.H., and S.J. Wiesner. 1992. Communication via One- and Two-Particle Operators on EPR States. *Physical Review Letters* 69: 2881–84.
Bohm, D. 1980. *Wholeness and the Implicate Order*. London: Routledge.
Bokulich, A. 2008. *Reexamining the Quantum-Classical Relation: Beyond Reductionism and Pluralism*. Cambridge: Cambridge University Press.
Borel, E. 1920. *Le hasard*. Paris: Alcan.
Born, M. 1949. *Natural Philosophy of Chance and Cause*. Oxford: Clarendon.
Bridgman, P.W. 1927. *The Logic Of Modern Physics*. New York: McMillan.
Brown, R., and T. Porter. 2006. Category Theory: An Abstract Setting for Analogy and Comparison. In *What is Category Theory?*, ed. G. Sica, 257–74. Milano: Polimetrica.
Busch, P., M. Grabowski, and P.J. Lahti. 1995. *Operational Quantum Physics*. Berlin: Springer.
Carnap, R. 1928. *Der logische Aufbau der Welt*. Meiner: Hamburg.
Coecke, B., and R. Lal. 2012. Categorial Quantum Mechanics Meets the Pavia Principles: Towards a Representation Theorem for CQM Constructions. in [Jacobs et al. 2012, 67–78].
Corrêa, R., M.F. Santos, C.H. Monken and P.L. Saldanha. 2015. Quantum Cheshire Cat' as Simple Quantum Interference. *New Journal of Physics* 17: 053042. https://doi.org/10.1088/1367-2630/17/5/053042.
D'Ariano, G.M., and A. Tosini. 2013. Emergence of Space–Time from Topologically Homogeneous Causal Networks. *Studies in History and Philosophy of Modern Physics* 44:294–99. https://doi.org/10.1016/j.shpsb.2013.04.003
D'Ariano, G.M., and H.-P. Yuen. 1996. Impossibility of Measuring the Wave Function of a Single Quantum System. *Physical Review Letters* 76: 2832–2835.
D'Ariano, M., G. Chiribella, and P. Perinotti. 2017. *Quantum Theory from First Principles: An informational approach*. Cambridge: Cambridge University Press.
Davies, E.B. 1976. *Quantum Theory of Open Systems*. London: Academic.
de Muynck, W.M. 2002. *Foundations of Quantum Mechanics, an Empiricist Approach*. New York: Kluwer Academic Publishers.
de Muynck, W.M., W.W. Stoffels, and H. Martens. 1991. Joint Measurement of Interference and Path Observables in Optics and Neutron Interferometry. *Physica B* 175: 127–32.
Denkmayr, T., H. Geppert S. Sponar, H. Lemmel, A. Matzkin, J. Tollaksen and Y. Hasegawa. 2014. Observation of a Quantum Cheshire Cat in a Matter–Wave Interferometer Experiment. *Nature Communications* 5: 4492. https://doi.org/10.1038/ncomms5492.
Deutsch, D. 2011. *The Beginning of Infinity: Explanations that Transform the World*. London: Penguin Books.
Dirac, P.A.M. 1945. On the Analogy between Classical and Quantum Mechanics. *Review of Modern Physics* 17: 195–99.
Eddington, A.S. 1934. On the Method of Theoretical Physics. *Philosophy of Science* 1:163–69; Eng. Tr. of [Eddington (1930)].

Elitzur, A.C. 1992. Locality and Indeterminism Preserve the Second Law. *Physics Letters* A167: 335–40. https://doi.org/10.1016/0375-9601(92)90268-q
Elitzur, A.C., S. Dolev, and N. Kolenda (eds.). 2005. *Quo Vadis Quantum Mechanics?*. Berlin: Springer.
Englert, B.-G., M.O. Scully, and H. Walther. 1995. Complementarity and Uncertainty. *Nature* 375: 367–68.
Englert, B.-G. 1996. Fringe Visibility and Which-Way Information: An Inequality. *Physical Review Letters* 77: 2154–57.
Epperson, M., and E. Zafiris. 2013. *Foundations of Relational Realism: A Topological Approach to Quantum Mechanics and the Philosophy of Nature*. Lexington Books.
Finkelstein, D.R. 1996. *Quantum Relativity*. Berlin: Springer.
Fano, Vincenzo (ed.). 2014. *Gino Tarozzi Philosopher of Physics: Studies in the Philosophy of Entanglement in his 60th Birthday, Milano*. Franco: Angeli.
Geroch, R. 1978. *General Relativity from A to B*. Chicago: University of Chicago Press.
Gottesman, D., and I.L. Chuang. 1999. Quantum Teleportation is a Universal Computational Primitive. *Nature* 402: 390–93. https://doi.org/10.1038/46503.
Greenberger, D.M., and A. Yasin. 1988. Simultaneous Wave and Particle Knowledge in a Neutron Interferometer. *Physics Letters* 128A: 391–94.
Grünbaum, A. 1973. *Philosophical Problems of Space and Time, 2nd ed 1974*. Dordrecht: Reidel.
Gao, Shan (ed.). 2014. *Protective Measurement and Quantum Reality*. Cambridge: Cambridge University Press.
Healey, R. 2017. *The Quantum Revolution in Philosophy*. Oxford: Oxford University Press.
Hegerfeldt, G.C. 1974. Remarks on Causality and Particle Localization. *Physical Review D* 10: 3320–21.
Hegerfeldt, G.C. 1985. Violation of Causality in Relativistic Quantum Theory? *Physical Review Letters* 54: 2395–98.
Heisenberg, W. 1958. *Physics and Philosophy*. New York: Harper.
Heisenberg, W. 1969. *Der Teil und Das Ganze*. München: Piper.
Howard, D. 1992. Einstein and Eindeutigkeit: A Neglected Theme in the Philosophical Background to General Relativity. In *Studies in the History of General Relativity*, ed. J. Eisenstaedt, and A.J. Kox. Bioston: Birkhäuser.
Jammer, M. 1974. *The Philosophy of Quantum Mechanics. The Interpretation of Quantum Mechanics in Historical Perspective*. New York: Wiley.
Joos, E., and H.D. Zeh. 1985. The Emergence of Classical Properties Through Interaction with the Environment. *Zeitschrift für Physik* B59: 223–43.
Jacobs, B., Selinger, P., and Spitters, B. (eds.). 2012. *Proceedings 8th International Workshop on Quantum Physics and Logic (2011)*, http://eptcs.web.cse.unsw.edu.au/content.cgi?QPL2011.
Kastner, R.E. 2013. *The Transactional Interpretation of Quantum Mechanics: The Reality of Possibility*. Cambridge: Cambridge University Press.
Kim, Y.-H., R. Yu, S.P. Kulik, Y.H. Shih, and M.O. Scully. 2000. A Delayed Choice Quantum Eraser. *Physical Review Letters* 84: 1–5.
Kraus, K. 1983. *States, Effects and Operations*. Berlin: Springer.
Kuhlmann, M. 2013. What is Real? *Scientific American* 8: 32–39.
Kuhn, T.S. 1957. *The Copernican Revolution: Planetary Astronomy in the Development of Western Thought*. Cambridge: Harvard University Press.
Laughlin, R.B. 2005. *A Different Universe (Reinventing Physics from the Bottom Down)*. New York: Basic Books.
Leibniz, G.W. 1710. *Essais de Théodicée*; in [Leibniz (1978), VI, 21–436].
Leibniz, P.S. 1978. Philosophische Schriften (Ed. Gerhardt), Halle, 1875. rep. Olms: Hildesheim.
Leon-Garcia, A. 2008. *Probability, Statistics, and Random Processes for Electrical Engineering*, 3rd ed. Upper Saddle River: Pearson Prentice Hall.
Lewontin, R.C. 2000. *The Triple Helix: Gene, Organism, and Environment*. Cambridge: Harvard University Press.

Lindblad, G. 1983. *Non-Equilibrium Entropy and Irreversibility*. Dordrecht: Reidel.
Lockwood, M. 2005. *The Labyrinth of Time: Introducing the Universe*. Oxford: Oxford University Press.
Ludwig, G. 1983. *Foundations of Quantum Mechanics*. Berlin: Springer; Eng. Tr. of [Ludwig 1954].
Ma, X.-S., J. Kofler, A. Qarry, N. Tetik, T. Scheidl, R. Ursin, S. Ramelow, T. Herbst, L. Ratschbacher, A. Fedrizzi, T. Jennewein and A. Zeilinger. 2013. Quantum Erasure with Causally Disconnected Choice. *Proceedings of the National Academy of Sciences USA* 110: 1221–26. https://doi.org/10.1073/pnas.1213201110.
Malament, D.B. 1995. In Defense of Dogma: Why There Cannot Be a Relativistic Quantum Mechanics of (Localizable) Particles (Clifton (1995), 1–10).
Margenau, H. 1950. The Nature of Physical Reality: A Philosophy of Modern Physics, 1977. *New York: McGraw-Hill; Woodbridge*. Conn.: Ox Bow Press.
Martens, H., and W.M. De Muynck. 1990. Nonideal Quantum Measurements. *Foundations of Physics* 20: 255–81.
Mattle, K., H. Weinfurter, P. Kwiat, and A. Zeilinger. 1996. Dense Coding in Experimental Quantum Communication. *Physical Review Letters* 76: 4656–59.
Maudlin, T. 1994. *Quantum Non-Locality and Relativity: Metaphysical Intimations of Modern Physics*. Blackwell.
Misner, C.W., K.S. Thorne, and J.A. Wheeler. 1970. *Gravitation*. San Francisco: Freeman and Co.
Nielsen, M.A., and I.L. Chuang. 1997. Programmable Quantum Gate Arrays. *Physical Review Letters* 79: 321–24. https://doi.org/10.1103/PhysRevLett.79.321
Nielsen, M.A., and I.L. Chuang. 2000. *Quantum Computation and Quantum Information*. Cambridge: Cambridge University Press.
Ollivier, H., and W.H. Zurek. 2001. Quantum Discord: A Measure of the Quantumness of Correlations. *Physics Review Letters* 88: 017901.
Peirce, C.S. 1885. An American Plato: Review of Royce's. *Religious Aspects of Philosophy*. in [Peirce 1982, V, 221–34].
Peirce, W. 1982. *Writings, Bloomington*, vol. I. Indiana University Press.
Peirce, C.S. 1898. *Reasoning and the Logic of Things: The Cambridge Conferences Lectures of 1898*. Manuscript; ed. K. L. Ketner. Cambridge: Harvard University Press.
Peirce, C.S. 1903. Lessons from the History of Philosophy—Nominalism. in [Peirce 1958, 1.15–26].
Peirce, C.P. 1958. *The Collected Papers*, vol. I–VI (eds. Charles Hartshorne and Paul Weiss). Cambridge: Harvard University Press. 1931–1935; vol. VII–VIII (ed. Arthur W. Burks), Cambridge: Harvard University Press.
Penrose, R. 2004. *The Road to Reality: A Complete Guide to the Laws of the Universe. J*. Cape: London.
Poincaré, H. 1899. Reflexions sur le calcul des probabilités. *Revue Générale des Sciences Pures et Appliqués* 10: 262–69; Report in [Poincaré (1991), 114–34].
Poincaré, H. 1897. Le rapport de l'analyse et de la physique mathématique. *Acta Mathematica* 21: 331–41; Report in [Poincaré 1991, 17–30].
Poincaré, H. 1902. *La science et l'hypothèse*. Paris: Flammarion.
Poincaré, H. 1905. *La Valeur de la science*. Paris: Flammarion.
Poincaré, H. 1907. Le hasard. *Revue du mois* 3:257–76; Report in [Poincaré (1991), 135–56].
Poincaré, H. 1991. *L'analyse et la recherche*. Paris: Hermann.
Pusey, M.F., J. Barrett and T. Rudolph. 2012. On the Reality of the Quantum State. *Nature Physics* 8: 476–79. https://doi.org/10.1038/nphys2309.
Rindler, W. 2001. *Relativity: Special, General, and Cosmological*. Oxford: Oxford University Press.
Rovelli, C. 1996. Relational Quantum Mechanics. *International Journal of Theoretical Physics* 35: 1637–78. https://doi.org/10.1007/bf02302261.
Rovelli, C. 2004. *Quantum Gravity*. Cambridge: Cambridge University Press.
Rovelli, C. 2005. Relational Quantum Mechanics. in [Elitzur et al. 2005, 113–20].
Royer, A. 1989. Measurement of Quantum States and the Wigner Function. *Foundations of Physics* 19: 3–30.

Shapere, A., and F. Wilczek (eds.). 1989. *Geometric Phases in Physics*. Singapore: World Scientic.
Schlosshauer, M. 2007. *Decoherence and the Quantum-to-Classical Transition*. Berlin: Springer.
Schrödinger, E. 1935a. Die gegenwärtige Situation in der Quantenmechanick. I–III. *Naturwissenschaften* 23: 807–12, 823–28, 844–49; Report in [Schrödinger 1984, III, 484–501].
Schrödinger, E. 1935b. Discussion of Probability Relations Between Separated Systems. *Proceedings of the Cambridge Philosophical Society* 32: 446–52; Report in [Schrödinger 1984, I, 433–39].
Schrödinger, E. 1958. *Mind and Matter*. Cambridge: Cambridge University Press; Report in [Schrodinger 1992, 91–164].
Schrödinger, E. 1984. *Gesammelte Abhandlungen*. Wien: Verlag der österreichischen Akademie der Wissenschaften.
Schrödinger, E. 1992. *What is Life? with Mind and Matter and Autobiographical Sketches*. Cambridge: Cambridge University Press (enlarged ed. of [?]).
Scully, M.O., B.-G. Englert, and H. Walther. 1991. Quantum Optical Tests of Complementarity. *Nature* 351: 111–16.
Smart, J.J.C. 1954. The Temporal Asymmetry of the World. *Analysis* 14: 79–83.
Spinoza, B. 1972. Opera (Ed. C. Gebhardt), vol. 4. Heidelberg; Report Heidelberg, Carl Winter, 1972.
Stefanov, A., H. Zbinden, N. Gisin, and A. Suarez. 2002. Quantum Correlations with Spacelike Separated Beam Splitters in Motion: Experimental Test of Multisimultaneity. *Physical Review Letters* 88: 120404.
Storey, E.P., S.M. Tan, M.J. Collett, and D.F. Walls. 1994. Path Detection and the Uncertainty Principle. *Nature* 367: 626–28.
Suarez, A. 2001. Is There a Real Time Ordering behind the Nonlocal Correlations?. arXiv:0110124v1.
'T Hooft, G., 2016. *The Cellular Automaton Interpretation of Quantum Mechanics*. Berlin: Springer.
Verlinde, E. 2011. On the Origin of Gravity and the Laws of Newton. arXiv: 1001.0785v1.
von Bertalanffy, L. 1950. An Outline of General System Theory. *British Journal of the Philosophy of Science* 1: 139–64; Report in [von Bertalanffy 1969, 54–88].
von Bertalanffy, L. 1969. *General System Theory: Foundations, Development, Applications*. New York: G. Braziller.
Wang, L.J., X.Y. Zou, and L. Mandel. 1991. Induced Coherence Without Induced Emission. *Physical Review* A44: 4614–22.
Wheeler, J.A. 1983. Law Without Law. in [Wheeler and Zurek (1983), 182–213].
Wheeler, J.A. 1988. World as System Self–Synthesized by Quantum Networking. *IBM Journal of Research and Development* 32: 4–15. https://doi.org/10.1147/rd.321.0004.
Wheeler, J.A. 1990. Information, Physics, Quantum: The search for Links. in [Zurek 1990, 3–28].
Wheeler, J.A., and W. Zurek (eds.). 1983. *Quantum Theory and Measurement*. Princeton: Princeton University Press.
Zwolak, M., and W.H. Zurek. 2013. Complementarity of Quantum Discord and Classically Accessible Information. *Scientific Reports* 3: 1729. https://doi.org/10.1038/srep01729.

Chapter 6
Information and Signal Exchanging in Universal Network



L'hypothèse ainsi renversée a–t–elle donc été stérile?
Loin de là, on peut dire qu'elle a rendu plus de services qu'une hypothése vraie.

Henri Poincaré, *La science et l'hypothése*

With his notion of the fabric of reality (Sect. 3.1.2), D. Deutsch, extending ideas that were originally proposed by R. Feynman (Sect. 4.1) and J. Wheeler (Sect. 3.2.3), has contributed to change the way in which we consider the basic elements of our universe. According to traditional physical wisdom, the universe is composed of corpuscles (atoms or other particles), whose combination allows us to produce everything else. As mentioned, it was understood as a physics of tokens. Although successful for the first steps in our understanding of the structure of matter and chemical reactions, such an explanation appears to be less congruent when we deal with fields and less helpful for the majority of structures, patterns, exchanges, functions, and so on, occurring in our universe, for which other methodologies or approaches are used, like the principles of chaos and self-organisation, information theory, evolution theory, and so on; in fact, also in the explanation of the structure of matter that framework has been essentially modified,[1] as we shall see in the following.

The traditional ontological assumption is based on the idea that the notion of physical *object*, taken in its generality, is sufficiently robust to account for the existence of those basic elements (Sect. 3.3.4). In fact, the notion of object is not self–consistent if we consider the problem of object composition, since, abstractly speaking, properties ascribed to whole 'objects' could be incompatible with properties assigned to their parts, as nicely shown by the EPR experiment, where an entanglement of possible outcomes is not compatible with the ascription of a property to one of the two particles in absence of detection. Since it is not our aim to annihilate the notion

[1] See Auletta (2011, Sects. 6.2–6.5) for a summary.

© Springer Nature Switzerland AG 2019

G. Auletta, *The Quantum Mechanics Conundrum*,
https://doi.org/10.1007/978-3-030-16649-6_6

of ontological referent as such (Sect. 2.1), in Sect. 5.2.6 I have assumed the notion of *system* to be basic: this is a more abstract entity that maps in its state the local environmental context but also expresses those quantum-mechanical characters that we have shown to be independent of specific physical contexts. This chapter is centred on both these characters and the contexts. Having renounced the ontological notion of object, in the following, I shall use that term for indicating entities in general, included rather abstract structures like those of mathematics.

In this chapter, I shall first explore the major conceptual consequences of quantum information (the global aspect). Then, we shall deal with the fundamental notion of signal (the local aspect). Finally, the reality is understood as a network of networks dealing with information (the dynamical entrenchment of the previous two aspects).

The American philosopher Harry Collins speaks of four ways in which strings (types) influence physical entities (type–tokens) (see also Sect. 3.4.2)[2]: through physical impact, inscription of their information content in another entity, communication, which can be either mechanical (like inserting a new piece of information on a computer) or interpreted. I essentially agree but, when strings act physically or mechanically, is not though their informational nature. Thus, I propose the following variation: types can (1) restrict the space of possible messages (mutual information), while types–tokens, when some events happen, can (2) broadcast information (signal), (3) alter or disrupt other codes (noise), (4) induce some consequent action or process in other entities (communication). Note that the latter three effects go under the same category: signals and their possible effects. The latter case is typical for biological systems.

6.1 A Quantum Theory of Information

Here and in the following sections, in agreement with the results of Sect. 3.2.4, we shall deal with a notion of information that is no longer statistical, as it is the case for Shannon entropy (3.43), but with a notion that deals with *single* outcomes, having obviously a certain probability to occur. In other words, we no longer focus on the statistical mean of these possible outcomes. This fits very well with the idea that quantum systems show an amount of information that is intrinsic and self-similar (Sect. 3.2.3).

6.1.1 Information is Conserved in Our Universe

The first point of this examination is about whether information is conserved in our universe, a view that all adepts of the MWI support.[3] Indeed, if we have a self-

[2] Collins (2010, pp. 16–17). I recall also Armstrong (1978).

[3] See e.g. Deutsch (2011, Chap. 11).

replication of the same informational structure and content, the obvious consequence is that quantum information is conserved at all scales and the only variation is represented by local accessibility to it that constitutes the essence of the local swap of information that we have considered in Sect. 5.2.3. Indeed, two fundamental theorems of QM tell us that information can be neither cloned nor deleted.

No-Cloning

Let us start with the *no-cloning* theorem.[4] Suppose first that such a cloning is possible, that is, an apparatus (or whatever device) is able to take the state of an input system and copy it into another system without destroying the former. For the sake of simplicity, let us restrict ourselves to a two-level system, for example the polarisation state of a photon. Then, consider an incoming system in the state $|1\rangle$ that could represent a photon in horizontal polarisation, a second system (the one onto which we want to make a copy of the photon's state) in a generic state $|\psi\rangle$, and the apparatus being in an initial "ready" state $|A_\psi\rangle$ (corresponding to the second system's state). What we we would like to obtain is the evolution

$$|1\rangle|\psi\rangle|A_\psi\rangle \longmapsto |11\rangle\ |A_1\rangle\ . \tag{6.1a}$$

Similarly, for the same initial state of the apparatus and of the third system, but with an incoming system in a state $|0\rangle$ that could be a photon in vertical polarisation, we would like to have

$$|0\rangle|\psi\rangle|A_\psi\rangle \longmapsto |00\rangle\ |A_0\rangle\ . \tag{6.1b}$$

Now, given the assumptions (6.1a)–(6.1b), consider the case of an incoming system in an unknown superposition state $c_0|0\rangle + c_1|1\rangle$, where $c_0, c_1 \in \mathbb{C}$ satisfy the normalisation condition. Due to linearity, the result will be

$$(c_0|0\rangle + c_1|1\rangle)\,|\psi\rangle|A_\psi\rangle \longmapsto c_0\,|00\rangle\ |A_0\rangle + c_1\,|11\rangle\ |A_1\rangle\ . \tag{6.2}$$

Now, it is easy to recognise that this result is simply an ordinary coupling (where the second and third system could be detector and apparatus: Sect. 5.2.3) and not a clone of the superposition represented by the state of the first system in the LHS of Eq. (6.2). In other words, the new-born states $|0\rangle\ , |1\rangle$ could represent any kind of independent code alphabet and thus do not necessarily represent the reproduction of a specific piece of information. In fact, cloning should be rather represented by the transformation

$$(c_0|0\rangle + c_1|1\rangle)\,|\psi\rangle|A_\psi\rangle \longmapsto (c_0|0\rangle + c_1|1\rangle)\,(c_0|0\rangle + c_1|1\rangle)\,|A_?\rangle\ , \tag{6.3}$$

where the apparatus is at the end of the operation in the unknown state $|A_?\rangle$. The previous proof shows that cloning would represent a violation of the superposition principle (i.e. of the linearity of QM). Summarising, what we have proved is that it

[4] Wootters and Zurek (1982), Dieks (1982). See also Auletta et al. (2009, Sect. 15.2).

is impossible to force an arbitrary state of a system *to become the copy* of the state of another system. In other words, we cannot establish correlations between the two systems that would allow copying the state of one of them into the state of the other. This is the conceptual–formal ground of the impossibility to share code alphabets but only to couple or pair them, and as a consequence, the non-local manipulation of codes is also impossible (Sect. 3.3.7).

D'Ariano–Yuen Theorem

The proof has been generalised by the Italian physicist Mauro D'Ariano and the American computer scientist and physicist Horace Yuen[5] by showing that cloning of two non-orthogonal states would represent a violation of the unitarity of the quantum-mechanical state evolution. A quantum–cloning machine that produces $n > 1$ copies of a generic state $|\psi\rangle$ from a given set of possible states, should determine a unitary evolution of the form

$$|\psi\rangle \otimes |b_1\rangle \otimes \cdots \otimes |b_{n-1}\rangle \otimes |A\rangle \longmapsto |\psi\rangle \otimes \cdots \otimes |\psi\rangle \otimes |A'(\psi)\rangle \ , \tag{6.4}$$

in analogy with Eq. (6.3), where $|A\rangle$ represents the initial state of the apparatus or the environment (and $|A'(\psi)\rangle$ the final state), $|b_1\rangle \otimes \cdots \otimes |b_{n-1}\rangle$ are the state preparation of the modes (which are the Fourier components of the field in the case of photons) which will support the clones, and, for the sake of simplicity, normalisation has not been considered. Note that $|A\rangle$ is the initial state of sufficiently enough other modes (environment and others), so that the transformation is unitary. Now, consider the case where the states to be copied are two non-orthogonal states $|\varphi\rangle$, $|\varsigma\rangle$, i.e. such that

$$0 < |\langle\varphi|\varsigma\rangle| < 1, \tag{6.5}$$

and suppose we know *a priori* that any one of the modes is equal to the input system to be cloned, i.e for any $1 \leq k \leq n-1$, we have either

$$|\,b_k\rangle = |\,\varphi\rangle \ \text{ or } \ |\,b_k\rangle = |\,\varsigma\rangle \ , \tag{6.6}$$

depending on the input. We know that the transformation (6.4) must preserve the scalar product in order to be unitary, according to the Wigner theorem (Sect. 1.2.5). In order to test this critical point, let us write Eq. (6.4) for $|\,\varphi\rangle$,

$$|\varphi\rangle \otimes |b_1\rangle \otimes \cdots \otimes |b_{n-1}\rangle \otimes |A\rangle \longmapsto |\varphi\rangle \otimes \cdots \otimes |\varphi\rangle \otimes |A'(\varphi)\rangle \ , \tag{6.7a}$$

and for $|\,\varsigma\rangle$,

$$|\varsigma\rangle \otimes |b_1\rangle \otimes \cdots \otimes |b_{n-1}\rangle \otimes |A\rangle \longmapsto |\varsigma\rangle \otimes \cdots \otimes |\varsigma\rangle \otimes |A'(\varsigma)\rangle \ , \tag{6.7b}$$

and take the scalar product of the LHS and RHS of these two equations, so that, in analogy with Eq. (3.33), we have

[5] See D'Ariano and Yuen (1996).

$$\langle\varphi|\varsigma\rangle = (\langle\varphi|\varsigma\rangle)^n \langle A'(\varphi)|A'(\varsigma)\rangle \,, \tag{6.8}$$

where all the $\left|b_j\right\rangle$'s $(1 \leq j \leq n-1)$ cancel out due to orthonormality. From this expression it immediately follows that

$$(\langle\varphi|\varsigma\rangle)^{n-1} \langle A'(\varphi)|A'(\varsigma)\rangle = 1 \,, \tag{6.9}$$

which would in turn require that

$$|\langle A'(\varphi)|A'(\varsigma)\rangle| > 1 \tag{6.10}$$

for $n > 1$, which is clearly a violation of unitarity. So far, we have proved that the cloning of two non-orthogonal states contradicts unitarity. Since the possibility to discriminate between two non-orthogonal states would in turn imply the possibility—through quantum-mechanical unitary transformations—of generating clones of the two original states, then also the *discrimination of two non-orthogonal states* in a single step would violate the unitarity of quantum-mechanical transformations.

Broadcasting

Note that so far we have proved the pure-state no-cloning theorem. A consequence is that it prohibits broadcasting pure states, for the only way to broadcast a pure state $|\,\psi\rangle$ is to put the two systems in the product state $|\,\psi\rangle \otimes |\,\psi\rangle$, which amounts to clone $|\,\psi\rangle$. This is again typical of the way in which quantum information is 'transmitted' (Sect. 3.2.4). Things are obviously more difficult with mixed states. However, it can be proved that we cannot broadcast non-commuting mixed states.[6] Consider two mixtures

$$\hat{\tilde{\rho}} = \sum_j w_j \hat{P}_j \text{ and } \hat{\tilde{\rho}}' = \sum_k w_k \hat{P}_k. \tag{6.11}$$

Then, they do not commute iff

$$\sum_{j,k} w_j w_k \hat{P}_j \hat{P}_k \neq 0. \tag{6.12}$$

No-Deleting

Let us now deal with the impossibility of a *deleting* machine.[7] Let us denote the initial state of such hypothetical machine by $|\,A\rangle$. The aim of this machine is to delete one of the two initial copies of $|\,\psi\rangle$ and replace it with some standard state $|\,\Sigma\rangle$ of a qubit. Then, for some initial arbitrary input $|\,\psi\rangle$ we should have a transformation like

$$|\,\psi\rangle\,|\,\psi\rangle\,|\,A\rangle \longmapsto |\,\psi\rangle\,|\,\Sigma\rangle\,\left|A_\psi\right\rangle \,, \tag{6.13}$$

[6] See Barnum et al. (1996) for details.

[7] Pati and Braunstein (2000).

where $\left|A_\psi\right\rangle$ is the state of the machine after the deletion of information, which again depends on the input. Now, the most obvious solution to this equation is simply to swap (see Eq. (4.28a)) the second and third states. However, in this case we would have obtained the standard reversible erasure (Sect. 5.2.3) and not a deletion. We need to find other possibilities by keeping the general form of the transformation. Let us consider two particular instances of it:

$$|00\rangle\,|A\rangle \longmapsto |0\rangle\,|\Sigma\rangle\,|A_0\rangle \text{ and } |11\rangle\,|A\rangle \longmapsto |1\rangle\,|\Sigma\rangle\,|A_1\rangle\,, \tag{6.14}$$

which, to a certain extent, could be considered the inverses of Eqs. (6.1). However, this should be true for various classes of inputs, also when they are different or entangled. For instance, we should have

$$\left|\Psi^+\right\rangle|A\rangle = \frac{1}{\sqrt{2}}\left(|01\rangle + |10\rangle\right)|A\rangle \longmapsto |\Xi\rangle\,, \tag{6.15}$$

where $|\Xi\rangle$ here is the resulting appropriate state of the three involved systems, and I have used Eq. (3.164a). Suppose now that each input is the qubit

$$|\psi\rangle = c_0\,|0\rangle + c_1\,|1\rangle\,, \tag{6.16}$$

where $c_0, c_1 \in \mathbb{R}$ for the sake of simplicity and satisfy normalisation of probabilities. Now, keeping in mind the linearity of transformations (6.14) and (6.15), we must have

$$\begin{aligned} |\psi\rangle\,|\psi\rangle\,|A\rangle &= \left[c_0^2\,|00\rangle + c_1^2\,|11\rangle + c_0c_1\left(|01\rangle + |10\rangle\right)\right]|A\rangle \\ &\longmapsto c_0^2\,|0\rangle\,|\Sigma\rangle\,|A_0\rangle + c_1^2\,|1\rangle\,|\Sigma\rangle\,|A_1\rangle + \sqrt{2}c_0c_1\,|\Xi\rangle\,. \end{aligned} \tag{6.17}$$

However, in order to satisfy the general formula (6.13), the final state should reduce to

$$|\psi\rangle\,|\Sigma\rangle\left|A_\psi\right\rangle = \left(c_0\,|0\rangle + c_1\,|1\rangle\right)|\Sigma\rangle\left|A_\psi\right\rangle, \tag{6.18}$$

which, to a certain extent, could be considered the inverse transformation relative to (6.3). The only possible solution is represented by

$$|\Xi\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle\,|\Sigma\rangle\,|A_1\rangle + |1\rangle\,|\Sigma\rangle\,|A_0\rangle\right). \tag{6.19}$$

In fact, substitution of the latter equation into the last row of Eq. (6.17) gives

$$c_0\,|\psi\rangle\,|\Sigma\rangle\,|A_0\rangle + c_1\,|\psi\rangle\,|\Sigma\rangle\,|A_1\rangle\,. \tag{6.20}$$

Thus, the linearity of the transformation is sufficient condition of

$$\left| A_\psi \right\rangle = c_0 \left| A_0 \right\rangle + c_1 \left| A_1 \right\rangle . \tag{6.21}$$

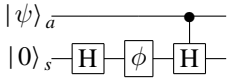
This shows that, also in the case of deletion, we have finally only swapped one 2D input onto the 2D subspace of the apparatus, and therefore we have performed a simple reversible erasure of the kind discussed in Sect. 5.2.3. As the authors say: we can only "move the information around without deleting it".

Information is Conserved

The unavoidable conclusion is that the whole information and therefore also the whole (informational) entropy of our universe is *conserved* if our universe does not interact (or very weekly interacts) with other possible universes. We can unit these two no-go theorems (no-cloning and no-deleting information) in a single theorem formulated in a positive way: quantum information is everywhere conserved. The problem is only its local accessibility. This is what *information selection* means in such a context: quantum systems are sensitive to different physical contexts (Sect. 5.2.6), so that they share, swap and locally loose information. As mentioned, this is a quite natural consequences of their fundamental self-similar nature. Moreover, also the reciprocal influence happens: quantum systems, with their sole presence, influence the environment. To throw light on this issue, I shall consider here an experiment, of the genus of the delayed-choice ones (Sect. 2.4.3), showing a superposition between "presence" and "absence" of an object.[8] Consider an ancilla photon in the state

$$\left| \psi \right\rangle_a = \cos\alpha \left| 0 \right\rangle_a + \sin\alpha \left| 1 \right\rangle_a , \tag{6.22}$$

where α is a complex parameter, and the circuit



(6.23)

This could represent an interferometer, where, however, the second beam splitter is a quantum-mechanical BS (represented by a controlled-Hadamard operation), which can be set in a superposition of *present* and *absent*, according to the state of an ancilla photon. This allows intermediate quantum behaviour of the signal photon s to be observed, with continuous transformation between particle and wave behaviour. The photon s that is in an initial state $\left| 0 \right\rangle$ after the first (ordinary or classical) 50–50 beam splitter and a phase shifter in the upper path is in the state (see the RHS of Eq. (2.28))

$$\left| \psi \right\rangle_s = \frac{1}{\sqrt{2}} \left(\left| 0 \right\rangle + e^{i\phi} \left| 1 \right\rangle \right)_s . \tag{6.24}$$

Now, the presence of the second beam splitter depends on the state of the ancillary photon. If the ancillary photon is prepared in the state $\left| 0 \right\rangle$, no BS is present, and therefore the interferometer is left open at the end. In this case, it is like let the

[8] Ionicioiu and Terno (2011), Peruzzo et al. (2012). See also Afriat and Tarozzi (2006).

identity operator acting on the state (6.24), and we can know the path taken by the signal photon. If, however, the ancillary photon is prepared in the state $|1\rangle$, the BS is present and the interferometer is therefore closed. Formally, this means to apply a Hadamard transformation to the signal photon resulting in the state (see Eq. (2.31))

$$\left|\psi'\right\rangle_s = \frac{1}{2}\left[(1+e^{i\phi})\,|0\rangle_s + (1-e^{i\phi})\,|1\rangle_s\right]. \tag{6.25}$$

The final measurement will give no information about about the path taken by the photon. Now, assuming that the ancilla is prepared in the superposition state (6.22), the global state of the system evolves into

$$|\Psi\rangle_{sa} = \cos\alpha\,|\psi\rangle_s\,|0\rangle_a + \sin\alpha\left|\psi'\right\rangle_s|1\rangle_a\,, \tag{6.26}$$

which is entangled whenever $0 < \alpha < \pi/2$. We have thus established a superposition of "BS absent" and "BS present". In other words, we make a piece of the environment depend on the state of a quantum system.

6.1.2 *No-programming Theorem*

No-Programming

The theorem tells us that it is impossible to build a quantum device that accepts two input states, namely a data state and a program state, and, depending on the data, performs on the latter unitary transformations exactly.[9] This is classically possible due to von Neumann's paradigm (after the name of John von Neumann) of *program as data*, i.e. of programs as coded binarily as any chunk of data. As a matter of fact, universal quantum computation can be achieved only approximately, and approximation improves as the dimensions of the program increase.

To prove the theorem, suppose that the initial state of the system is

$$|D\rangle \otimes |P\rangle\,, \tag{6.27}$$

where $|D\rangle$ is a state of the m-qubit data register, and $|P\rangle$ is a state of the n-qubit program register, where it is noted that the initial state is a product state. We assume that the dynamics of the whole programmable gate array is given by a unitary operator, $\hat{U}_G$, so that

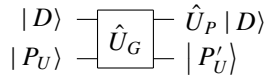
$$|D\rangle \otimes |P\rangle \longmapsto \hat{U}_G\,[\,|D\rangle \otimes |P\rangle\,]\,. \tag{6.28}$$

A unitary operator, $\hat{U}_P$, acting on the m data qubits, is said to be implemented by this gate array if there exists a state $|P_U\rangle$ of the program register such that

[9] Nielsen and Chuang (1997).

$$\hat{U}_G\,[\,|\,D\rangle\,\otimes\,|\,P_U\rangle\,]=\left(\hat{U}_P\,|\,D\rangle\right)\otimes\left|\,P'_U\right\rangle \tag{6.29}$$

for all states $|\,D\rangle$ of the data register and for some output state $\left|\,P'_U\right\rangle$ of the program register (meaning, e.g. a "stop"). This can be summarised in the circuit



(6.30)

It is quite natural to assume that $\left|\,P'_U\right\rangle$ depended on $|\,D\rangle$. To show that this is not possible, consider the following transformations on the two data registers $|\,D_1\rangle$ and $|\,D_2\rangle$:

$$\hat{U}_G\,[\,|\,D_1\rangle\,\otimes\,|\,P_1\rangle\,]=\left(\hat{U}_P\,|\,D_1\rangle\right)\otimes\left|\,P'_1\right\rangle\,, \tag{6.31a}$$

$$\hat{U}_G\,[\,|\,D_2\rangle\,\otimes\,|\,P_2\rangle\,]=\left(\hat{U}_P\,|\,D_2\rangle\right)\otimes\left|\,P'_2\right\rangle\,. \tag{6.31b}$$

Taking the inner product side by side of these two equations (see also Eq. (3.33))

$$\langle D_1\,|\,D_2\rangle=\langle D_1\,|\,D_2\rangle\left\langle P'_1\,|\,P'_2\right\rangle, \tag{6.32}$$

we see that, whenever $\langle D_1\,|\,D_2\rangle\neq 0$, we have $\left\langle P'_1\,|\,P'_2\right\rangle=1$, what in turn implies that $\left|\,P'_1\right\rangle=\left|\,P'_2\right\rangle$, that is, the state of the program does not depend on the state of the data. If $\langle D_1\,|\,D_2\rangle=0$ (the two data sets are orthogonal), then the scalar product $\left\langle P'_1\,|\,P'_2\right\rangle$ can be arbitrary, what shows again that there is no dependence on such data. In other words, in any case there is no constraint by the data upon the programs. This is further evidence of the impossibility to share codes (Sect. 3.3.7).

No Deterministic Universal Quantum Gate Array

The previous result depends on the fact that the program does not evolve deterministically in a classical sense (Sects. 2.4.1–2.4.2 and 3.4.2). The set of unitary operators on m qubits can be parametrised by 2^{2m} independent real numbers, which is fewer than the $2^{2m+1}-1$ real numbers needed to parametrise a set of $2m$ qubits. This seems to imply that it is possible to build a universal quantum gate array able to implement any unitary operation. Universal gate arrays are certainly possible for classical computers, since by counting the number of possible functions we see that an arbitrary function on m bits can be specified using $m2^m$ bits, and it is straightforward to design a classical circuit which will take $m2^m$ program bits as input and implement the corresponding function on m data bits. However, the authors of the study show that every implementable unitary operation requires an extra Hilbert space dimension in the program register.[10] Since the number of possible unitary operations on m qubits is infinite, it follows that a universal gate array would require an infinite number of qubits in the program register, and thus no such array exists.

[10] Nielsen and Chuang (1997).

Suppose distinct (up to a global phase) unitary operators $\hat{U}_1, \ldots, \hat{U}_n$ are implemented by some programmable quantum gate array. Then, there are two fundamental consequences: (i) the program register is at least n dimensional, that is, contains at least $\lg n$ qubits; (ii) the corresponding programs $|P_1\rangle, \ldots, |P_n\rangle$ are mutually orthogonal, in order that the dimension of the Hilbert space of the program register be $= n$. It suffices to prove these results for 2D, by supposing that $|P\rangle$ and $|Q\rangle$ are two programs which implement unitary operators $\hat{U}_P$ and $\hat{U}_Q$, respectively, which are distinct up to global phase changes. Then, for arbitrary data $|D\rangle$, in agreement with Eq. (6.31), we have

$$\hat{U}_G\left[|D\rangle \otimes |P\rangle\right] = \left(\hat{U}_P |D\rangle\right) \otimes \left|P'\right\rangle, \tag{6.33a}$$

$$\hat{U}_G\left[|D\rangle \otimes |Q\rangle\right] = \left(\hat{U}_Q |D\rangle\right) \otimes \left|Q'\right\rangle. \tag{6.33b}$$

Taking again the scalar product of the two equations side by side, we get

$$\langle Q \mid P\rangle = \left\langle Q' \mid P'\right\rangle \left\langle D \left| \hat{U}_Q^\dagger \hat{U}_P \right| D\right\rangle. \tag{6.34}$$

Suppose that we have $\left\langle Q' \mid P'\right\rangle \neq 0$. This allows us to write

$$\frac{\langle Q \mid P\rangle}{\left\langle Q' \mid P'\right\rangle} = \left\langle D \left| \hat{U}_Q^\dagger \hat{U}_P \right| D\right\rangle. \tag{6.35}$$

Since the LHS has no dependence on $|D\rangle$, we have

$$\hat{U}_Q^\dagger \hat{U}_P = \alpha \hat{I}, \tag{6.36}$$

where $\alpha = \langle Q \mid P\rangle / \left\langle Q' \mid P'\right\rangle$, and thus $\alpha \in \mathbb{R}$. In other words, we have $\left\langle Q' \mid P'\right\rangle \neq 0$ when $\hat{U}_P$ and $\hat{U}_Q$ are the same up to a global phase factor. But we have assumed that this is not the case, and thus it follows that $\left\langle Q' \mid P'\right\rangle = 0$. Then, from Eq. (6.34) it immediately follows that also

$$\langle Q \mid P\rangle = 0. \tag{6.37}$$

As a consequence, the programs are orthogonal, and, therefore no deterministic universal quantum gate array exists.

Probabilistic Universal Quantum Gate Array

However, we can implement a universal quantum gate array in a probabilistic way, according to the probabilistic evolution of quantum systems, where I recall that quantum determinism is not of classical nature. Let us show this for a single qubit ($m = 1$). Suppose that the program input is the two-qubit Bell state (3.164c) and we have

$$|\,P_U\rangle = \left(\hat{I}\otimes\hat{U}\right)|\,\Phi^+\rangle = \frac{1}{\sqrt{2}}\left(|\,0\rangle\ (\hat{U}\,|\,0\rangle\,) + |\,1\rangle\ (\hat{U}\,|\,1\rangle\,)\right), \tag{6.38}$$

for some unitary operation $\hat{U}$. Assume now that the data input is

$$|\,D\rangle = c_0\,|\,0\rangle + c_1\,|\,1\rangle\ , \tag{6.39}$$

with $c_0, c_1 \in \mathbb{R}$ satisfying the normalisation condition. Then, taking advantage of Eqs. (4.100), we have for the whole data+program input

$$\begin{aligned}
|\,D\rangle\otimes|\,P_U\rangle &= \frac{1}{2}\left\{c_0\left[\left(|\,\Phi^+\rangle + |\,\Phi^-\rangle\right)\hat{U}\,|\,0\rangle + \left(|\,\Psi^+\rangle + |\,\Psi^-\rangle\right)\hat{U}\,|\,1\rangle\right]\right. \\
&\quad \left. + c_1\left[\left(|\,\Psi^+\rangle - |\,\Psi^-\rangle\right)\hat{U}\,|\,0\rangle + \left(|\,\Phi^+\rangle - |\,\Phi^-\rangle\right)\hat{U}\,|\,1\rangle\right]\right\} \\
&= \frac{1}{2}\left[|\,\Phi^+\rangle\left(c_0\hat{U}\,|\,0\rangle + c_1\hat{U}\,|\,1\rangle\right) + |\,\Phi^-\rangle\left(c_0\hat{U}\,|\,0\rangle - c_1\hat{U}\,|\,1\rangle\right)\right. \\
&\quad \left. + |\,\Psi^+\rangle\left(c_0\hat{U}\,|\,1\rangle + c_1\hat{U}\,|\,0\rangle\right) - |\,\Psi^-\rangle\left(c_0\hat{U}\,|\,1\rangle - c_1\hat{U}\,|\,0\rangle\right)\right] \\
&= \frac{1}{2}\left[|\,\Phi^+\rangle\left(\hat{U}\hat{I}\,|\,D\rangle\right) + |\,\Phi^-\rangle\left(\hat{U}\mathrm{Z}\,|\,D\rangle\right) + |\,\Psi^+\rangle\left(\hat{U}\mathrm{X}\,|\,D\rangle\right) - |\,\Psi^-\rangle\left(\hat{U}i\mathrm{Y}\,|\,D\rangle\right)\right],
\end{aligned} \tag{6.40}$$

in analogy with the teleportation case (Sect. 3.3.6). Indeed, we can recover $\hat{U}\,|\,D\rangle$. However, since the success of this recovering depends on the result of a Bell measurement performed on the data input and the first qubit of the program input (always in analogy with the teleportation case), already in the $m = 1$ case, as predicted, the gate array is clearly nondeterministic, and succeeds (depending on the measurement outcome) with probability 1/4.

6.1.3 No-Bit Commitment

Bit–Commitment Protocol

Suppose that, in a cryptography protocol (Sect. 3.3.6), Alice has a bit in mind to send and she wishes to provide Bob with a piece of evidence that (i) she has already chosen the bit to send and (ii) cannot change it after such a choice. Meanwhile, Bob should not be able to tell from that evidence what this bit is. At a later time, however, it must be possible for Alice to open the commitment, that is, Alice must be able to show Bob which bit she has committed to and convince him that this is indeed the genuine bit that she had in mind when she committed. This is the bit-commitment protocol, which, therefore, units two different requirements[11]:

- It must be *binding*, meaning that Alice could not cheat about the value of the bit for which she is committed to Bob, and
- It must be *concealing*, meaning that Bob cannot have access to such a bit before Alice decides to disclose it.

[11] D'Ariano et al. (2017, Sect. 9.3).

No QBC: The Conceptual Scheme

Now, it can be shown that a dishonest party can exploit the non-local EPR type correlations in QM to cheat successfully.[12] To do so, she generally needs to maintain the coherence of her shared quantum system by using a quantum computer (Chap. 4). What is meant by cheating? As an example, Alice may choose a particular value of the bit during the commitment phase and tell Bob another value during the opening phase. A quantum bit commitment (QBC) scheme is secure against Alice's cheating only if such a fake commitment can be discovered by Bob.

Let us consider the BB84 protocol reported in Sect. 3.3.6 and use it for QBC. In this case, Alice and Bob first agree on a security parameter, a positive integer s. The sender, Alice, chooses the value (0 or 1) of the committed bit, B. We then have two possibilities:

- If $B = 0$, she prepares and sends Bob a sequence of s photons each of which is randomly chosen to be either horizontally or vertically polarized. Of course, the value of B is kept secret during the commitment phase. Moreover, the actual polarisation of each photon chosen by Alice is not announced to Bob.
- If $B = 1$, she prepares and sends Bob a sequence of s photons each of which is randomly chosen to be one of the diagonal polarisation basis (see Eqs. (3.176)) while once again both the value of B and the actual polarisation of each photon are kept secret by Alice.

Of course, Bob chooses randomly between the rectilinear and diagonal bases. A calculation can show that the two density matrices describing the s photons corresponding to $B = 0$ and $B = 1$, respectively, are exactly the same (and are proportional to the identity matrix). Consequently, Bob cannot learn anything about the value of B.

At a later moment, Alice discloses her commitment to Bob. Since Bob has chosen his basis (rectilinear or diagonal) of measurement randomly for each photon in the commitment phase, on average, only half of the s photons have been measured by him in the correct basis. For those photons, Bob can verify that Alice's announced polarisations match his measurement results. However, a cheating Alice may, for example, send rectilinear photons in the commitment phase (hence commits to $B = 0$) but tell Bob that they are diagonal photons in the opening phase (hence announces $B = 1$). Alice then has to make a random guess for the polarisations of the photons that Bob has measured along the diagonal basis. Since Bob, on average, measures $s/2$ photons along the diagonal basis, Alice, with such a cheating strategy, has apparently only a probability of $(1/2)^{\frac{s}{2}}$ for success.

No QBC: Quantum-Mechanical Formalism

Let us see what happens in quantum-mechanical terms. According to whether Alice chooses 0 or 1 as the value of B, she prepares the two registers (the two entangled particles) in one of the two following states:

[12] Lo and Chau (1997), Mayers (1997). See also Clifton et al. (2003).

$$|0\rangle = \sum_j \alpha_j \left|a_j\right\rangle_A \otimes \left|b_j\right\rangle_B , \tag{6.41a}$$

$$|1\rangle = \sum_k \beta_k \left|a'_k\right\rangle_A \otimes \left|b'_k\right\rangle_B , \tag{6.41b}$$

where Alice's states (the $\left|a_j\right\rangle_A$'s) are orthogonal but those of Bob (the $\left|b_j\right\rangle_B$'s) not necessarily so. Both partners know the states $|0\rangle , |1\rangle$, what implies that they also know the states of the two registers in both the above cases. Supposing that Alice be honest, she measures her register and determines either the value of j if $B = 0$ or the value of k if $B = 1$, and sends the second register (e.g. the qubits in states $\left|b_j\right\rangle_B$'s) to Bob as evidence of her commitment (the two registers are in fact correlated). Finally, Bob performs his own measurement and, thanks to the correlation between the piece of information provided by Alice and his experimental result, should tell whether Alice has cheated or not.

Up to now we have followed a canonical, for example cryptography, protocol. However, a key weakness of Bennett and Brassard's scheme (and similar ones) is that Alice can always cheat successfully by using EPR-pairs. Alice can prepare s EPR-pairs of photons and send a member of each pair to Bob during the commitment phase. She *skips* her measurements and decides on the value of B only at the beginning of the opening phase, that is, she decides to perform (a delayed) measurement (Sect. 5.2.2). As mentioned, to do so, Alice generally needs a quantum computer.

In order to have a successful cheat, Bob's register should contain little information about the state of Alice's register. Ideally, supposing that it contains no information at all, as announced, the two reduced density matrices of Bob's register should be the same for both cases:

$$\mathrm{Tr}_A \, |0\rangle \langle 0| = \hat{\varrho}_0^B = \mathrm{Tr}_A \, |1\rangle \langle 1| = \hat{\varrho}_1^B . \tag{6.42}$$

Making use of the Schmidt decomposition (1.398), we can rewrite Eqs. (6.41) as

$$|0\rangle = \sum_n \sqrt{w_n} \, |A_n\rangle \otimes |B_n\rangle , \tag{6.43a}$$

$$|1\rangle = \sum_n \sqrt{w_n} \left|A'_n\right\rangle \otimes |B_n\rangle , \tag{6.43b}$$

where I have dropped the subscripts A, B for the sake of the notation; the w_n are the eigenvalues of the reduced density operators (6.42), and the states of both registers represent now two orthogonal bases for the respective Hilbert spaces. Note that the only difference between the two previous equations is represented by the states of Alice's register. Consider now a unitary transformation $\hat{U}_A$ acting on Alice's register only and mapping $|A_n\rangle$ to $\left|A'_n\right\rangle$, what, for the above reason, represents also a map from $|0\rangle$ to $|1\rangle$. Since this can be done by Alice without any awareness by Bob, she can cheat about the value of B in the opening phase of the protocol.

Thus, there is no way to guess that Alice is cheating. It has been proved that a cheating strategy works also in the non-ideal case with a large probability. In conclusion, the previous proof shows that there do not exist quantum protocols that are both binding and concealing. This implies that there is no way to ensure once and for all a certain specific coupling between Alice's states and Bob's states, which is again another evidence that code alphabets cannot be shared as well as they are freely and locally selected (Sect. 3.3.7).

6.1.4 Information Accessibility

Accessibility and Robustness

The different environmental contexts in which a system is determine what is the information that we can have access to (Sect. 5.2.6). When measuring a quantum system, we can locally get the final classical information. Nevertheless, I have also remarked that we can also measure, although in an indirect sense, the quantum features of that system (Sects. 5.1.2–5.1.4). As stated by the physicists D. Deutsch and P. Hayden,[13] *locally inaccessible information* can be defined as information which is present in a system but does not affect the probability of any outcome of any possible measurement on that system alone. Therefore, we can say that *locally* we can only have access to the classical part of the information represented by a quantum system, but, when, for example *comparing* two (series of) local measurements on the two (sets of identical) particles constituting (each couple) an entangled state we can also have access to the quantum part. In fact, we have already shown that entanglement does not alter the local probability distributions but modifies the space of the possible joint events making some outcomes impossible.

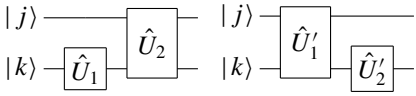
Local Observables?

On this basis, the issue raised by D. Deutsch and P. Hayden is whether the evolution of observables can be described as a local process. Thus, they make use of the Heisenberg formalism for dealing with this kind of evolution. They note that, if we do not consider the evolution of the state *and* consider observables confined to a single subsystem (therefore avoiding non-local observables like the Schrödinger's ones (3.102)), these observables are local in the sense understood by EPR, namely of the independence of the operations performed on one system relative to those performed on the other system (Sect. 2.3.1). This is in agreement with our analysis that observables need to be taken as elements of reality provided that we are able to exhibit experimental contexts (Sect. 3.3.2). In other words, "the descriptor at time t of a given system will depend, apart from the history of operations applied to it alone, only on its previous interactions and on the histories and past interactions of the systems it has interacted with. This property has been called contiguity."[14] In

[13] Deutsch and Hayden (2000).

[14] Timpson (2013, p. 105).

particular, consider the two following circuits:



(6.44)

They represent local operations according to Deutsch and Hayden. On the left, we have that the subsequent evolution of qubit j does depend both on the unitary operator $\hat{U}_2$ and on the unitary operation $\hat{U}_1$ performed at a previous time on k; on the right, the subsequent evolution of qubit j is unaffected by the unitary operation $\hat{U}'_2$ on qubit k. The conclusion of the philosopher C. Timpson, who has carefully analysed this argument, is that it is nothing else than the no-signalling requirement (Sect. 3.3.7). In fact, the different protocols discussed so far, like superdense coding, teleportation, entanglement swapping (Sects. 3.3.6 and 5.2.2), and so on, show that, in the presence of entanglement, local unitary operations can have a very significant effect on the global state of the system.[15]

At the opposite, Deutsch and Hayden aim to show that there can be no effect on the global state from local unitary operations, even in the presence of entanglement; and this indeed follows, in the trivial sense, if we fix a *particular* initial state and track the time evolution via the quantum gates, adopting the Heisenberg picture (Sect. 4.4), so that that state remains unchanged during this time evolution (Sect. 2.1.1). Deutsch and Hayden suppose in fact that the initial state is always $|00\ldots0\rangle$, according to the number of qubits involved. But what we described in the Schrödinger picture as a change in the global state following a local operation now merely becomes, in the Heisenberg picture, a change in the expectation values precisely for some *joint* observables that cannot be understood in terms of changes in the expectation values for observables pertaining to single subsystems. This point is also implicitly acknowledged by Deutsch and Hayden when they say that the information contained in local observables does "not contain a copy of all the information" contained in correlated qubits. This, however, shows that non-locality is a genuine irreducible concept as far as we take into consideration the state of a quantum system.

As pointed out by Timpson, a stronger (ontological) interpretation of the previous proposal would be that the intrinsic properties of a subsystem, denoted by certain descriptors like the unitary operators introduced in circuits (6.44), which could be Pauli matrices, be fundamental. This means that there would be an answerable question about which properties a given system actually possesses at any stage; and thus also, another answerable question about what the true descriptor of the system is. However, that interpretation raises the problem of which properties are in fact observable: "The observable properties are those that are given by expectation values. But this means that we can never in fact know the true descriptor of a system. We only have empirical access to expectation values and to the density matrices of systems, but continuously many different gates will be compatible with this data. The true descriptor of a system could be any one of the many that would provide consistency

[15] See also Timpson (2013, Sect. 5.3).

with both the density matrix of the subsystem and that of the total system. Thus the facts about the true descriptors, and hence about the intrinsic properties that systems actually possess, although supposedly the fundamental reality, are empirically inaccessible."[16]

It is not by chance that Deutsch has developed a kind of local, EPR, dynamics inside *a single world* while acknowledging *non-local correlations among worlds*. This has to do with the characteristic under-evaluation of the role of correlation inside a single world to avoid Leibniz's problem with possibilities, as already pointed out in Sect. 3.1.2.

6.1.5 Final Considerations

The Theorems Resumed

What we have considered up to now are several theorems about quantum information. This is what we can call the *global aspect* of information. In fact, all the discussed theorems deal ultimately with correlations. The no-cloning theorem tells us that we cannot unitarily evolve quantum correlations in order to get a copy of the state of one system into the state of another system. In fact, the global level of correlations does not allow local changes (Sect. 5.1.4). Since cloning is logically equivalent to the possibility of individual state determination out of correlations, the consequence was superluminal signalling (Sect. 5.1.5). The impossibility to clone avoids, therefore, any conflict with relativity. The no-deleting theorem focusses on the inverse situation but whose general terms remain the same: quantum correlations only allow for information swapping from one system to another but not for using one for deleting the information contained in the state of another. It may be noted that information is conserved at any scale and for any system or subsystem provided that it can be well described by quantum theory.

The no-programming theorem tells us something that is close to the two previous theorems but perhaps more general: we cannot deterministically get a state of a system depending on the state of another system. At most, we get results according to the probabilistic laws of QM. Finally, the no- bit commitment tells us that correlations cannot hinder local manipulations of information when some kind of measurement is involved. Such a manipulation represent the local independence in the choice of code alphabets, so that the latter cannot be quantum-mechanically transferred but only coupled (Sects. 3.3.6–3.3.7). The general conclusion is that the global *does not determine* the local, while reciprocally we have *no access* to the global starting from the local.

[16]Timpson (2013, pp. 111–112). However, I do not share the negative conclusion of the author about information accessibility.

Information and Entropy

This examination allows us to make more precise what happens during local decoherence processes: there are both a local *shift in entropy* and a non-local *swap of information*, which helps us to clarify the relation between these two concepts (Sects. 3.2.2–3.2.4).

Information is a formal quantity. As all kinds of constraints, it consists in the arrangement of elements and parts, in their combinatorics (Sect. 3.4.2). In particular, as stressed, is a type and not a token. However, it is such typical arrangement of elements that can exert a (non-dynamical) causal influence. Obviously, we humans target such an effect when we control experiments or produce strings of bits that tell us something about something else, but we cannot assume that this also happens in our external physical world. Nevertheless, we have reached the conclusion that the formal nature of information is such that this causal-constraining influence spontaneously arises when there are appropriate conditions and physical processes. In particular, the spontaneous establishment of correlations among quantum systems, although still reversible (Sect. 5.2.3), determines these different conditions of accessibility. Thus, this informational structure has in turn consequences on the structure of reality.

While information is *intrinsic* to the system and is the only quantity to be intrinsic to the system, *entropy* deals not only with the relation among the components of the system (what is also true for information) but also of a system with other systems. This is why for a quantum system entropy is typically zero. In other words, entropy, although tightly connected with the notion of information due to the relevance of correlations for both concepts, is by definition a *relative* concept, while information is monadic. This is why entropy (especially in its form of Boltzmann entropy) is dynamic while information is only formal. In fact, in order to have a dynamics you need to have a change of a system relative to other systems or a background. To understand this, let us recall the case (mentioned in Sect. 3.2.3) of a quantum system trapped in a potential well that is discontinuously turned off: in that case the system could be everywhere in the universe, which is something that happens without propagation of a signal at all. This is not a dynamical process in the strict sense of the word.

In fact, entropy means ultimately (mix of) order and disorder: by rewriting the second of Eqs. (5.12) as

$$S(Y) = I(X : Y) + S(Y|X), \tag{6.45}$$

we see that the entropy of a system Y is a mix of mutual information and conditional entropy relative to another system X. This can be generalised when the other system is an environment. Now, order and disorder can be displaced. This is well known: in the simplest case, this can be done with energy expenditure, like with the Maxwell's demon. (Sect. 3.2.4).[17] So, when we are downloading quantum features into the environment, we are in fact displacing entropies.

[17] Landauer (1961), Bennett (1982).

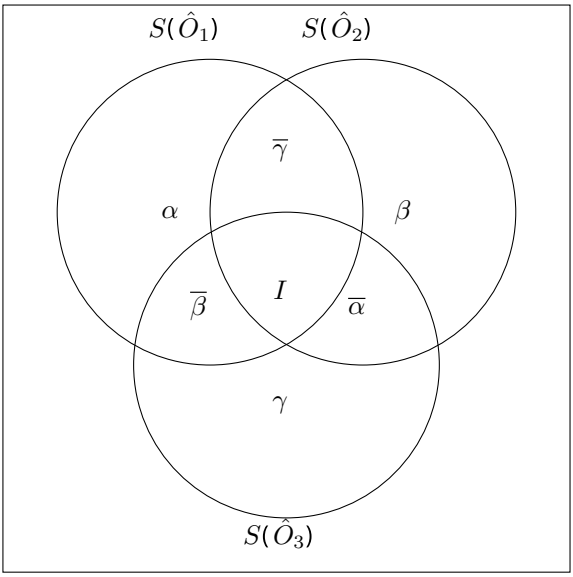


Fig. 6.1 Conditional entropies represented by means of Venn diagrams

Negative Conditional Entropy

We have seen that the Araki–Lieb inequality (3.76) formally guarantees us that there can be local positive entropy of the subsystems (provided that it is equal) even when the global system has zero entropy. Thus, if the world conserves the same amount of entropy (and information), we may wonder whether the local growth of entropy is compensated elsewhere, in other pieces of the environment, by a form of negative entropy in order to always keep a global zero value (we have already mentioned this in Sect. 3.3.6). This is precisely the case for quantum systems. To show this, let us consider the von Neumann entropy of observables (Sect. 3.2.2) and generalise the notion of mutual information (see the right equation of (5.12)) to three observables $\hat{O}_1, \hat{O}_2, \hat{O}_3$[18]:

$$I(\hat{O}_1 : \hat{O}_2 : \hat{O}_3) = I(\hat{O}_1 : \hat{O}_2) - I(\hat{O}_1 : \hat{O}_2|\hat{O}_3). \tag{6.46}$$

Note that using observables means to pick out their own eigenbases, circumventing in this way the problem raised in Sect. 5.1.3. Again, we have evidence that observables represent a higher degree of determination than states. Let us have a look at Fig. 6.1 and denote the three conditional entropies as follows:

$$\alpha := S(\hat{O}_1|\hat{O}_2, \hat{O}_3),\ \beta := S(\hat{O}_2|\hat{O}_1, \hat{O}_3),\ \gamma := S(\hat{O}_3|\hat{O}_1, \hat{O}_2), \tag{6.47}$$

and the three conditional mutual informations as follows:

[18] Cerf and Adami (1997).

$$\overline{\alpha} := I(\hat{O}_2 : \hat{O}_3|\hat{O}_1),\ \overline{\beta} := I(\hat{O}_1 : \hat{O}_3|\hat{O}_2),\ \overline{\gamma} := I(\hat{O}_1 : \hat{O}_2|\hat{O}_3); \quad (6.48)$$

the overall mutual information (6.46) is defined as $I := I(\hat{O}_1 : \hat{O}_2 : \hat{O}_3)$. For a classical system all entropies are nonnegative (while both the classical and quantum mutual information need to be always ≥ 0, according to Eq. (3.48)). Thus, we can write

$$0 \leq \alpha + \overline{\alpha} = S(\hat{O}_1|\hat{O}_2, \hat{O}_3) + I(\hat{O}_2 : \hat{O}_3|\hat{O}_1), \quad (6.49a)$$
$$0 \leq \beta + \overline{\beta} = S(\hat{O}_2|\hat{O}_1, \hat{O}_3) + I(\hat{O}_1 : \hat{O}_3|\hat{O}_2), \quad (6.49b)$$
$$0 \leq \gamma + \overline{\gamma} = S(\hat{O}_3|\hat{O}_1, \hat{O}_2) + I(\hat{O}_1 : \hat{O}_2|\hat{O}_3). \quad (6.49c)$$

It can be shown that these inequalities are in fact Bell inequalities (Sect. 3.3.5), although formulated in informational-entropic terms.[19] Now, a violation of one of inequalities (6.49) means that one of the sums of type $\alpha + \overline{\alpha}$ is less then zero. But since $\overline{\alpha}, \overline{\beta}, \overline{\gamma}$, being (conditional) mutual informations, are always ≥ 0, it follows that at least one of the conditional entropies α, β, γ must be negative. This is quite understandable if we consider that, informationally speaking, conditional entropies express equivocation, as mentioned in Sect. 5.1.3. Indeed, quantum features guarantee a level of order that is classically unknown. This result is strictly connected with the presence of the quantum discord (5.19). In other words, quantum features, are responsible also for the negative values of the conditional entropies computed on observables. Thus, it appears meaningful to rewrite the previous inequalities for systems $\mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_3$ as

$$0 \leq \alpha + \overline{\alpha} = S(\mathcal{S}_1|\{\hat{P}_j\}_{23}) + I(\mathcal{S}_2 : \mathcal{S}_3|\{\hat{P}_k\}_1), \quad (6.50a)$$
$$0 \leq \beta + \overline{\beta} = S(\mathcal{S}_2|\{\hat{P}_j\}_{13}) + I(\mathcal{S}_1 : \mathcal{S}_3|\{\hat{P}_k\}_2), \quad (6.50b)$$
$$0 \leq \gamma + \overline{\gamma} = S(\mathcal{S}_3|\{\hat{P}_j\}_{12}) + I(\mathcal{S}_1 : \mathcal{S}_2|\{\hat{P}_k\}_3), \quad (6.50c)$$

where the projectors could also project on the eigenbases of the previous observables (which in such a case would pertain each to one of the three subsystems).

Information Distance

A similar problem concerns the information distance (3.50). It can be proved that $\forall X, Y, Z$, the informational distance satisfies the following triangular inequality

$$\delta(X, Y) + \delta(Y, Z) \geq \delta(X, Z)\ . \quad (6.51)$$

This allows us to write a quadrilateral inequality, the so-called quadrilateral information-distance Bell inequality for the four Bell-like quantum observables[20] (Fig. 6.2)

[19] Beltrametti and Maczyński (1991). For a general summary on entropic Bell inequalities see Auletta (2000, Sect. 42.7).

[20] Schumacher (1990, 1991).

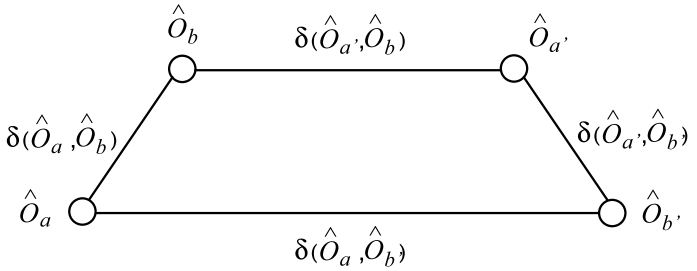


Fig. 6.2 Schematic information-theoretic representation of quantum non-separability. The information distance $\delta(\hat{O}_a, \hat{O}^{b'})$ is greater than is allowed by the classical metric properties of informational distance. Adapted from Auletta et al. (2009, p. 636)

$$\delta(\hat{O}_a, \hat{O}_b) + \delta(\hat{O}_b, \hat{O}_{a'}) + \delta(\hat{O}_{a'}, \hat{O}_{b'}) \geq \delta(\hat{O}_a, \hat{O}_{b'}) \ . \tag{6.52}$$

If we represent those observables with co-planar vectors, fix the information distance between $\hat{O}_a$ and $\hat{O}_b$ (separated by angle θ) to be

$$\delta(\hat{O}_a, \hat{O}_b) = 2f\left(\frac{\theta}{2}\right) \ , \tag{6.53}$$

where $f(\phi) = -\cos^2\phi \ln\cos^2\phi - \sin^2\phi \ln\sin^2\phi$, and take the angle between $\hat{O}_a$ and $\hat{O}_b$, between $\hat{O}_b$ and $\hat{O}_{a'}$, between $\hat{O}_{a'}$ and $\hat{O}_{b'}$ to be equal to $\pi/8$, we arrive at the following values:

$$\delta(\hat{O}_a, \hat{O}_b) = \delta(\hat{O}_{a'}, \hat{O}_b) = \delta(\hat{O}_{a'}, \hat{O}_{b'}) = 2f\left(\frac{\pi}{16}\right) \simeq 0.323, \tag{6.54a}$$

$$\delta(\hat{O}_a, \hat{O}_{b'}) = 2f\left(\frac{3\pi}{16}\right) \simeq 1.236 \ . \tag{6.54b}$$

But since $0.323 + 0.323 + 0.323 < 1.236$ it is evident that the Bell inequality (6.52) is violated.

Now, we have to consider the other part of our problem, i.e. the local operations themselves. In fact, such a theory of information (which not by chance consists in no-go theorems) does not tell us how information is in fact locally selected and acquired. For treating this problem, we need to consider the role of signals in quantum theory.

6.2 Signal Exchanging

Signals

As stressed in the two previous chapters, in order to have a physical universe one needs (i) well-defined physical states of the involved systems (in other words prepa-

ration procedures and their spontaneous analogues), (ii) constraining influences (in the sense of local contexts potentially able to give rise to certain events), and (iii) the events themselves. In few words, a physical universe is the context in which information is manipulated, exchanged and acquired. This is different relative to the possible information, characterising all quantum systems as far as they are free systems. In fact, in Sect. 3.4.2 we have agreed that dynamical interactions and causal processes happen with exchange of signals.

I recall that a *signal* is any oscillation, modification or perturbation evolving in time, thus building patterns. Any signal must be vehicled through a physical system, like a photon or electromagnetic wave. For this reason, although, as we shall see, signals are generated by quantum sources, a pure quantum-mechanical signal does not exist, as it is evident through the discussion of the EPR paper (Sect. 3.3), especially concerning the role of correlations. Moreover, the information that it vehicles must be somehow codified in the structure of the signal. For instance, the spectrum of stellar electromagnetic radiation carries with itself the information about the element that emitted it.

Signals show important similarities with information. Signals are scale invariant, that is, they represent the same information (and pattern) when multiplied by some positive constant.[21] A signal $s(t)$ is normalised if

$$\int s^2(t)dt = 1. \tag{6.55}$$

Moreover, orthonormal signals $s(t)$, $s'(t)$ satisfy

$$\int s(t)s'(t)dt = 0. \tag{6.56}$$

6.2.1 Channels for Exchanging Information

Two Requirements

Exchanging (acquiring) information needs codification procedures and needs a channel. About the first requirement, in fact any information exchange (acquiring) is made through signals that can be codified and need to be codified during any information acquisition, although often do not fulfil the requirements of linear codification, typical of quantum systems. With this, it is meant that a signal is not necessarily part of a code with some elementary codifying units (Sect. 3.2.3), since a signal can represent any kind of alteration in some physical or chemical 'stuff'. Nevertheless, it can always be coded at the reception, since, as mentioned, it presents in itself patterns that in general allow some kind of Fourier decomposition in 'pure tones', and these pure tones can be understood as the elements that allow the codification. Then, we can use

[21] Battail (2014, Sect. 3.1).

Fourier transforms and their inverses (Sect. 1.2.3). This is also true for non-linear signals, as it has been recently shown.[22] The reason for the second requirement is that any kind of transformation 'goes through' a kind of channel, according to what said in Sect. 5.2.4. This is especially relevant when the effect of other systems on the information transmission is considered. Through an operation $\hat{T}$ and its dual (Sect. 5.2.5) we can formally describe a generalised transmission of states from present to the future (through some kind of channel), where this transmission is deterministic only if it is trace-preserving, otherwise it undergoes some kind of local dispersion (and the channel is noisy). Although QM is deterministic in its specific sense (Sects. 2.4.2 and 3.2.4), we can distinguish here between deterministic–probabilistic transformations, which are trace-preserving and in this context can be called *deterministic* in short, and *probabilistic* transformations that are trace-decreasing.

Encoding/Decoding

Let us first make some basic considerations on coding.[23] When we transmit information, we need first to apply a suitable deterministic operation $\hat{\mathcal{E}}$ called (linear) *encoding*, thanks to which we transform an input state $\hat{\rho}$ into

$$\left|\hat{\rho}'\right) = \hat{\mathcal{E}}\left|\hat{\rho}\right) . \tag{6.57}$$

We say that such an encoding is lossless for the state $\left|\hat{\rho}\right)$ iff there exists another deterministic operation $\hat{\mathcal{D}}$, called (linear) *decoding*, such that, $\forall \left|\hat{\rho}\right) \in F_{\hat{\rho}}$, where $F_{\hat{\rho}}$ is the set of all states that are compatible with (non-orthogonal to) $\hat{\rho}$, we have

$$\hat{\mathcal{D}}\hat{\mathcal{E}}\left|\hat{\rho}\right) = \left|\hat{\rho}\right) . \tag{6.58}$$

In the general case, both encoding and decoding are operations that could be represented by means of amplitude operators. Since encoding is given by the transformation (6.57), this means that any state in the set $F_{\hat{\rho}}$ is encoded in a state of the set $F_{\hat{\rho}'}$, that is,

$$\hat{\mathcal{E}} F_{\hat{\rho}} \subseteq F_{\hat{\rho}'} . \tag{6.59}$$

The symbol $\subseteq$ means "subset of" (including itself) while $\subset$ denotes a proper subset (non including itself). When we have equality, we say that the encoding is *efficient*. In other words, here the encoding operation does not waste space and the set $F_{\hat{\rho}'}$ is as small as possible. When the encoding is both lossless and efficient, we say that it is *ideal*. The ultimate limit of ideal compression (Sect. 3.2.4) of a state $\left|\hat{\rho}\right)$ is reached when every state $\left|\hat{\rho}'\right)$ of the encoding system is of the form $\hat{\mathcal{E}}\left|\hat{\rho}\right)$ with some $\left|\hat{\rho}\right) \in F_{\hat{\rho}}$. Thus, a transformation $\hat{\mathcal{E}}$ is an ideal encoding for the state $\left|\hat{\rho}\right)$ of system $\mathcal{A}$ iff there exists a decoding $\hat{\mathcal{D}}$ for system $\mathcal{A}'$ such that we have both

$$\hat{\mathcal{D}}\hat{\mathcal{E}} =_{\hat{\rho}} \hat{I}_A \text{ and } \hat{\mathcal{E}}\hat{\mathcal{D}} =_{\hat{\rho}'} \hat{I}_{A'} , \tag{6.60}$$

[22] Singh et al. (2017).

[23] For the following considerations see D'Ariano et al. (2017, Sects. 8.2–8.5).

where the subscripts denote the density matrices on which the transformations apply. The first requirement tells us that $\hat{\mathcal{E}}$ is lossless, what implies the reverse of Eq. (6.59)

$$\hat{\mathcal{D}} F_{\hat{\rho}'} \subseteq F_{\hat{\rho}}; \tag{6.61}$$

the second condition guarantees that that every $\hat{\varsigma}' \in F_{\hat{\rho}'}$ can be written as

$$\left|\hat{\varsigma}'\right) = \hat{\mathcal{E}}\hat{\mathcal{D}}\left|\hat{\varsigma}'\right) = \hat{\mathcal{E}}\left|\hat{\varsigma}\right), \tag{6.62}$$

with

$$\left|\hat{\varsigma}\right) = \hat{\mathcal{D}}\left|\hat{\varsigma}'\right) \in F_{\hat{\rho}'}. \tag{6.63}$$

This result tells us that $\left|\hat{\rho}\right)$ can be ideally encoded into $\left|\hat{\rho}'\right)$ iff $\left|\hat{\rho}'\right)$ can be ideally encoded into $\left|\hat{\rho}\right)$. Thus, ideally encodings can be identified with deterministic transformations that map pure states into pure states. The deterministic transformations (6.60) represent a *quantum channel* between systems $\mathcal{A}$ and $\mathcal{A}'$ or also between past and future states of the same system. In the case of the channel between two different systems we should rather speak of quantum-classical channel, provided that classical correlations are integral part of entanglement (Sect. 5.1.3). In the pure classical case, we have a similar situation but with the (classical) deterministic transformation happening between bits and classical states. Of course, many transformations are not deterministic, especially when other systems influence somehow the dynamics of the system or systems under consideration, as it is evident with the master equation and similar formalism (Sect. 5.2.4). As said, such an influence manifests itself as a *noise* disturbing our channel. In fact, any signal is characterised by the signal–noise ratio.

Noisy Channel and Error Correction

As I have mentioned, signals spread and this implies that noise can somehow interferes with their transmission. It may then be asked in which conditions we would ideally recover the whole information represented by the signal. Suppose that we want to send some piece of information carried by a system $\mathcal{A}$ through a noisy channel $\hat{\mathcal{N}}$, which could, for example represent the effect of an external environment. The goal is to find two operations (encoding $\hat{\mathcal{E}}$ and decoding $\hat{\mathcal{D}}$) such that in the end this information is left intact:

$$\hat{\mathcal{D}}\hat{\mathcal{N}}\hat{\mathcal{E}} = \hat{I}_A. \tag{6.64}$$

We say that $\hat{\mathcal{E}}$ is a good encoding for $\hat{\mathcal{N}}$ if the above condition is satisfied for some $\hat{\mathcal{D}}$. It can then be proved that every good encoding for $\hat{\mathcal{N}}$ is a convex combination of ideal encodings, each of which is a good encoding for $\hat{\mathcal{N}}$. In order to have good encoding, we must be able to undo the action of noise on all states on the set $F_{\hat{\rho}}$. Indeed, defining a *recovery operation* as

$$\hat{\mathcal{R}} := \hat{\mathcal{E}}\hat{\mathcal{D}}, \tag{6.65}$$

and making use of Eq. (6.64), $\forall \left| \hat{\varsigma} \right)$ in the state space of $\mathcal{A}$ we have

$$\hat{\mathcal{R}}\hat{\mathcal{N}}\hat{\mathcal{E}} \left| \hat{\varsigma} \right) = \hat{\mathcal{E}} \left(\hat{\mathcal{D}}\hat{\mathcal{N}}\hat{\mathcal{E}} \right) \left| \hat{\varsigma} \right) = \hat{\mathcal{E}} \left| \hat{\varsigma} \right) , \tag{6.66}$$

where $\hat{\mathcal{N}}\hat{\mathcal{E}} \left| \hat{\varsigma} \right)$ represents the application of the transformation represented by the transmission of the coded information $\hat{\mathcal{E}} \left| \hat{\varsigma} \right)$ through a noisy channel. Since, given Eq. (6.57), $\hat{\mathcal{E}} \left| \hat{\varsigma} \right)$ is an arbitrary state $\in F_{\hat{\rho}'}$, this is a sufficient condition of

$$\hat{\mathcal{R}}\hat{\mathcal{N}} =_{\hat{\rho}} \hat{I}_{A'}. \tag{6.67}$$

In other words, the recovery operation removes the noise. As a consequence, we can regard any noise as an *error*, and consider our recovery as *error correction*. This allows us to understand deterministic non-dissipative and reversible dynamics as the limiting case in which the noise is dropped through a successful error correction.

6.2.2 *No Information Acquisition Without Disturbance*

The Question

Measurement is a special case of information transmission. Thus, it may be asked whether it is possible to acquire information during measurement in a reversible way. Although we have already provided a negative answer to this question, a detailed analysis is necessary here in connection with the issue noise–disturbance error. It is in fact possible to show that an ideal measurement is reversible iff no information about the initial state is obtained, and then, if so, it cannot be considered a true measurement (nor as information acquisition). Thus, if we were able to fully correct the local disturbance generated by a detection, we needed to renounce to acquire any information at all. In other words, a full error correction makes information acquisition impossible.

In a pioneering work, the Japanese physicist Masahito Ueda and information engineer Masahito Kitagawa[24] showed that a measurement on photons can be made reversible if it is unsharp (Sect. 5.2.4), and if it is sensitive to the vacuum field fluctuations. However, they did not consider a conserved quantity (Sect. 1.2.4) and, for this reason, the argument was not definitive. A conserved quantity, the photon-number operator, was considered in a paper by the Turkish–American physicist Atac Imamoğlu,[25] who investigated the possibility of repetitive (logically reversible) measurements. Let us consider these concepts.

[24] See Ueda and Kitagawa (1992).

[25] See Imamoğlu (1993).

Number Operator

The presentation of basic concepts of quantum-optics (the theory dealing with properties of the quantum electromagnetic field) is relevant for understanding the model under discussion but also for the whole development of this part of the book.[26] Up to now, I have introduced raising and lowering operators for angular momentum (Sect. 1.2.4) and for 2D Hilbert spaces only when discussing the model of a two-level atom coupled with a field (Sect. 5.2.4). A generalisation to n dimensions (with the possibility to have up to n photons in the field) is provided by the annihilation and creation operators

$$\hat{a} := \sqrt{\frac{m}{2\hbar\omega}} \left(\omega \hat{x} + i \frac{\hat{p}_x}{m} \right) , \tag{6.68a}$$

$$\hat{a}^\dagger := \sqrt{\frac{m}{2\hbar\omega}} \left(\omega \hat{x} - i \frac{\hat{p}_x}{m} \right) , \tag{6.68b}$$

respectively, where $\hat{x}$, $\hat{p}_x$ are position and momentum observables with 1 Cartesian dimension and ω is the angular frequency (5.87). I also recall that, although massless, photons display an associated momentum as shown in Eq. (1.3). It should be emphasised that the operators (6.68) are not Hermitian— in fact, one is the Hermitian conjugate or adjoint of the other— and therefore are not observables. Note also that $\hat{a} + \hat{a}^\dagger$ is proportional to $\hat{x}$ while $\hat{a} - \hat{a}^\dagger$ is proportional to $\hat{p}_x$. These operators allow us to derive the *number operator* (which is an observable: see property (1.82f)) as

$$\begin{aligned} \hat{N} := \hat{a}^\dagger \hat{a} &= \frac{m}{2\hbar\omega} \left(\omega \hat{x} - i \frac{\hat{p}_x}{m} \right) \left(\omega \hat{x} + i \frac{\hat{p}_x}{m} \right) \\ &= \frac{m}{2\hbar\omega} \left(\omega^2 \hat{x}^2 + \frac{i\omega}{m} \hat{x}\hat{p}_x - \frac{i\omega}{m} \hat{p}_x \hat{x} + \frac{\hat{p}_x^2}{m^2} \right) \\ &= \frac{m}{2\hbar\omega} \left(\omega^2 \hat{x}^2 + \frac{i\omega}{m} \left[\hat{x}, \hat{p}_x \right] + \frac{\hat{p}_x^2}{m^2} \right) \\ &= \frac{1}{\hbar\omega} \left(\frac{1}{2} m\omega^2 \hat{x}^2 + \frac{\hat{p}_x^2}{2m} \right) - \frac{1}{2} , \end{aligned} \tag{6.69}$$

where I have made use of the commutation relation (1.209). The term in brackets on the last row of Eq. (6.69) is, for example the harmonic-oscillator Hamiltonian (one of the simplest systems described by these operators) (Fig. 6.3). Thus, we can write

$$\hat{H} = \hbar\omega \left(\hat{N} + \frac{1}{2} \right) . \tag{6.70}$$

Therefore, in the energy representation, the number operator is diagonal and its nth diagonal term is just equal to n. This is the reason why it is called the *number* operator,

[26] See Auletta et al. (2009, Sect. 4.4).

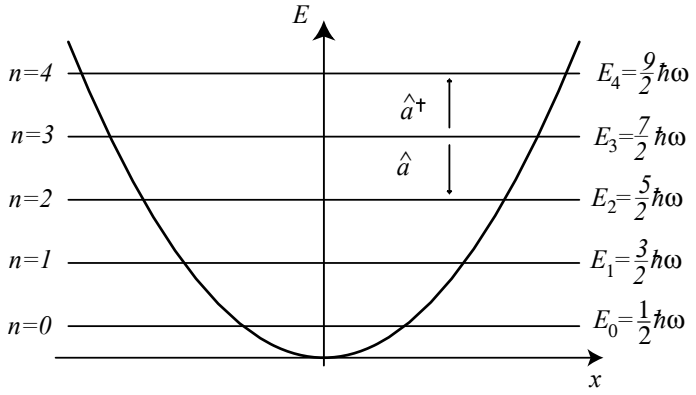


Fig. 6.3 Bell-like potential and energy levels of the harmonic oscillator. The action of the creation and annihilation operators is illustrated by the arrows. Adapted from Auletta et al. (2009, Sect. p. 154)

i.e.

$$\hat{N}\,|n\rangle = n\,|n\rangle \quad , \tag{6.71}$$

where $|n\rangle$, according to the Bose–Einstein statistics (Sect. 1.3.1), is a (normalised) state with n photons, related to energy eigenstates (1.197) and (4.110) through Eq. (6.70), showing that we have a case of degeneracy (Sect. 1.2.4). The $|n\rangle$'s are called *number states* or also *Fock states*, after the name of the Russian physicist and mathematician Vladimir Fock (1898–1974). This allows us to consider the previous formula as the eigenvalue equation of the number operator. The reader might find the explicit expressions for the number, annihilation, and creation operators helpful:

$$\hat{N} = \begin{bmatrix} 0 & 0 & 0 & 0 & \dots \\ 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 2 & 0 & \dots \\ 0 & 0 & 0 & 3 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{bmatrix} , \tag{6.72a}$$

$$\hat{a} = \begin{bmatrix} 0 & \sqrt{1} & 0 & 0 & \dots \\ 0 & 0 & \sqrt{2} & 0 & \dots \\ 0 & 0 & 0 & \sqrt{3} & \dots \\ 0 & 0 & 0 & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \end{bmatrix} , \tag{6.72b}$$

$$\hat{a}^\dagger = \begin{bmatrix} 0 & 0 & 0 & 0 & \dots \\ \sqrt{1} & 0 & 0 & 0 & \dots \\ 0 & \sqrt{2} & 0 & 0 & \dots \\ 0 & 0 & \sqrt{3} & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{bmatrix} , \tag{6.72c}$$

where the matrix expression for the number operator is quite natural and the matricial expressions for annihilation and creation operators follow when Eq. (6.69) is considered. Note in fact that the only non–zero elements of the annihilation and creation operators are located immediately above and below the diagonal (since they express each time a single photon lost and acquired), respectively, which shows that they are indeed generalisations of the binary lowering and raising operators (5.89), with the two vectors $|0\rangle$, $|1\rangle$ replaced by $|n\rangle$ vectors on the outline of basis (1.70). In other words, we have the following series for the annihilation operator

$$\hat{a}\,|0\rangle = 0,\ \hat{a}\,|1\rangle = \sqrt{1}\,|0\rangle\ ,\ \hat{a}\,|2\rangle = \sqrt{2}\,|1\rangle\ ,\ldots, \tag{6.73a}$$

where the eigenvectors represent (up to the corresponding eigenvalue) the columns of $\hat{a}$, and similarly for the the creation operator:

$$\hat{a}^\dagger\,|0\rangle = \sqrt{1}\,|1\rangle\ ,\ \hat{a}^\dagger\,|1\rangle = \sqrt{2}\,|2\rangle\ ,\ \hat{a}^\dagger\,|2\rangle = \sqrt{3}\,|3\rangle\ ,\ldots, \tag{6.73b}$$

where the eigenvectors represent (up to the corresponding eigenvalue) the rows of $\hat{a}^\dagger$. These series allow us to derive the following fundamental relations in an easy way (they can be obviously directly derived by using the explicit expressions (6.68)):

$$\hat{a}\,|n\rangle = \sqrt{n}\,|n-1\rangle\ \ ,\ \ \hat{a}^\dagger\,|n\rangle = \sqrt{n+1}\,|n+1\rangle\ \ , \tag{6.74}$$

from which we have

$$\begin{aligned}\hat{a}^\dagger\hat{a}\,|n\rangle &= \hat{a}^\dagger\left(\sqrt{n}\,|n-1\rangle\right) = n\,|n\rangle\\ &= \hat{N}\,|n\rangle\ ,\end{aligned} \tag{6.75}$$

in accordance with Eq. (6.71). Taking advantage of the fact that

$$\hat{a}\hat{a}^\dagger\,|n\rangle = \hat{a}\left(\sqrt{n+1}\,|n+1\rangle\right) = (n+1)\,|n\rangle\ , \tag{6.76}$$

from which

$$\hat{a}\hat{a}^\dagger = \hat{I} + \hat{N}, \tag{6.77}$$

we derive the commutation relation

$$\left[\hat{a},\hat{a}^\dagger\right] = \hat{a}\hat{a}^\dagger - \hat{a}^\dagger\hat{a} = \left(\hat{I}+\hat{N}\right) - \hat{N} = \hat{I}\ , \tag{6.78a}$$

and similarly

$$\left[\hat{a}^\dagger,\hat{a}\right] = -\hat{I}\ . \tag{6.78b}$$

The sequence $\hat{a}^\dagger\hat{a}$ is called normal ordering while the sequence $\hat{a}\hat{a}^\dagger$ antinormal ordering.

No Reversible Measurement

Coming back to the problem of reversible measurement, Imamoglu has proved that, under certain conditions, the state of the electromagnetic field before and after the measurement is unchanged.[27] However, those conditions contradict the results of our analysis of measurement, which requires some event that in turn induces some disturbance. To see why it is so, let us consider also the subsequent work of the American physicist Hideo Mabuchi and the Austrian physicist Peter Zoller,[28] who have shown under which conditions reversibility is possible in the case of a system coupled to an environment. In short, they prove that the action of a jump operator describing the collapse of a wave function cannot in general be inverted as such. Inversion is possible if one considers the system as pertaining to a subset of the original Hilbert space and the jump is unitary—in this case, however, no new information is obtained. Let us consider the action of the annihilation operator $\hat{a}$ (which causes the jump by photon emission, for example). We assume now that this annihilation operator acts as an unitary operator *on a subspace* $\mathcal{H}_1$ of the original Hilbert space $\mathcal{H}$ of the system (since it cannot be unitary on the whole Hilbert space). With the expression $\mathcal{H}_1$ is meant that at least one photon is initially present. In other words, we postulate

$$\hat{a}|\psi\rangle = \hat{U}|\psi\rangle, \quad \langle\psi|\hat{a}^\dagger = \langle\psi|\hat{U}^\dagger \,, \tag{6.79}$$

where the generic ket $|\,\psi\rangle \in \mathcal{H}_1$ describes the system. Due to the unitarity of $\hat{U}$, we also have

$$1 = \langle\psi\,|\,\psi\rangle = \langle\psi|\hat{U}^\dagger\hat{U}|\psi\rangle = \langle\psi|\hat{a}^\dagger\hat{a}|\psi\rangle = \langle\psi|\hat{N}|\psi\rangle \,, \tag{6.80}$$

which means that the expectation value of the photon number is equal to unity for an arbitrary state from the specified subset $|\psi\rangle \in \mathcal{H}_1$. If we expand $|\psi\rangle$ by making use of the orthogonal basis represented by the eigenstates of the number operator, according to Eq. (6.71), we get

$$|\psi\rangle = c_1|1\rangle + \cdots + c_n|n\rangle \,, \tag{6.81a}$$

with $c_j \in \mathbb{C}$ $(j = 0, 1, \cdots, n)$, subject to the usual normalisation condition. Then, considering its bra

$$\langle\psi\,| = \langle 1\,|\,c_1^* + \cdots + \langle n\,|\,c_n^*, \tag{6.81b}$$

and taking the scalar product of these two equations we can rewrite Eq. (6.80) as

$$\wp_1 + \wp_2 + \wp_3 + \cdots + \wp_n = 1 \,, \tag{6.82}$$

with $\wp_n = |c_n|^2$. For this equation to be fulfilled, at least one of the numbers $\wp_1, \wp_2, \ldots \in \mathbb{R}$ must be non-zero. Since an annihilation operator diminishes the number of photons by a unity, the fact that a jump occurred gives the informa-

[27] For what follows see Auletta et al. (2009, Sect. 15.3).

[28] See Mabuchi and Zoller (1996). See also Mensky (1996).

tion that the initial state was not the vacuum, but this piece of information was already contained in the assumption that the initial state belonged to the subspace $\mathcal{H}_1$. Since $|\psi\rangle \in \mathcal{H}_1$ depends on the preparation, no new information is gained by such a reversible quantum jump. The same argument can be extended to a double jump (annihilation) operator and in fact to any number of jumps.

Some Lessons

On this basis, we can conclude what follows:

- In general, the action of a jump operator (which may describe certain dissipative processes and certain kinds of measurement) is not locally unitary and cannot be inverted.
- If we restrict the possible states to a properly chosen subset of the original Hilbert space of the system, the action of a jump operator may become unitary and therefore, under those conditions, reversible.
- However, in all those cases in which the action of a jump operator is reversible, the measurement described by the jump operator itself gives no new information besides that already included in the preparation of the initial state.

A Generalisation

Up to now, we have considered a specific model. The American physicists Michael Nielsen and Carlton Caves[29] provided a powerful formal generalisation of the previous result. Let us recall here the formula (5.153a), for an arbitrary density matrix $\hat{\rho} = |\psi\rangle\langle\psi|$ (with $|\psi\rangle \in \mathcal{H}$) describing the system:

$$\hat{\mathcal{T}}\left|\hat{\rho}\right) = \sum_k \hat{\vartheta}_k \hat{\rho} \hat{\vartheta}_k^\dagger \,. \tag{6.83}$$

If we have a single (not necessarily unitary) operator such that

$$\hat{\mathcal{T}}\left|\hat{\rho}\right) = \hat{\vartheta}\hat{\rho}\hat{\vartheta}^\dagger \text{ and } \hat{\vartheta}^\dagger\hat{\vartheta} = \hat{E} \,, \tag{6.84}$$

then we speak of an *ideal* measurement (perfect readout of the state of the apparatus). We then have a unitarily reversible measurement—on a subspace $\mathcal{H}_0$ of the Hilbert space $\mathcal{H}$ of the original problem—if there exists a unitary operator $\hat{U}_0$ acting on $\mathcal{H}_0$ such that

$$\left|\hat{\rho}\right) = \frac{\hat{\mathcal{U}}_0\hat{\mathcal{T}}\left|\hat{\rho}\right)}{\left(\hat{\mathcal{T}}\left|\hat{\rho}\right)\right)} = \hat{U}_0 \frac{\hat{\vartheta}\hat{\rho}\hat{\vartheta}^\dagger}{\mathrm{Tr}\left(\hat{\vartheta}\hat{\rho}\hat{\vartheta}^\dagger\right)} \hat{U}_0^\dagger, \tag{6.85}$$

$\forall \left|\hat{\rho}\right)$ whose support lies in $\mathcal{H}_0$. The denominator is here for normalising the output state. This means that the effect $\hat{E}$, when restricted to $\mathcal{H}_0$, is a positive multiple of the identity operator on $\mathcal{H}_0$, i.e.

[29] See Nielsen and Caves (1997).

$$\hat{P}_{\mathcal{H}_0}\hat{E}\,\hat{P}_{\mathcal{H}_0} = \eta\hat{P}_{\mathcal{H}_0}, \tag{6.86}$$

where $\hat{P}_{\mathcal{H}_0}$ is a projector onto the subspace $\mathcal{H}_0$ and $\eta \in \mathbb{R}$ is a constant satisfying $0 < \eta \leq 1$. In other words, we have

$$\hat{P}_{\mathcal{H}_0} + \hat{P}_{\mathcal{H}_0^\perp} = \hat{I}, \tag{6.87}$$

where $\hat{P}_{\mathcal{H}_0^\perp}$ projects onto the subspace that is complementary of $\mathcal{H}_0$. Then, the operator $\hat{\vartheta}$ can be written as

$$\begin{aligned}\hat{\vartheta} &= \hat{\vartheta}\hat{I} \\ &= \hat{\vartheta}\left(\hat{P}_{\mathcal{H}_0} + \hat{P}_{\mathcal{H}_0^\perp}\right) \\ &= \eta\hat{U}\,\hat{P}_{\mathcal{H}_0} + \hat{\vartheta}\hat{P}_{\mathcal{H}_0^\perp}\end{aligned} \tag{6.88}$$

where $\hat{U}$ is some unitary operator acting on the whole Hilbert space $\mathcal{H}$. As a consequence, recalling expression (5.75), we have

$$\left(\hat{E}\,\middle|\,\hat{\rho}\right) = \mathrm{Tr}\left(\hat{\vartheta}\hat{\rho}\hat{\vartheta}^\dagger\right) = \eta^2 \tag{6.89}$$

for all density operators $\left|\,\hat{\rho}\right)$ whose support lies in $\mathcal{H}_0$, and where η^2 takes the meaning of the probability of occurrence of the result represented by $\hat{\vartheta}$. In other words, the probability of occurrence of any measurement result represented by $\hat{\vartheta}$ is the same for all states that are normalised in that subspace.

Since any event necessarily represents a disturbance, although not always of the measured system (in fact, in Sect. 2.1.2 we have seen that also an interaction-free measurement needs a detection), the impossibility to measure in a reversible way excluding local irreversible events means the impossibility to acquire information without disturbance.[30] In fact, we say that a collection of effects $\{\hat{E}_j\}$ is a non-disturbing test on the input $\left|\,\hat{\rho}\right)$ if its coarse graining is proportional to the identity.[31]

6.2.3 *Informational Completeness and Tomography*

Informational Completeness

The previous examination raises the question of the extent to which we can extract information from a quantum state and thus how much an observable's measurement

[30] In Fuchs and Peres (1996) the trade-off (in terms of measurement uncertainty relation) between information gain and state disturbance has been computed. See also Pfister and Wehner (2013).

[31] D'Ariano et al. (2017, Sect. 2.10.3 and Sect. 7.7).

informs us about the state of a given system.[32] Stated in other terms, we may raise the question of whether the probability distributions of a certain set of observables are sufficient to determine the state of a quantum system, i.e. to discriminate between different states. Such a question leads naturally to the concept of *informational completeness* and brings us back to the relationship between sharp and unsharp observables (Sect. 5.2.4).

A family of self-adjoint operators $\{\hat{O}_k\}$ is said to be informationally complete if, $\forall k$ (see e.g. Eq. (5.148)),

$$\left(\hat{O}_k \,|\hat{\rho}\right) = \left(\hat{O}_k \,|\hat{\rho}'\right) \tag{6.90}$$

implies that $\hat{\rho} = \hat{\rho}'$ on the Hilbert space $\mathcal{H}$ of the system. Since unsharp observables are obtained by a smoothing operation (a kind of coarse graining) on sharp observables, we could infer that they provide less information than sharp ones. This inference is in fact not true, as I shall show now.

Pauli Problem

In fact, take as an example the relationship between the momentum $\hat{p}_x$ and the position $\hat{x}$ of a 1D system. The famous *Pauli problem,*[33] after the name of W. Pauli, is summarised in the following question: do the position and momentum distributions determine the wave function uniquely? We can now reformulate it in the language of informational completeness of the canonically conjugate position and momentum observables: are sharp complementary observables informationally complete? As we shall prove by making use of a counterexample, this is not the case.

Take[34] a normalised 1D wave function $\psi(x) \in L^2(x)$ of the kind (3.104):

$$\psi(x) = |\psi(x)|e^{i\phi(x)} \,, \tag{6.91}$$

where

$$|\psi(x)| = |\psi(-x)| \,. \tag{6.92}$$

The phase function $\phi(x)$ satisfies

$$0 \leq \phi(x) < 2\pi \text{ and } \phi(x) + \phi(-x) \neq \text{constant} \,(\text{mod}\, 2\pi) \,. \tag{6.93}$$

Then, thanks to Eq. (6.92), the wave function

$$\psi'(x) = \psi^*(-x) = |\psi(x)|e^{-i\phi(-x)} \tag{6.94}$$

represents a state different from $\psi(x)$, where a look at Fig. 4.5, Sect. 4.2.3, may be helpful, but with

[32] For this whole subsection see Auletta et al. (2009, Sects. 13.4–5 and 15.4).

[33] See Pauli (1980, 17).

[34] See Prugovečki (1977).

$$|\psi'(x)|^2 = |\psi(x)|^2 \text{ and } |\tilde{\psi}'(p_x)|^2 = |\tilde{\psi}(p_x)|^2 \,, \tag{6.95}$$

where $\tilde{\psi}(p_x)$ is the Fourier transform of $\psi(x)$ onto the momentum representation given by Eqs. (1.175). Therefore, the pair $\hat{p}_x$, $\hat{x}$ is not able to distinguish between the states ψ and ψ'.

Unsharp Observables

As said, unsharp observables are the result of a smearing operation on sharp ones, say $\hat{\mathbf{p}}$ and $\hat{\mathbf{r}}$, that can be understood as a coarse-graining operation. This operation can have an informationally complete refinement as a result: in fact, a set of unsharp complementary observables can always be chosen that is informationally complete: the probability distributions of sharp momentum and position (or relative quadratures, which I shall introduce now) do not cover the whole phase space of the system, whereas they do if are taken to be unsharp. Our task is to find counterparts of classical variables that would allow the joint treatment of conjugate observables. This can be done with a family of functions that are phase–space representations (see also Fig. 2.7, Sect. 2.2.2) that are called quasi–probability distributions, whose most famous one is the *Wigner function*, after the name of the Hungarian physicist Eugene Wigner. This function is of great use in quantum optics.

Quadratures

In order to deal with quasi-probability distributions, it is convenient to introduce the quadrature operators,[35] defined by

$$\hat{X}_1 := \frac{1}{\sqrt{2}}\left(\hat{a}^\dagger + \hat{a}\right) = \sqrt{\frac{2m\omega}{\hbar}}\hat{x}, \quad \hat{X}_2 := \frac{i}{\sqrt{2}}\left(\hat{a}^\dagger - \hat{a}\right) = -i\sqrt{\frac{2}{\hbar\omega}}\hat{p}_x, \tag{6.96}$$

where $\hat{a}$, $\hat{a}^\dagger$ are the annihilation and creation operators (6.68). The quadrature operators may be viewed as the dimensionless electric and magnetic fields of a single field mode. In the context of a 1D particle subjected to a harmonic-oscillator potential, $\hat{X}_1$ and $\hat{X}_2$ represent the dimensionless position and momentum operator, respectively. From Property (6.78a), it trivially follows that the quadratures obey the commutation relation

$$\begin{aligned}\left[\hat{X}_1, \hat{X}_2\right] &= \frac{i}{2}\left(\hat{a}^\dagger + \hat{a}\right)\left(\hat{a}^\dagger - \hat{a}\right) - \frac{i}{2}\left(\hat{a}^\dagger - \hat{a}\right)\left(\hat{a}^\dagger + \hat{a}\right) \\ &= \frac{i}{2}\left[\left(\hat{a}^\dagger\right)^2 - \hat{N} + (\hat{I} + \hat{N}) - \hat{a}^2\right] - \frac{i}{2}\left[\left(\hat{a}^\dagger\right)^2 + \hat{N} - (\hat{I} + \hat{N}) - \hat{a}^2\right] \\ &= i\hat{I} \,. \end{aligned} \tag{6.97}$$

Moreover, from Eqs. (6.97) and (1.286), we obtain the quadrature uncertainty relation

$$\Delta\hat{X}_1 \Delta\hat{X}_2 \geq \frac{1}{2} \,. \tag{6.98}$$

[35] See Yuen (1976).

Coherent States

I have already mentioned (see Eq. (3.41)) *coherent* states (particularly important for photons pumped by lasers, which represent in fact coherent light), but let us now introduce them formally. Since the vacuum is an eigenstate of the annihilation operator with eigenvalue 0, i.e.

$$\hat{a}\,|0\rangle = 0\,|0\rangle = 0\,. \tag{6.99}$$

we assume the existence of states $|\alpha\rangle$ that are eigenstates of $\hat{a}$:

$$\hat{a}\,|\alpha\rangle = \alpha\,|\alpha\rangle \quad , \tag{6.100a}$$

and, in agreement with Property (1.82c), similarly

$$\langle\alpha|\,\alpha^* = \langle\alpha|\,\hat{a}^\dagger\,, \tag{6.100b}$$

with α and α^* being the mean values of $\hat{a}$ and $\hat{a}^\dagger$, respectively, taken on the state $|\alpha\rangle$:

$$\langle\alpha\,|\hat{a}|\,\alpha\rangle = \alpha \text{ and } \langle\alpha|\,\hat{a}^\dagger\,|\alpha\rangle = \alpha^*. \tag{6.101}$$

I recall that $\hat{a}$, $\hat{a}^\dagger$ are in fact not observables. Now, let us make use of the completeness relation for eigenstates of the number operator:

$$|\alpha\rangle = \sum_n |n\rangle\,\langle n\,|\,\alpha\rangle \quad , \tag{6.102}$$

where the amplitudes $\langle n\,|\,\alpha\rangle$ are the coefficients of this expansion. Multiplying Eq. (6.100a) by $\langle n\,|$ from the left, we get (for $\alpha \neq 0$) the explicit form of these coefficients

$$\langle n\,|\,\alpha\rangle = \frac{\langle n\,|\hat{a}|\,\alpha\rangle}{\alpha} = \frac{\sqrt{n+1}}{\alpha}\,\langle n+1\,|\,\alpha\rangle \quad , \tag{6.103}$$

where I have made use of Eqs. (6.74). Since

$$\langle n+1\,| = \langle 0|\,\frac{\hat{a}^{n+1}}{\sqrt{(n+1)!}}\,, \tag{6.104}$$

the RHS of Eq. (6.103) becomes

$$\langle n\,|\,\alpha\rangle = \frac{\langle 0\,|\hat{a}^{n+1}|\,\alpha\rangle}{\alpha\sqrt{n!}} = \frac{\alpha^n}{\sqrt{n!}}\,\langle 0\,|\,\alpha\rangle \quad , \tag{6.105}$$

since, from Eq. (6.100a), we have $\hat{a}^{n+1}\,|\alpha\rangle = \alpha^{n+1}\,|\alpha\rangle$. By inserting this result into Eq. (6.102), we obtain

$$|\,\alpha\rangle = \sum_n \langle 0\,|\,\alpha\rangle\,\frac{\alpha^n}{\sqrt{n!}}\,|\,n\rangle \quad . \tag{6.106}$$

By imposing the normalisation condition

$$\begin{aligned} 1 = \langle\alpha\,|\,\alpha\rangle &= \sum_{n,m} |\,\langle 0\,|\,\alpha\rangle\,|^2 \frac{\alpha^n(\alpha^*)^m}{\sqrt{n!m!}}\,\langle m\,|\,n\rangle \\ &= |\,\langle 0\,|\,\alpha\rangle\,|^2 \sum_n \frac{|\alpha|^{2n}}{n!} \\ &= |\,\langle 0\,|\,\alpha\rangle\,|^2 e^{|\alpha|^2} \quad , \end{aligned} \tag{6.107}$$

where $\langle m\,|\,n\rangle = \delta_{mn}$ and in the last step I have used the Taylor expansion (1.18), and, by setting the arbitrary phase of the complex number $\langle 0\,|\,\alpha\rangle$ equal to zero, we obtain

$$\langle 0\,|\,\alpha\rangle = e^{-\frac{1}{2}|\alpha|^2} \quad , \tag{6.108}$$

which, substituted into Eq. (6.106), allows us to finally derive the canonical form of coherent states

$$\begin{aligned} |\alpha\rangle &= e^{-\frac{|\alpha|^2}{2}} \sum_n \frac{\alpha^n}{\sqrt{n!}}\,|\,n\rangle \\ &= e^{-\frac{|\alpha|^2}{2}} \sum_n \frac{\alpha^n\left(\hat{a}^\dagger\right)^n}{n!}|0\rangle \\ &= e^{-\frac{|\alpha|^2}{2}} e^{\alpha\hat{a}^\dagger}\,|\,0\rangle \quad , \end{aligned} \tag{6.109}$$

where I have used again the Taylor expansion (1.18). In other words, coherent states $|\,\alpha\rangle$ are coherent superpositions of eigenstates of the number operator. Note that it is possible to derive

$$|\,\alpha\rangle = \hat{D}(\alpha)\,|\,0\rangle \quad , \tag{6.110}$$

where

$$\hat{D}(\alpha) = e^{\alpha\hat{a}^\dagger - \alpha^*\hat{a}} \tag{6.111}$$

is known as the unitary *displacement operator*: Eq. (6.110) shows that an arbitrary coherent state can be 'generated' by displacing the vacuum state, i.e. the field state with no photons (Fig. 6.4). Thanks to the Stone theorem (Sect. 1.2.5), it is easy to prove that this operator is unitary.

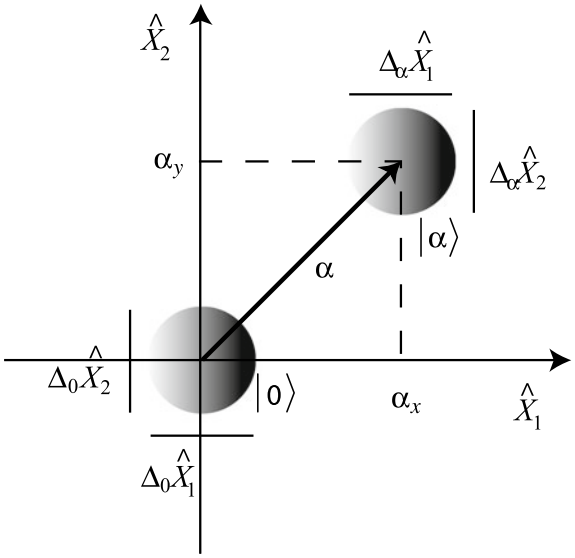


Fig. 6.4 Pictorial representation of the coherent states and of the action of the displacement operator. Here the phase space is determined by the quadrature operators. From Eq. (6.101), it follows that for the vacuum state $|0\rangle$ we have $\left\langle \hat{X}_1 \right\rangle_0 = \left\langle \hat{X}_2 \right\rangle_0 = 0$. Moreover, for this state we have the minimum value of the uncertainties (6.98), $\Delta_0 \hat{X}_1 = \Delta_0 \hat{X}_2 = 1/\sqrt{2}$, and so the vacuum state may be schematically depicted as a circle of radius $1/2\sqrt{2}$ centred around the origin (see Fig. 1.25, Sect. 1.2.4). Similarly, any coherent state may be represented by a circle of radius $1/2\sqrt{2}$ centred around the point $\alpha = \alpha_x + i\alpha_y$ in the complex plane $(\hat{X}_1, \hat{X}_2) \equiv (\Re(\alpha), \Im(\alpha))$. Therefore, the action of the displacement operator $\hat{D}(\alpha)$ may be interpreted as that of displacing the vacuum state from the origin to the point $\alpha = (\alpha_x, \alpha_y)$ in the complex plane, as shown by the arrow. Adapted from Auletta et al. (2009, p. 469)

Wigner Function

Let us first introduce the probability that a classical random variable ξ takes on a value that is less than x.[36] This is given by the distribution function

$$\mathcal{F}(x) = \wp(\xi < x) \ . \tag{6.112}$$

The *characteristic function* of the classical random variable ξ is defined as the expectation of the random variable $e^{i\eta\xi}$, where $\eta \in \mathbb{R}$ is a parameter (see Eq. (1.126)):

$$\chi_\xi(\eta) = \int d\mathcal{F}(x) e^{i\eta x} \ . \tag{6.113}$$

The quantum characteristic function (being the counterpart of the classical one) is expressed as (see Eqs. (1.380) and (5.145))

[36] Gnedenko (1969, Chaps. 4–5).

$$\chi_{\hat{a}^\dagger,\hat{a}}(\eta,\eta^*) = \mathrm{Tr}\left[\hat{\rho}e^{\eta\hat{a}^\dagger-\eta^*\hat{a}}\right] = \left(e^{\eta\hat{a}^\dagger-\eta^*\hat{a}}\,|\hat{\rho}\right), \tag{6.114}$$

where $|\hat{\rho})$ represents the state of the system, and $\eta \in \mathbb{C}$ is a parameter that does not need to be specified here. Clearly, the operator $e^{\eta\hat{a}^\dagger-\eta^*\hat{a}}$, which has the form of the displacement operator (6.111), replaces the classical random variable $e^{i\eta\xi}$. The Wigner distribution function[37] for the state $|\hat{\rho})$ may be written as the Fourier transform (1.176) of the characteristic function $\chi_{\hat{a}^\dagger,\hat{a}}(\eta,\eta^*)$

$$W(\alpha,\alpha^*) = \frac{1}{\pi^2}\int d^2\alpha e^{-\eta\alpha^*+\eta^*\alpha}\chi_{\hat{a}^\dagger,\hat{a}}(\eta,\eta^*)\,, \tag{6.115}$$

Thanks to the formulae (6.68) and (6.114), we can also express the Wigner function as dependent on position and momentum:

$$W(x,p_x) = \frac{1}{\pi h}\int_{\mathbb{R}} dx'\langle x+x'|\hat{\rho}|x-x'\rangle e^{2i\frac{p_x x'}{h}}\,. \tag{6.116}$$

If the Wigner function is integrated with respect to $\hat{p}_x$, it gives the correct probability distribution (marginal distribution) of $\hat{x}$ and *vice versa* (and similarly for quadratures (6.96)). However, due to the presence of quantum features (Sects. 5.1.2–5.1.3), the Wigner function can also assume negative values, as displayed in Fig. 6.5, which is a signature of the fact that it is not a true probability distribution but a *quasi-probability distribution*, as anticipated. What is really important is that the Wigner function (and similar quasi-probability distributions) allows us to deal with intermediate situations between canonical conjugate observables, thus overcoming the dichotomy characterising sharp observables and therefore to describe the area shown in Fig. 2.7b, Sect. 2.2.2.

In other words, the Wigner function is, among others a quasi-probability distribution for classical conjugate variables or, equivalently, a phase space representation of the density operator. It is then clear that measuring the Wigner function is completely equivalent to the measurement of the density operator. Indeed, there are circumstances where a direct measurement of the Wigner function is simply more convenient. The Canadian physicist Antoine Royer[38] analysed the problem in general terms, which can be cast as follows: given a well-defined preparation procedure and a certain number of identical systems, is it possible to determine experimentally (to measure) the state which such a procedure selects? Due to the one-to-one correspondence between the Wigner function and the density matrix of a system, this is possible if one is able to determine the Wigner function. Such a general method is the basis of a wide range of techniques for measuring the state of a quantum system known as quantum *tomography*.

---

[37] Wigner (1932). Note that the Wigner function is only one among a whole family of interrelated quasi-probability distributions.

[38] See Royer (1989).

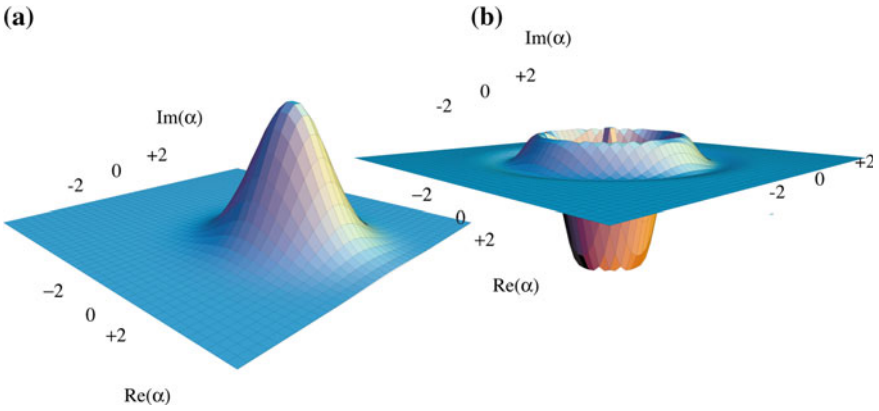


Fig. 6.5 **a** Representation of the Wigner function (6.115) of a coherent state $|\alpha\rangle = |2\rangle$. It is a bidimensional Gaussian centered at the point $\alpha_0 = (2, 0)$ in the complex plane $(\Re(\alpha),\Im(\alpha))$. **b** Representation of the Wigner function of a number state $|n\rangle$ with $n = 4$, i.e. with the number of photons equal to the mean number of photons in the coherent state (**a**). Its annular shape shows the phase invariance of the number state. Note that there are regions where the function becomes negative. Note also that the two figures do not have the same z-scale. Adapted from Auletta et al. (2009, p. 480)

Displaced Parity Operator

Let us limit ourselves to the consideration of a 1D system, whose phase state is represented by position $\hat{x}$ and momentum $\hat{p}_x$ (Fig. 1.25, Sect. 1.2.4). Let us first consider the parity operator about the origin, where I recall that a parity transformation is the flip in the sign of phase space coordinates Eq. (1.313):

$$\hat{\Pi} = \int_{-\infty}^{+\infty} dx |-x\rangle\langle x| = \int_{-\infty}^{+\infty} dp_x |-p_x\rangle\langle p_x| \,, \tag{6.117}$$

where $-\infty < x < +\infty$, $-\infty < p_x < +\infty$. In other words, we have

$$\hat{\Pi}\,|x\rangle \;=\; \int_{-\infty}^{+\infty} dx\,|-x\rangle\;|x\rangle\,\langle x\,| = |-x\rangle \tag{6.118a}$$

$$\hat{\Pi}\,|p_x\rangle \;=\; \int_{-\infty}^{+\infty} dp_x\,|-p_x\rangle\;|p_x\rangle\,\langle p_x\,| = |-p_x\rangle \,, \tag{6.118b}$$

where I have used the completeness relations (1.121) and (1.145). By making use of the displacement operator (6.111) expressed in the position–momentum coordinates

$$\hat{D}_{xp} = e^{\frac{i}{\hbar}(p_x\hat{x} - x\hat{p}_x)} \ , \tag{6.119}$$

where $x,\ p_x \in \mathbb{R}$, we can build now $\hat{\Pi}_{xp}$ as a parity operator about the phase–space point $(x,\ p_x)$:

$$\hat{\Pi}_{xp} = \hat{D}_{xp}\hat{\Pi}\hat{D}_{xp}^{-1} \ , \tag{6.120}$$

where the explicit form of $\hat{\Pi}_{xp}$ is

$$\begin{aligned}
\hat{\Pi}_{xp} &= \frac{\hbar}{2}\int_{-\infty}^{+\infty} dx' e^{ix'p_x} \left| x + \frac{1}{2}\hbar x' \right\rangle \left\langle x - \frac{1}{2}\hbar x' \right| \\
&= \frac{\hbar}{2}\int_{-\infty}^{+\infty} dp'_x e^{ixp'_x} \left| p_x + \frac{1}{2}\hbar p'_x \right\rangle \left\langle p_x - \frac{1}{2}\hbar p'_x \right| \\
&= \frac{\hbar}{4\pi}\int_{-\infty}^{+\infty} dp'_x \int_{-\infty}^{+\infty} dx' e^{ip'_x(\hat{x}-x) - ix'(\hat{p}_x - p_x)} \ ,
\end{aligned} \tag{6.121}$$

where the reader may verify the validity of the last step. Note that

$$\left| x + \frac{1}{2}\hbar x' \right\rangle \text{ and } \left| p_x + \frac{1}{2}\hbar p'_x \right\rangle \tag{6.122}$$

are the 'displaced' eigenkets of position and momentum, respectively. This is convenient when we consider the form (6.116) of the Wigner function. Indeed, from Eqs. (6.117)–(6.120), it follows that

$$\hat{\Pi}_{xp}(\hat{x} - x)\hat{\Pi}_{xp} = -(\hat{x} - x), \ \ \hat{\Pi}_{xp}(\hat{p}_x - p_x)\hat{\Pi}_{xp} = -(\hat{p}_x - p_x) \ . \tag{6.123}$$

Due to the form of $\hat{\Pi}_{xp}$, these states are clearly eigenstates of an unsharp observable (Sect. 5.2.4).

Displaced Kets

The crucial point is that the Wigner function (6.116) turns out to be the expectation value of the parity operator $\hat{\Pi}_{xp}$ on some input state $\left|\hat{\rho}\right)$ at time t

$$W_{\hat{\rho}}(x,\ p_x,\ t) = \left(\hat{\Pi}_{xp}\left|\hat{\rho}(t)\right.\right) = \frac{1}{\pi\hbar}\left\langle\hat{\Pi}_{xp}\right\rangle_{\hat{\rho}(t)} \ . \tag{6.124}$$

Since obviously $\hat{\Pi}_{xp}^2 = \hat{I}$ (two spatial reflections bring back to the input state), $\hat{\Pi}_{xp}$ is a (POVM) observable whose eigenvalues are ± 1. Consider an arbitrary complete orthogonal set, spanning the n-dimensional Hilbert space of the system, of kets $|\psi^{(n)}\rangle$ of definite parity about the origin, so that

$$\psi^{(n)}(-x) = (-1)^n \psi^{(n)}(x) \ . \tag{6.125}$$

Then, another complete set of eigenstates $|\psi_{xp}^{(n)}\rangle$, $n = 1, 2, \ldots$, satisfying, in agreement with Eqs. (6.118), the eigenvalue equation

$$\hat{\Pi}_{xp}|\psi_{xp}^{(n)}\rangle = (-1)^n |\psi_{xp}^{(n)}\rangle \ , \tag{6.126}$$

may be obtained by displacing in phase space the kets $|\psi^{(n)}\rangle$:

$$|\psi_{xp}^{(n)}\rangle = \hat{D}_{xp}|\psi^{(n)}\rangle \ , \tag{6.127}$$

where, in analogy with Eq. (6.121), we have

$$\hat{\Pi}_{xp} = \sum_n (-1)^n |\psi_{xp}^{(n)}\rangle\langle\psi_{xp}^{(n)}| \ . \tag{6.128}$$

We have now expressed the parity operator as a discrete sum, so that, taking into account Eqs. (6.128) and (6.124) can be rewritten as

$$W_{\hat{\rho}}(x, p_x, t) = \frac{1}{\pi\hbar} \sum_n (-1)^n \langle\psi_{xp}^{(n)}|\hat{\rho}(t)|\psi_{xp}^{(n)}\rangle \ . \tag{6.129}$$

We try now to measure $W_{\hat{\rho}}(x, p_x, t)$ at some definite time (e.g. $t = 0$). This can be done by measuring each mean value $\langle\psi_{xp}^{(n)}|\hat{\rho}(0)|\psi_{xp}^{(n)}\rangle$.

A Simple Method

A simple approach is possible if we choose the $|\psi^{(n)}\rangle$'s to be eigenstates of the Hamiltonian (Eq. (1.12))

$$\hat{H} = \frac{\hat{p}_x^2}{2m} + V(\hat{x}) \ , \tag{6.130}$$

where $V(-x) = V(x)$ is a symmetric potential. Then, the displaced kets $|\psi_{xp}^{n}\rangle$'s are eigenstates of the displaced Hamiltonian

$$\hat{H}_{xp} = \hat{D}_{xp}\hat{H}\hat{D}_{xp}^{-1} = \frac{(\hat{p}_x - p_x)^2}{2m} + V(\hat{x} - x) \ , \tag{6.131}$$

so that measuring the set $\langle\psi_{xp}^{(n)}|\hat{\rho}(0)|\psi_{xp}^{(n)}\rangle$ (or $\hat{\Pi}_{xp}$) becomes equivalent to measure the Hamiltonian $\hat{H}_{xp}$. A suitable method to measure $\hat{H}_{xp}$ almost in the strict sense is as follows: first, we place ourselves in a (primed) reference frame moving with uniform speed $v = p_x'/m$ relative to the preparation apparatus $\mathcal{A}$ (Sect. 2.3.2). By virtue of the Galilei transformations (2.62a), the observed density operator (for $t \leq 0$) in the moving frame is displaced as

$$\hat{\rho}'(t) = \hat{D}_{vt,p_x'}^{-1}\hat{\rho}(t)\hat{D}_{vt,p_x'} \ . \tag{6.132}$$

At time $t = 0$, we turn on the potential $V(x - x')$ in the moving frame. The eigenstates of

$$\hat{H}_{x',0} = \frac{\left(\hat{p}'_x\right)^2}{2m} + V(\hat{x}' - x') \tag{6.133}$$

are also displaced relative to the energy eigenstates of the Hamiltonian (6.130)

$$\left|\psi^{(n)}_{x',0}\right\rangle = \hat{D}_{x',0}\left|\psi^{(n)}\right\rangle \ , \tag{6.134}$$

with corresponding energies E_n (see Eq. (1.197)). Then, at times $t \geq 0$ we obtain

$$\begin{aligned}\hat{\rho}'(t) &= e^{-\frac{i}{\hbar}t\hat{H}_{x',0}}\hat{\rho}'(0)e^{\frac{i}{\hbar}t\hat{H}_{x',0}} \\ &= \sum_{m,n} e^{-\frac{i}{\hbar}(E_n - E_m)t}|\psi^{(n)}_{x',0}\rangle\langle\psi^{(m)}_{x',0}|\langle\psi^{(n)}_{x',0}|\hat{\rho}'(0)|\psi^{(m)}_{x',0}\rangle \\ &= \sum_{m,n} e^{-\frac{i}{\hbar}(E_n - E_m)t}|\psi^{(n)}_{x',0}\rangle\langle\psi^{(m)}_{x',0}|\langle\psi^{(n)}_{x',p'_x}|\hat{\rho}(0)|\psi^{(m)}_{x',p'_x}\rangle \ ,\end{aligned} \tag{6.135}$$

where in the last step I have made use of the transformations given by Eq. (6.132) for $\hat{\rho}(0)$. Now, the transition probabilities (which account for shifts in time of the energy's eigenvalues)

$$\langle\psi^{(n)}_{x',0}|\hat{\rho}'(0)|\psi^{(n)}_{x',0}\rangle = \langle\psi^{(n)}_{x'p'_x}|\hat{\rho}(0)|\psi^{(n)}_{x'p'_x}\rangle \tag{6.136}$$

are in fact time independent, so that we have a long time available to perform a measurement of $\hat{H}_{x',0}$ referring to the set $\{|\psi^{(n)}_{x',0}\rangle\}$ and 'find' the particle in one of the states pertaining to this set. Repeating the measurement many times will allow to build the distribution (6.136), from which $W_{\hat{\rho}}(x',\, p'_x,\, 0)$ can be deduced by means of Eq. (6.129). What has been done is a measurement of $W_{\hat{\rho}}(x',\, p'_x,\, 0)$ by measuring

$$W_{\hat{\rho}'}(x',\, 0,\, t) = W_{\hat{\rho}}(x' + vt,\, p'_x,\, t) \tag{6.137}$$

at $t = 0$ in the moving frame. In conclusion, applying the same procedure over and over again with different values x' and p'_x, it is in principle possible to reconstruct the Wigner function on any relevant region of the phase space.

Having established a procedure for showing that unsharp observables or quasi-probability functions can be informationally complete, let us mention that tomographic procedures apply also to effects. In fact, a set of effects that is separating for states is informationally complete for states. However, since according to Eqs. (5.70), also states separate for effects, it is evident that we can build a tomographic procedure for effects.[39]

[39] See D'Ariano et al. (2017, Sects. 2.4.2 and 2.8.4).

6.2.4 *Green's Functions, Path Integrals, Quantum Trajectories*

Since when we exchange signals, an irreversible aspect is always present apart from the case in which no information is gained, it is opportune to come back here to the notion of operation and generalised dynamics that it has been discussed in Sect. 5.2.5. In particular, I shall introduce two different ways to deal with this problem: the path-integral approach and the quantum-trajectories approach.

Green's Functions

The so-called Green's functions, after the name of the British mathematician George Green (1793–1841), constitute a mathematical object that, in its applications to physics, take particular aspects.[40] In this case, it is essentially a *propagator*, that is, it describes the spread of a signal out of a source. Let us start by the consideration that the elements of the time-translation unitary matrix can be written as[41] (Sect. 1.2.5)

$$\left\langle k \left| e^{\frac{i}{\hbar}\hat{H}(t_0)} e^{-\frac{i}{\hbar}\hat{H}(t)} \right| j \right\rangle = \left\langle k \left| e^{-\frac{i}{\hbar}\hat{H}(t-t_0)} \right| j \right\rangle = i\mathrm{G}(k,t;j,t_0) \ , \tag{6.138}$$

where $|\, j\rangle$ at some initial time t_0 and $|\, k\rangle$ at some later time t are some state vectors describing a quantum system. The functions G are the *Green's functions*. In order to appreciate their importance, let us consider following equation:

$$\left|\psi(t')\right\rangle = e^{-\frac{i}{\hbar}\hat{H}(t'-t)} \,|\psi(t)\rangle \ , \tag{6.139}$$

which relates the state vector at time t' to the state vector at time t. If we multiply both sides of Eq. (6.139) times the 3D eigenstates of position $\left\langle \mathbf{r}' \right|$ from the left and make use of the resolution of the identity in the form $\hat{I} = \int d\mathbf{r}\,|\mathbf{r}\rangle\,\langle\mathbf{r}\,|$, we obtain, for $\psi(\mathbf{r}',t') = \left\langle \mathbf{r}' \mid \psi(t)\right\rangle$,

$$\begin{aligned} \psi(\mathbf{r}',t') &= \int d\mathbf{r} \left\langle \mathbf{r}' \right| e^{-\frac{i}{\hbar}\hat{H}(t'-t)} \,|\mathbf{r}\rangle\, \langle \mathbf{r} \mid \psi(t)\rangle \\ &= i \int d\mathbf{r} \mathrm{G}(\mathbf{r}',t';\mathbf{r},t)\psi(\mathbf{r},t) \ , \end{aligned} \tag{6.140}$$

where I have made use of Eq. (1.157). Equation (6.140) represents an instance of Huygens' principle, after the name of the Dutch mathematician and physicist Christiaan Huygens (1629–1695): if the wave function $\psi(\mathbf{r},t)$ is known at a time t, it may be found at any later time t' by assuming that each point $\mathbf{r}$ at time t is a source of waves which propagate outward from $\mathbf{r}$. Thus, the previous equation tells us that the strength of the wave amplitude arriving at point $\mathbf{r}'$ at time t' from the point $\mathbf{r}$ will

[40] Byron and Fuller (1969–70, II, Chap. 7).

[41] See Auletta et al. (2009, Sect. 3.5.5).

be proportional to the original wave amplitude $\psi(\mathbf{r}, t)$ and the constant of proportionality is given by $i\mathrm{G}(\mathbf{r}', t'; \mathbf{r}, t)$. Note that Eq. (6.140) is a consequence of the first-order character of the Schrödinger equation and of its linearity, since the knowledge of $\psi(\mathbf{r}, t)$ for all values of $\mathbf{r}$ and one particular time t is enough to determine $\psi(\mathbf{r}', t')$ for all values of $\mathbf{r}'$ and any (subsequent or previous) time t', and the relation between the two wave functions is linear.

Using Green's functions, it is possible to write the evolution of a free particle in space and time in the form[42]

$$\psi(\mathbf{r}', t') = i \int d\mathbf{r} \mathrm{G}_0(\mathbf{r}', t'; \mathbf{r}, t)\psi(\mathbf{r}, t) \ , \tag{6.141}$$

for $t' > t$. The expression G_0 is called the *free Green function* and its explicit form can be proved to be

$$\mathrm{G}_0(\mathbf{r}', t'; \mathbf{r}, t) = -i \left[\frac{m}{2\pi i \hbar (t' - t)} \right]^{\frac{3}{2}} e^{\frac{im|\mathbf{r}' - \mathbf{r}|^2}{2\hbar(t' - t)}} \ . \tag{6.142}$$

Green's Functions and Action

The probability for a 1D quantum system $\mathcal{S}$, which is in an initial position eigenstate $|\, x(t_0)\rangle$ at time t_0, to be found in a state $|\, x(t)\rangle$ at time t is given by $|\langle x(t) \mid x(t_0)\rangle|^2$ and its relative amplitude by the Green's function[43]

$$\mathrm{G}(x(t), t; x(t_0), t_0) = \langle x(t) \mid x(t_0)\rangle = e^{\frac{i}{\hbar} S(x(t), t; x(t_0), t_0)} \ , \tag{6.143}$$

where the time unitary evolution operator has been incorporated in the time dependence of both the input and output states. One can prove that in the limit in which $\hbar \to 0$ (and we reach the classical regime: Sect. 1.3.2) the function $S(x(t), t; x(t_0), t_0)$ will coincide with the classical action evaluated along the classical trajectory going from $x(t_0)$ to $x(t)$ (see Eq. (1.23)). Then, Eq. (1.52) tells us that the Hamiltonian of the corresponding classical system is

$$H^{\mathrm{c}}\left(x(t_0), p_x(t_0)\right) = -\frac{\partial S}{\partial t_0} \ , \tag{6.144}$$

which is formally identical to the quantum Hamiltonian, and, in agreement with Eq. (1.50), we also have

$$p_x(t_0) = -\frac{\partial S}{\partial x(t_0)} \ . \tag{6.145}$$

[42] Further details regarding the subject of this subsection can be found in Bjorken and Drell (1964, 78–89) Prugovečki (1971, 520–42).

[43] See Dirac (1930, 125–30). For what follows see also Auletta et al. (2009, Sect. 10.5.2).

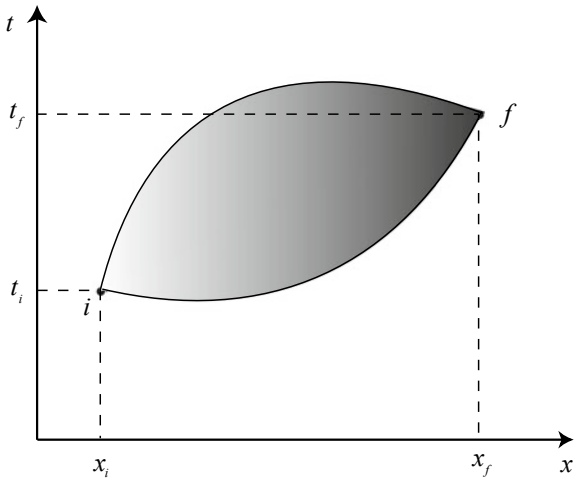


Fig. 6.6 A pictorial representation of a bundle of trajectories in 1D configuration space. Of course, all the (infinite) possible paths connecting the initial position i to the final f contribute to the probability amplitude of moving from i to f. Adapted from Auletta et al. (2009, p. 390)

To obtain the quantum analogue of the classical Lagrangian, we must consider an infinitesimal time interval $t = t_0 + \delta t$, allowing us to write the amplitude $\langle x(t_0 + \delta t) \mid x(t_0)\rangle$ as the analogue of $e^{\frac{i}{\hbar}L(t_0)\delta t}$. In such a case, one should consider $L(t_0)$ as a function of the coordinate $x(t_0 + \delta t)$ at time $t_0 + \delta t$ and of the coordinate $x(t_0)$ at time t_0 rather than as a function of position and momentum, as is usually assumed. Thus, the quantum amplitude (6.143) may be written in the limit of a small $\hbar$ as

$$\langle x(t) \mid x(t_0)\rangle = e^{\frac{i}{\hbar}S(x(t),t;x(t_0),t_0)} = e^{\frac{i}{\hbar}\int_{t_0}^{t} dt' L(t')} \ . \tag{6.146}$$

Note that in the classical limit we recover the Hamilton–Jacobi Equation (3.108).

Path Integrals

The idea of path integrals has been already mentioned in Sect. 2.2.4. It is Richard Feynman who introduced this method in quantum physics.[44] Let us consider the evolution of a quantum system from a certain initial point $i = (x_i, t_i)$ of the configuration space to a certain final point $f = (x_f, t_f)$. As we know, each possible trajectory (Fig. 2.7, Sect. 2.2.2) contributes with a different phase to the total probability amplitude of the transition $i \mapsto f$ (Fig. 6.6). According to Eq. (6.143), the probability $\wp(f, i)$ is the absolute square of the Green's function $\mathrm{G}(f, i)$. Note that I have omitted here the imaginary unit. This has obviously no consequence on the absolute square of the Green's functions, even though it yields to a slightly different form of Eq. (6.140). The total probability amplitude may then be interpreted as the sum of the contributions (amplitudes) $\vartheta[x(t)]$ of each possible path connecting i and f in the configuration space. Thus, in the 1D case, the Green's function may be written as

$$\mathrm{G}(f, i) = \sum \vartheta\,[x(t)] \ , \tag{6.147}$$

[44] Feynman (1948), Feynman and Hibbs (1965). See also Auletta et al. (2009, Sect. 10.8).

where the sum is taken over all possible paths. On the most general grounds, we can assume that each path from i to f has equal probability, and therefore contributes by an equal amount to the total probability amplitude. In agreement with Eq. (6.143), the contribution $\vartheta[x(t)]$ of each path has a phase proportional to the action S

$$\vartheta[x(t)] = Ce^{\frac{i}{\hbar}S[x(t)]} , \tag{6.148}$$

where C is a constant. Since the number of paths is certainly infinite, it is natural to ask how to assign the correct measure to the space of these paths. To this purpose, we can use the properties of the Riemann integral, after the name of Bernhard Riemann, for solving our problem. We can choose a subset of the paths (Fig. 6.7) and divide the independent time variable into small steps (slices) of width ϵ. At each time t_j we select some point x_j (from the starting point x_i to arrival point x_f). We build then a path by connecting all positions that have been found by connecting all the points (x_j, t_j) by straight lines. Given that we have n time intervals of length ϵ, i.e. $n\epsilon = t_f - t_i$, we may define a sum over all paths by taking a multiple integral over all possible values of x_j at each node

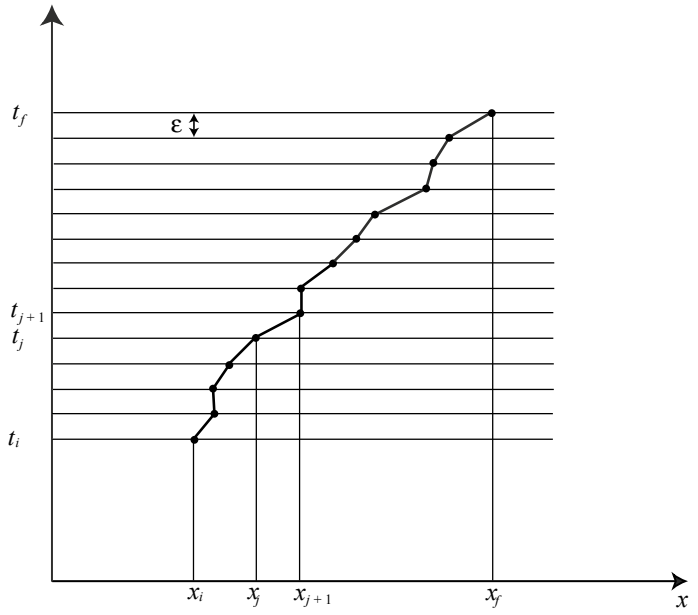


Fig. 6.7 The sum over paths is defined as a limit in which at first the path is specified by the coordinates at a large number of times separated by very small intervals ϵ. The sum over the paths is then an integral over all possible values of these intermediate coordinates. In order to achieve the correct measure, the limit is taken for $\epsilon \to 0$. Adapted from Auletta et al. (2009, p. 392)

$$\mathrm{G}(x_n, t_n; x_0, t_0) \approx \int\int \cdots \int dx_1 dx_2 \cdots dx_{n-1} \vartheta[x(t)] \,, \tag{6.149}$$

where $t_0 = t_i$, $t_n = t_f$ and $x_0 = x_i$, $x_n = x_f$. Of course, there is no need to integrate over x_0 and x_n, because these points are known and fixed. In the general case, it is very difficult to compute the correct normalisation factor, but, in all cases where the action can be derived by integrating the Lagrangian Eq. (1.25)

$$L(\dot{x}, x, t) = \frac{m}{2}\dot{x}^2 - V(x, t) \,, \tag{6.150}$$

it is possible to prove that the normalisation factor is given by $\mathcal{N}^{-n}$ where

$$\mathcal{N} = \left(\frac{2\pi i \hbar \epsilon}{m}\right)^{\frac{1}{2}} \,. \tag{6.151}$$

With this factor the limit exists and it produces the correct value of the Green's function $\mathrm{G}(f, i)$. Hence, as a generalisation of Eq. (6.146), we can write

$$\mathrm{G}(x_n, t_n; x_0, t_0) = \lim_{\epsilon \to 0} \frac{1}{\mathcal{N}} \int\int \cdots \int \frac{dx_1}{\mathcal{N}} \frac{dx_2}{\mathcal{N}} \cdots \frac{dx_{n-1}}{\mathcal{N}} e^{\frac{i}{\hbar} S(f,i)} \,, \tag{6.152}$$

where, in agreement with Eq. (1.23), we have

$$S(f, i) = \int_{t_0}^{t_n} dt L(\dot{x}, x, t) \,, \tag{6.153}$$

and the integral is computed using a trajectory $x(t)$ that is piecewise linear and passes through the points (x_j, t_j).

Given Eq. (6.150), the action S for each trajectory is given by

$$S = \epsilon \sum_j \left[\frac{\left(x_j - x_{j-1}\right)^2}{2\epsilon^2} m\right] - \int_0^1 du V\left[(1-u)x_j + u x_{j-1}\right] \,, \tag{6.154}$$

where u parametrises the trajectory between x_{j-1} and x_j. Assuming the trajectory to be continuous, when ϵ goes to zero also the difference $x_j - x_{j-1}$ does, and the action remains finite. In this case, neglecting terms that vanish when ϵ goes to 0, the integral in the previous formula may be approximated by $V(x_j)$ or, equivalently, by $\frac{1}{2}\left(V(x_j) + V(x_{j-1})\right)$.

There are also other ways to define a subset of all paths between i and f. However, the concept of the sum (or integration) over all paths is rather general, and, independently from the method we use for defining the integral, it is usually written in the following notation:

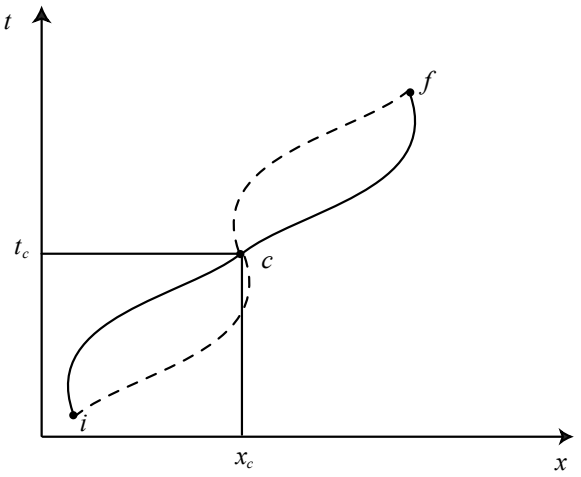


Fig. 6.8 Two possible paths from i to f both passing through the same central point c. Adapted from Auletta et al. (2009, p. 394)

$$\mathrm{G}(f,i) = \int_i^f d[x(t)] e^{\frac{i}{\hbar} S(f,i)} \,, \tag{6.155}$$

and is called *path integral*, with the meaning given to it by expression (6.152).

Sequence of Events

Suppose now that we have two events in succession (say, a particle moving first from location x_i at time t_i to location x_c at time t_c, and then to location x_f at time t_f), such that $S(f,i) = S(f,c) + S(c,i)$, where $c = (x_c, t_c)$. The action is an integral in time and the Lagrangian does not depend on derivatives higher than the velocity. Making use of Eq. (6.155) we can write

$$\mathrm{G}(x_n, t_n; x_0, t_0) = \int d[x(t)] e^{\frac{i}{\hbar}[S(f,c)+S(c,i)]} \,. \tag{6.156}$$

Thus, it is possible to split any path into two parts (Fig. 6.8), allowing us to write

$$\mathrm{G}(x_n, t_n; x_0, t_0) = \int dx_c \int_c^f d[x(t)] e^{\frac{i}{\hbar} S(f,c)} \mathrm{G}(x_c, t_c; x_0, t_0) \,, \tag{6.157}$$

where the integration is now performed not only over all possible paths from c to f, but also over the variable central point x_c. Then, we carry out the integration over paths between x_0 and an arbitrary x_c and between x_c and x_n, i.e.

$$\mathrm{G}(x_n, t_n; x_0, t_0) = \int dx_c \mathrm{G}(x_n, t_n; x_c, t_c) \mathrm{G}(x_c, t_c; x_0, t_0) \,. \tag{6.158}$$

This result can be summarised as follows: All alternative paths going from i to f can be labelled by specifying the position x_c through which they pass at time t_c. Then, the integral over the paths, or the *kernel* $\mathrm{G}(x_n, t_n; x_0, t_0)$ for a particle going from the point i to the point f, can be computed according to the following rules:

- The kernel from i to f is the sum over all possible values of c of the amplitudes for the particle to go from i to f passing through c.
- The probability amplitude for a particle to go from i to c and then to f is given by the product of the kernel from i to c times the kernel from c to f.

In other words, we have a sequence event–propagation–event, where the events can be reception or emission of a signal (due to the role played by action) that propagates according to the multipath dynamics of QM.

In the classical limit of the path integral approach, represented by the limit for the action being much larger than h (Sect. 1.3.2), only one trajectory exists, precisely the one that minimises the action. In the classical limit, the phase $\mathrm{S}/\hbar$ is very large: small changes at the classical scale in the path will produce large variations in the phase contributions, making it a rapidly oscillating function. As a consequence, the total contribution resulting from the paths that are far away from the classical path will add to zero. Instead, in the vicinity of the classical path—where the action is at a minimum—small variations in the path itself give rise to no change in the action in the first order, according to Eq. (1.26), and the paths in that region will be in phase and give rise to a non-zero net contribution. In conclusion, in the classical limit the only path that needs to be considered is precisely that minimising the action, according to the principle of least action (1.23). This makes us understand both the relation and the difference between multipath quantum dynamics and single-path classical dynamics.

The Concept of Quantum Jump

Let us come back to the connection between the amplitude-operators formalism and the master equation and try to generalise the notion of jump, which I have already introduced in a simplified form in Sect. 6.2.2. I have mentioned that the former is more general than the master-equation formalism (Sect. 5.2.4). Let us now show this in detail. The following development is deeply connected with Feynman's path integral method, as far as it considers the whole dynamical evolution of a system as alternating local collapses with global evolution. This approach takes its roots from an analogy with classical statistical physics, where one may describe a system as an ensemble of stochastic trajectories, each of which is generated by a set of stochastic differential equations. In the QM of open systems, quasi-probability distributions may be used in the place of the classical probability distributions (Sect. 6.2.3).

Let us then suppose for the time being to be able to monitor in a perfect way an open quantum system.[45] Even though we know that this is not physically possible, this procedure will help us in the derivation of the new formalism of *quantum trajectories* (which are obviously different relative to classical trajectories). If, in controlled

[45] I follow here Auletta et al. (2009, Sect. 14.4).

situation, our monitoring is perfect, then we are able to detect any quanta lost by the system. For example, in the case of the electromagnetic field inside a cavity, we would be able to detect any single photon lost by the cavity. We can then record the times at which the quanta are released. Therefore, we may assume that between two successive detections the system evolves in a continuous (without emission of quanta) and, at the limit, reversible way. In this hypothetical, perfectly monitored, quantum trajectory, the dynamics of the system would then consist of a succession of continuous evolutions and discrete emissions of quanta (each one representing a quantum event), which, as we know, are called *quantum jumps*. In spite of the quantum character of these trajectories, the evolved density operator at a certain time can be determined as the weighted ensemble average of all possible quantum trajectories.

Jumps Alternating with Free Evolution

Two ingredients are necessary for the following procedure: first, we need two types of operators ruling each one of the two types of evolution, one without jumps and the other with a jump; second, we need the specific set of times for the jumps. Assume that an open quantum system is described by the reduced density matrix $\hat{\varrho}$ and its dynamics is ruled by the master Equation (5.83)

$$\frac{d}{dt}\left|\,\hat{\varrho}(t)\right) = \hat{\mathcal{L}}\left|\,\hat{\varrho}(t)\right) \ , \tag{6.159}$$

where (see also Eq. (5.84))

$$\hat{\mathcal{L}}\left|\,\hat{\varrho}\right) = \left(\hat{\mathcal{L}}_{\mathrm{d}} + \hat{\mathcal{L}}_{\mathrm{nd}}\right)\left|\,\hat{\varrho}\right) \tag{6.160}$$

is the Lindblad evolution superoperator acting on both the system and its environment. The differential Eq. (6.159) can be formally solved as

$$\left|\,\hat{\varrho}(t)\right) = e^{\hat{\mathcal{L}}t}\left|\,\hat{\varrho}(t_0)\right) \ . \tag{6.161}$$

The jump operator, i.e. the superoperator describing the loss of a quantum by the system (see also Eq. (5.102b)), can be defined as

$$\hat{\mathcal{J}}\left|\,\hat{\varrho}\right) = \hat{a}\hat{\rho}\hat{a}^{\dagger} \ , \tag{6.162}$$

where the annihilation and creation operators are given by Eqs. (6.72).[46] We may then add and subtract the operator $\hat{\mathcal{J}}$ to $\hat{\mathcal{L}}$ in the exponent in the RHS of Eq. (6.161) to obtain

$$\left|\,\hat{\varrho}(t)\right) = e^{[(\hat{\mathcal{L}}-\hat{\mathcal{J}})+\hat{\mathcal{J}}]t}\left|\,\hat{\varrho}(t_0)\right) \ , \tag{6.163}$$

[46]Truly speaking, the choice of $\hat{\mathcal{J}}$ is not unique and may well depend on how the system is thought to be monitored. This remark, however, does not alter the essence of the following argument.

in order to split the action of the two superoperators. Let us now introduce some mathematical considerations. We shall try to make a Taylor expansion of the exponential of the previous equation. An exponential of the form $e^{(\hat{O}+\eta\hat{O}')\delta)\xi}$ can be rewritten as

$$(e^{(\hat{O}+\eta\hat{O}')\delta})^{\frac{\xi}{\delta}} \text{ with } \delta \neq 0 \,. \tag{6.164}$$

We make the expansion around $\eta = 0$. Taking the limit for $\delta \to 0$ and making use of the approximation (1.19), we can write

$$\lim_{\delta\to 0}(e^{(\hat{O}+\eta\hat{O}')\delta})^{\frac{\xi}{\delta}} = \lim_{\delta\to 0}(1+\delta(\hat{O}+\eta\hat{O}'))^{\frac{\xi}{\delta}} \,. \tag{6.165}$$

Without loss of generality we can evaluate the limit $\delta \to 0$ when $\zeta = \xi/\delta$ is an integer. Using this definition of ζ we can rewrite the previous expression as

$$\lim_{\delta\to 0}\prod_{j=1,\zeta}(1+\delta(\hat{O}+\eta\hat{O}')) \,. \tag{6.166}$$

We can now compute the expansion in powers of η of the previous formula: it is the product of terms linear in η. Let us compute the coefficient of the order η^2. The term η^2 may come from both the jth and the lth terms of the three-product (where all the other factors give a contribution equal to $1/\delta\hat{O}$):

$$\begin{aligned}\delta^2 \sum_{j,l=1,\zeta;\, j<l} \left(\left(\prod_{a=1}^{j-1}(1+\delta\hat{O})\right)\hat{O}'\left(\prod_{b=j+1}^{l-1}(1+\delta\hat{O})\right)\hat{O}'\left(\prod_{c=l+1}^{\zeta}(1+\delta\hat{O})\right)\right) \\ = \delta^2 \sum_{j,l=1,\zeta;\, j<l}(1+\delta\hat{O})^{j-1}\hat{O}'(1+\delta\hat{O})^{l-j-2}\hat{O}'(1+\delta\hat{O})^{\zeta-l-1}\end{aligned} \tag{6.167}$$

where the jth and the lth terms are represented by the two occurrences of $\hat{O}'$. In the limit $\delta \to 0$ each individual term in the sum over j and l is irrelevant and we can assume that both j and l are of order δ^{-1}. We can thus substitute the sums with integrals; neglecting terms going to zero with δ we obtain

$$\int_0^\xi d\xi_1 \int_{\xi_1}^\xi d\xi_2 (1+\delta\hat{O})^{\frac{\xi_1}{\delta}}\hat{O}'(1+\delta\hat{O})^{\frac{\xi_2-\xi_1}{\delta}}\hat{O}'(1+\delta\hat{O})^{\frac{\xi-\xi_2}{\delta}} \,. \tag{6.168}$$

We can now perform the limit $\delta \to 0$ and obtain

$$\int_0^\xi d\xi_1 \int_{\xi_1}^\xi d\xi_2 \, e^{\xi_1\hat{O}}\hat{O}'e^{(\xi_2-\xi_1)\hat{O}}\hat{O}'e^{(\xi-\xi_2)\hat{O}} \,. \tag{6.169}$$

This can be generalised to higher (and lower) values k of ξ_k. If we limit ourselves to convergent series, we are finally authorised to rewrite Eq. (6.163) as

$$\begin{aligned}\left|\hat{\varrho}(t)\right\rangle\!\!\rangle &= e^{[(\hat{\mathcal{L}}-\hat{\mathcal{J}})+\hat{\mathcal{J}}]t}\left|\hat{\varrho}(t_0)\right\rangle\!\!\rangle \\ &= \sum_{n=0}^{\infty}\int_0^t dt_n \int_0^{t_n} dt_{n-1}\cdots\int_0^{t_2} dt_1 e^{(\hat{\mathcal{L}}-\hat{\mathcal{J}})(t-t_n)}\hat{\mathcal{J}}e^{(\hat{\mathcal{L}}-\hat{\mathcal{J}})(t_n-t_{n-1})}\hat{\mathcal{J}}\cdots \\ &\quad \hat{\mathcal{J}}e^{(\hat{\mathcal{L}}-\hat{\mathcal{J}})(t_1-t_0)}\left|\hat{\varrho}(t_0)\right\rangle\!\!\rangle \ . \end{aligned} \tag{6.170}$$

The expression to be integrated, i.e.

$$e^{(\hat{\mathcal{L}}-\hat{\mathcal{J}})(t-t_n)}\hat{\mathcal{J}}e^{(\hat{\mathcal{L}}-\hat{\mathcal{J}})(t_n-t_{n-1})}\hat{\mathcal{J}}\cdots\hat{\mathcal{J}}e^{(\hat{\mathcal{L}}-\hat{\mathcal{J}})(t_1-t_0)}\left|\hat{\varrho}(t_0)\right\rangle\!\!\rangle = \left|\hat{\tilde{\varrho}}(t)\right\rangle\!\!\rangle \tag{6.171}$$

can be considered as the unnormalised conditioned density operator of the system under observation. Moving from right to left, this term can be interpreted in the following manner: the initial density operator evolves in the time interval between $t_0 \leq t < t_1$ (when there are no loss of quanta) under the propagator

$$e^{(\hat{\mathcal{L}}-\hat{\mathcal{J}})(t_1-t_0)}, \tag{6.172}$$

then jumps under the action of $\hat{\mathcal{J}}$ at the time of the first emission (t_1), evolves during the next interval ($t_1 \leq t < t_2$) under the propagator

$$e^{(\hat{\mathcal{L}}-\hat{\mathcal{J}})(t_2-t_1)}, \tag{6.173}$$

jumps again at t_2 under the action of $\hat{\mathcal{J}}$, and so on. In this way, we are building a step-wise trajectory for the conditioned density operator, in full correspondence to the formalism of path integrals.

Final Considerations

The formalism presented is particularly interesting since it shows that we can conceive the generalised quantum dynamics as a sequence of irreversible event, reversible free multipath evolution followed by another irreversible event, followed by reversible multipath evolution, and so on. In particular, path integrals depict very well the dynamics of (subsequent) particle scattering(s). In fact, we can build different Green's functions $G_0, G_1, G_2, \ldots$ according to whether a particle has been not scattered, scattered once, two, and so on, between an initial and final point. Thus, propagators like the Green functions (and relative path integrals) well describe the quantum-classical dynamics of signals exchanged by quantum systems. In fact, the propagation is a multipath one, according to QM, but it has always a *starting* and an *arrival* point, which are events and therefore involve a classical contribution making a bit available (Fig. 6.9). Indeed, as long as we remain confined to the pure quantum-mechanical side, we cannot speak of starting and arrival points. I have often said that

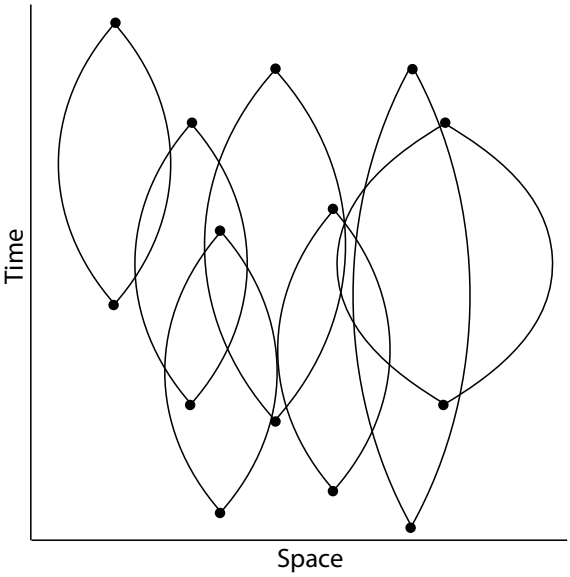


Fig. 6.9 Graphic representation of how all what "has happened" in the past is influenced by events happening in the present as to what can be observed. The upper tip of each "leaf" could stand for the act of registration. The lower end of each "leaf" could stand for source event. As is shown, there is not a single path leading to a single event. Adapted from Wheeler (1983)

(free) quantum systems occupy the whole space that they can occupy (Sects. 3.2.3 and 6.1.5).

Moreover, the formalism of quantum trajectories provides us with a decomposition of the quantum dynamics into an infinite number of quantum trajectories, which are generated by the times at which jumps happen, and between which jumps, although watched for, do not occur. It is important to note that, although a jump means in the paradigmatic case the collapse on a single eigenstate of the measured observable, such an eigenstate (supposed that the system survives the collapse) is fully able to give rise to a new multipath dynamics since it does not loose the general characters of the quantum superposition of possibilities, best expressed by the notion of self-correlation (see also Sects. 5.1.3). In conclusion, in the limit in which dissipative and discontinuous processes tend to zero, we recover the continuous and deterministic dynamics of a free system.

6.2.5 Information Sources, Causality, Reversibility

Quantum Events as Sources

Supposing that signals follow the 'alternate' dynamics described in Sect. 6.2.4, the natural question arises about the relation between events and quantum dynamics in the propagation process of the signal. In which way an initial event contributes to the further evolution and especially to the subsequent event which could well be a signal reception? For answering this question, let us consider the relation between events

and causes in the most general terms. In fact, at least classically, they are identified, as we have seen in Sect. 2.4.2, since events (understood here in a relativistic sense) are the sources of signals and therefore also of causal interconnections (Sect. 2.3.2). The first point to be stressed is that any information source is in fact a source of signals, since it must determine some difference in the surroundings that make information acquisition possible at all: I insist that to speak of an information source from which we cannot acquire information makes no sense (Sect. 5.1.1).

Up to this point, there is general agreement. Now, one of the biggest confusions here is to consider a source of signals not only as a source of causal relations (through the effects that the signal produces on the surroundings) but also as a cause in the ordinary (efficient) sense of the word, namely as something that generates as such a certain output (like the reception event): it is a source determinism.[47] However, quantum events, although giving rise to signals, cannot be treated as causes in the ordinary sense. In the fortunate case in which, for example a particle is scattered few times without being absorbed before reaching the target it seems that we could keep some amount of information contained in the original signal. However, we should distinguish between Markovian processes and non-Markovian ones. A Markovian dynamical sequence, after the name of the Russian mathematician and statistician Andrej A. Markov (1856–1922), is such that each step preserve the memory only of the previous step. Now, the microscopic event of the signal emission is something that cannot be reduced to a previous event that would constitute the former as signal, due to the fundamental random nature of every quantum event. In fact, although a signal is propagated through a succession or chain of events (happening in a particular physical context), each event, at the quantum-mechanical scale, is an irreducible source of signals although it has been generated by the dynamical interaction of the previous signal with another quantum. For instance, a photon is generated by a laser in the state $|1\rangle$ (first event, according to Fig. 5.18, Sect. 5.2.4). Then, it undergoes beam splitting according to the experimental setup. Suppose, for simplicity that it is detected at detector D3 (second event). Then, it delivers a certain information about the path followed, and this information need to be broadcast to us, and this happens through a third event (such event of information acquisition involves in general several steps and so it is rather a coarse-grained event). Now, in the best case, we get the information about the path followed in the interferometer but we cannot know anything about the first event, since the photon could have entered the apparatus also in the state $|2\rangle$ (with the laser pumping the photon vertically). Then, the chain of quantum events need to be Markovian. Thus, although we cannot positively prove that a sequence of quantum events is random (Sect. 3.2.4), we are nonetheless authorised to assume that such a sequence has in most cases such character.

Thus, any quantum information source is only something that, provided that there is a channel and under a suitable operation (which, we have shown to be an information selection), will deliver information (which can eventually be used and *then* have causal effects of classical type). Such a source will not provide this information by itself: additional requisites (for instance, an experimental setting) are nec-

[47] Examination in Auletta (2011, Sects. 2.1–2.2).

essary (Sect. 3.4.2). Summarising, there is *no source determinism* (in the classical sense) in QM.[48] Note that it is indeterminism (violation of *omnimoda determinatio*: Sects. 1.2.1, 1.2.4, and 3.3.4) to hinder that quantum correlations become superluminal communication, as it is evident in the case of cloning (Sect. 6.1.5).

The 'alternating' dynamics of signals, a kind of quantum telegraph (dot–line–dot–$\cdots$), is in full agreement with the causal constraints of SR (Sect. 2.3.2). In fact, the sequence of events is a local–causal propagation of signals, so that the transmission of the signal cannot be instantaneous and there are limitations on its propagation speed.[49] I recall that the local propagation of the signal corresponds to the requirement that no action-at-a-distance (no instantaneous propagation of effects) be allowed. Clearly, the 'gaps' represented by the multipath evolution fall outside this scheme. Nevertheless, we shall see that they do not alter such a causal structure.

Signal Reception

Signal reception is in general at the end of a chain of events. For instance, when measuring some observable of a quantum system, we do this through some apparatus-detector that is itself composed of particles. Since it must provide a macroscopic answer to our test, it is necessary that (i) at least one of the latter particles had caught a signal from the object system, (ii) such a signal has been emitted when our detector and the object system have interacted at sub-atomic scale (giving rise to local increase in entropy), and (iii) the signal has been subsequently amplified in order to reach our sensory apparatus.[50] For instance, interferometry experiments often use photomultipliers (see e.g. Fig. 2.16, Sect. 2.4.3): the photon (the object system) impinges a metal plate and, thanks to the photoelectric effect, at least one electron is ejected that gives rise to a cascade that can be macroscopically detected. Now, if the quantum-classical evolution of a signal at the quantum scale is constituted by a Markov chain, then also the event represented by signal reception is not determined by quantum source events. Thus, there is *no collapse determinism* either in QM,[51] i.e. some classical cause that would univocally determine the reception of the signal in such and such way.

Even classically, we can take the deterministic reception of a signal as a limiting ideal case. Indeed, any information source in our universe only represents information that can be *potentially* acquired (or is actionable) only in certain physical contexts, also in the classical case (Sect. 3.4.3). Thus, the previous statement about the necessity to involve particular physical contexts is true also for the reception of information out of classical sources, and could not be otherwise since a (coarse-grained) relativistic event is based on exchange of signals that have originated from quantum sources (Sect. 5.1.1). This means that the widening of causality to causal constraints is necessary not only for dealing with QM, but also with a classical theory of signal exchanging. I recall that this was also Born's standpoint (Sect. 3.4.2). It is the

[48] Elitzur (1992).

[49] Elitzur (1992).

[50] Cohen–Tannoudji (1991, pp. 60–61).

[51] Elitzur (1992).

classical copying model (a consequence of a causal-efficient model of information), according to which information reception is a kind of reproduction of the information source, that hinders a clear understanding of this state of affairs (Sect. 3.3.6).

In fact, also classically, a source of signals is not by itself a sufficient cause of information reception, nor of the modality of such a reception. For instance, when we assume that a star will provide us with a certain amount of information without any constraints, we actually forget that this information is delivered to us thanks to many additional physical factors and parameters that permit or make it possible that the light reach us (Sect. 2.4.3). In other words, the light that reaches us is not only the light that was sent but it includes also *all that happened or was present* (in terms of physical contexts) in between (like the presence of other bodies, of fields, and so on: e.g. gravitation can bend light signals). In a delayed-choice context, in order to avoid an unacceptable retro-causation, we are forced to admit that we can deal with this information only through the actual setting that we choose at the present (Sects. 2.4.3 and 5.2.3). In the words of J. Wheeler, "The 'past' is theory. The past has no existence except as it is recorded in the present."[52] The signal that has been sent in the past is information that can be made active only through the selection of a complex of formal–causal constraints (the experimental context that we set now) and the actual act of detection.

Coarse-Grained Classical Events

The propagation of the information that spreads from a source as a signal (as the signal moves away from the source) necessarily involves (1) more and more *classical correlations* with other systems and (2) further *selection* and therefore further *source events*, any time that the signal interacts not destructively with something else. As a matter of fact, according to Huygens' principle (Sect. 6.2.4), we always observe signals *spreading out of a source* and never implode from the exterior into the 'interior', apart from light rays that collapse into black holes (but this does not change the ordinary processes of information processing and exchanging because it is light trapped in something else and not a signal 'collapsing in itself'[53]). And relativity tells us that any causal connection has the basic structure of signal sending. In fact, as already stressed, one of the most important of Einstein's contributions with SR is to have provided physics with a robust *causal scaffold* (Sect. 2.4.2).

These considerations are especially true when considering classical events. As we have seen, the chain of quantum events is Markovian, but when classical (relativistic) events are considered. it is not necessarily so. In fact, when we consider classical events, a considerable amount of information of previous steps can be kept. The reason is that classical events, as coarse-grained, although themselves not predetermined in the full sense of the word, contain also information about the (classical) relations among the quantum events of which they are a coarse graining, and this information can be kept. This is the case for sufficiently complex chemical reaction that can involve several quantum events that are interconnected, like the formation of

[52] Wheeler (1983, p. 194).

[53] Lloyd and Ng (2004).

cyanohydrin from a carbonyl compound and sodium cyanide, involving (a double) migration of several electrons.[54]

However, provided that there is a coarse graining, the consequence is a dissipative dynamics and the *the accumulation of noise* (Sects. 6.2.1–6.2.2). So, in such a case we can keep memory of past events but at the price of a (partial) disruption of such information due to the fragility of classical correlations (Sects. 5.1.2–5.1.4). In fact, since classical correlations (the source of order for classical systems), being not intrinsic, display interdependencies among different systems, in a world in which local shifts in order and disorder are allowed, are not sufficiently robust to survive local dynamics, at the opposite of quantum ones, which are robust. Thus, since the spreading out of the signal necessarily implies some information dispersion into the environment (or shifts in entropies, in the quantum-mechanical case), as understood by Brillouin (Sect. 3.2.4), this has as a consequence that, both in the classical and quantum case, we always have experience only of *a part* of the possible effects produced by any event (in the quantum case with its multipath dynamics, this is obvious). Thus, when a signal spreads from a source producing several effects on the environment, this implies that any spatial–temporal moving away from the source adds further interference with the original signal, and this is what we perceive as noise. Note that the distinction between noise and message, and thus also the ratio signal/noise, depends on the receiver,[55] in agreement with the previous analysis.

Thus, during irreversible local processes the signal spreads and information is locally lost but it is conserved at a global level. This can explain why we can recover more and more information according to the capability of science and technology to enlarge the context sufficiently. But can we recover the whole information? Can we correct the error completely? Can we eliminate any disturbance? From a quantum-mechanical point of view, a full reversible process does not permit any information acquisition. Nevertheless, since both CM and QM are ruled by reversible dynamical laws (and it must be so because the laws of Nature need to fulfil symmetries as much as possible: Sects. 2.4.2 and 3.2.4), it may be asked why causal processes (signal exchanges) cannot be reversed. Quantum-mechanically, a system can in principle undergo an indefinite number of transformations in whatever state it is, as shown in Sect. 6.2.4. However, a sequence of quantum events is Markovian, that is, we loose memory of the sequence. Moreover, when the number of involved systems grows, it becomes more and more unlikely to restore exactly certain past physical conditions.[56] In general, I note that to reverse a quantum process demands to reverse not only time but also charge and parity, according to the CPT theorem (Sect. 1.2.5). A weaker form to express the previous standpoint is to say that it is the high improbability of the latter transformation (that we also measure in terms of local loss of information) to ground the factual (and observed) local irreversibility, although the dynamics is reversible in itself. The previous argument was already used by L. Boltzmann

[54] Clayden et al. (2001, pp. 137–139).

[55] Battail (2014, p. 58).

[56] Chibbaro et al. (2014).

for justifying the second law.[57] However, it does not apply rigorously to classical physics of dissipative processes. In fact, classical systems show a natural tendency to disorder, whose statistical nature is a phenomenal aspect but rooted in the fact that local lack or significant reduction of quantum coherence cannot preserve order but induces spontaneous tendency to degradation, and therefore variability (Sect. 5.2.4). Although Poincaré's recurrence theorem, after the name of H. Poincaré, tells us that a system can go back to its initial state after a cycle (that in general is exceedingly long),[58] in fact it goes never exactly to the same state (as it happens in QM for the geometric phase: Sect. 5.1.2), what explains why we have an unattainable limit here.

Thus, in both classical and quantum physics we deal with local growth of entropy. In fact, in dissipative processes, "entropy will increase even under the complete reversal of all molecular motions since the initial state is extremely unlikely to be restored."[59] So, QM makes of the second law a fundamental principle when we deal with signal spreading. The main difference, relative to quantum systems, is represented by the fact that, when there is local growth of entropy, we cannot purify classical systems. The two aspects, namely (i) to classically keep information throughout several steps in the dynamical propagation of the signal and (ii) the impossibility to purify the systems, are closely related.

Wiener's Classical Analysis

Thus, both quantum-mechanically and classically, both signal propagation and information acquiring imply local irreversibility. This is a quite general character that depends on basic considerations on how we can deal with information, as already clearly pointed out in a classical context by the American mathematician and philosopher Norbert Wiener (1894–1964), the father of cybernetics.[60] According to him, in all classical phenomena where considerations of probability and prediction enter, the problems become asymmetrical. In fact, one can only bring a system from the past into the present in such a way that one determines certain quantities (preparation), assumes that other quantities have known statistical distributions (premeasurement), and then observes the statistical distribution of results after a given time (final detection): it may be instructive to compare these statements with the words of Heisenberg quoted in Sect. 5.2.1. The reason is that selection necessarily follows preparation (and in fact preparation is independent of selection: Sect. 5.2.2), and the very act of selection always consists in an actual reduction out of a space of possibilities. In order to reverse this sequence, we should spread a single event in a multidimensional possibility. The result would be the annihilation of the event as such, i.e. the transformation of its singularity to a sort of generality. I recall again the words of Wheeler: "The past has no existence except as it is recorded in the present" through an act of information selection.

[57] Boltzmann (1896, Sect. 6).

[58] Poincaré (1890).

[59] Elitzur (1992).

[60] Wiener (1948).

To Observe Is Essentially the Same, Classically and Quantum-Mechanically

In any acquisition of information there is always an inferential component.[61] In Sect. 5.1.2 I have in fact pointed out that, when signals spread, inferential processes play a role in information acquisition and thus also in our experience of the physical world. Let us see now which are the most general characters that inferences need to satisfy by using our preferred model: observation (a particular but instructive case of signal reception). Now, to observe is to acquire information in both the classical and the quantum case. Quantum-mechanically, we need the coupling with an apparatus in order to recover the information about an object system. Classically, as said, we have a similar situation, although not traditionally acknowledged in CM.[62] Here, we have an unknown parameter $K \subset \mathbb{R}$, whose value k we wish to know and some data d at our disposal pertaining to the set $D \subset \mathbb{R}$. These data could be the result of several detections. Once we have observed or acquired data, we must perform an information extrapolation that allows us to have an 'informed guess' about the value k of the parameter K. This is the process of *information selection*. As we know, the joint probability $\wp(j,k)$ that we select the event j (that we decide to attribute the value j to the system's parameter) while having an event represented by the value of an unknown parameter K (i.e., the probability that both event k and event j occur) is given by

$$\wp(j,k) = \wp(j|k)\wp(k), \tag{6.174}$$

where $\wp(j|k)$ is the conditional probability of the selection event j given the source event represented by k (Eq. (3.185)). Now, by taking into account the data d that are somehow the interface between the source event k and our selection event j we may express the probability $\wp(j|k)$ as[63]

$$\wp(j|k) = \sum_{d\in D} \wp(j|d)\wp(d|k), \tag{6.175}$$

where the summation is over all the data d pertaining to the set D. Note that this expression is the classical analogue of Eq. (5.75), when taking into account Eq. (5.62). By substituting the above expression into Eq. (6.174) we obtain

$$\begin{aligned}\wp(j,k) &= \sum_{d\in D} \wp(j|d)\wp(d|k)\wp(k)\\ &= \sum_{d\in D} \wp(j|d)\wp(d,k).\end{aligned} \tag{6.176}$$

Note that Eq. (6.176) can be considered as a generalisation of the well-known formula, valid $\forall j$,

---

[61] On this problem the reader may have a look at Auletta (2005a).

[62] See Auletta (2011, Sect. 2.3).

[63] Helstrom (1976).

$$\wp(j) = \sum_{d \in D} \wp(j|d)\wp(d), \tag{6.177}$$

and it reduces to the latter when $\wp(k) = 1$, i.e., when the value of the parameter K is known with certainty. It is important to stress that the two conditional probabilities $\wp(j|d)$ and $\wp(d|k)$ are quite different:

- The probability $\wp(d|k)$ represents how *faithful* our data are given the event k, that is, how reliable our apparatus (or sensory system) is, while
- The probability $\wp(j|d)$ represents our ability *to select* a *single* event j which can be used to interpret the data d in the best way.

Bayes Theorem

In such a context the *Bayes theorem*, after the name of the British statistician, philosopher, and Presbyterian minister Thomas Bayes (1701–1761), tells us that[64]:

$$\wp(j|k) = \frac{\wp(k|j)\wp(j)}{\wp(k)}, \tag{6.178}$$

where we can now say that the expression $\wp(j|k)$ represents the *a posteriori* probability that we choose the hypothesis or select the event denoted by j given the value k, while $\wp(k)$ is the probability distribution of the parameter K, $\wp(j)$ represents the *a priori* probability that the chosen hypothesis or selected event j is true, and $\wp(k|j)$ expresses the probability that the values k of the parameter have indeed occurred supposing that the hypothesis j is true. Inferences that satisfy the previous equation are called *Bayesian inferences*. Then, using Eq. (6.175), we can now express $\wp(k|j)$ in terms of $\wp(j|d)$ and $\wp(d|k)$ as

$$\begin{aligned} \wp(k|j) &= \frac{\wp(k)\wp(j|k)}{\wp(j)} \\ &= \frac{\wp(k)}{\wp(j)} \sum_{d \in D} \wp(j|d)\wp(d|k). \end{aligned} \tag{6.179}$$

In other words, we can invert the kind of question we pose and try to infer the unknown value of the parameter conditioned on having selected the event j. In fact, this is also called *inverse Bayesian inference*. Bayesian inferences are crucial for any kind of information acquisition, and are in fact performed in any experimental context. Thus, we assume for the following that all inferences used in (quantum) physics need to be Bayesian.[65]

---

[64] Bayes (1763). See also Auletta (2011, Sects. 7.6.2 and 18.4.4).

[65] There is another attempt at interpreting quantum mechanics by using Bayesian probabilities by Caves et al. (2002). It is also known as quantum Bayesianism, and clearly there are points in common with the view supported here. However, it is characterised by a strong subjectivity, according to which the state of a quantum system represents the degrees of belief an agent has about the possible outcomes of measurements.

Information Acquisition Principle

The previous results show that we never have direct access to things (whose properties are described by the parameter K) but always to things *through data* D. These data can be represented by the position of the pointer of our measuring apparatus or simply by the impulse that our sensory system has received (as in ordinary experience), or even by the way we receive information about the position of the pointer through our sensory system. It does not matter how long this chain may be (or the context in which we consider this problem). The important point is a matter of principle that can be formulated as an *information acquisition principle*[66]:

> We can receive information about objects and events only conditionally on the their effects (the data at our disposal).

Whatever event is at the source, this can be conceived as a signal that spreads out and is caught by means of the effects that produces in the exterior, including ourselves and our settings, therefore (not deterministically) generating the data to which we have access to. This implies that such information is necessarily acquired with a certain delay (Sect. 5.1.1). In other words, any signal sent by whatever source can be known only through a *later* effect due to the finite speed of light. As a matter of fact, all of our perceptual experience is mediated and slightly delayed in time.[67] This is also true for quantum-mechanical systems once that they have been submitted to experimental procedures.

We could express this by saying that any event in itself tells nothing about nothing (Sect. 5.1.1). Paraphrasing Kant, who spoke of representation as blind without the forms of the intellect,[68] it could be said that any event is *dumb*. It happens, and it is all. It is only thanks to a correlation, for example a coupling between object system and apparatus that an *objective condition* is established that allows us to subsequently acquire information and thus *infer* a certain property (see also Sects. 3.3.7 and 5.2.3–5.2.4). As we shall see later, this distinction has a formal justification. Thus, a signal tells us in general something (allows us to ascribe properties) about something (including other contexts), like a photon that is detected in certain experimental conditions that tell us which path it has taken. Paraphrasing the previous quotation of Wheeler, we can say that "A source event is inferred. It has no existence except the reception(s) of the signal that it has generated".

Thus, Bohr's Kantian epistemology seems correct in the specific sense that it points out that we make experience of *phenomena* and not of the source events directly, although I prefer to use a slightly different terminology relative to Kant's phenomenon/noumenon distinction (Sect. 3.1.1).[69] Of course, there are also many differences. In conclusion, we are obliged to understand *any* source of signals as the unattainable limit of the observation process: any source (quantum) event is a kind of irreducible singularity beyond the space-time web. The delay jointly with

[66] Formulated in Auletta (2011, Sect. 2.3). See also Auletta and Wang (2014, Sect. 11.2).

[67] Auletta (2011, Chaps. 2 and 12).

[68] Kant (1787, p. 75).

[69] Kant (1787).

the partiality of information acquisition make any relation with whatever event both perspective-like and participatory (Sect. 5.1.5). In such a context, Leibniz's idea of monads can be recalled as a kind of precognition in this sense (Sects. 2.3.3 and 2.4.1).

Quantum Mechanics, Special Relativity, Statistical Mechanics

Summarising so far, we are authorised to treat signals as the basic element of the dynamical–causal interconnections in our world (according to SR) provided that we understand that both the context and way in which these signals are received and detected are integral part of the efficient causal process itself. Thus, such an examination shows that QM and SR fully agree on the causal structure of our world as far as signals exchanging is taken into account as well as QM and thermodynamics fully agree on the fact that there is a time arrow once that some signal propagates.[70]

These concepts can be made more precise by considering the significance of, and the relation among, universal constants. According to Eq. (1.8), there is a relation between the quantum constant h and the relativistic constant c:

$$h = (m\lambda_C)\, c, \tag{6.180}$$

where the (variable) proportional factor has the dimensions [M][L]. This can be generalised by replacing the Compton wavelength by the Planck length (3.42b). Even more interesting is the relation between h and Boltzmann constant (1.355) on the basis of de Broglie's thermal wavelength (3.39), that is given by

$$h = \left(\frac{\sqrt{2\pi m T}}{\lambda_T}\right)\sqrt{k_{\mathrm{B}}}, \tag{6.181}$$

where the (variable) proportional factor has the dimensions $[\mathrm{M}^{1/2}][\Theta^{1/2}][\mathrm{L}^{-1}]$.

Consider that the Boltzmann constant k_{B}, like Boltzmann entropy, has the dimensions $[\mathrm{E}][\Theta^{-1}]$, i.e. energy over temperature. On the other hand, according to Eq. (1.2), the Planck constant h has dimension of action, i.e., energy in time: [E][T] or also of an energy over frequency since the latter has dimension $[\mathrm{T}^{-1}]$. Temperature and frequency are related as in Wien's displacement law, after the name of the German physicist Wilhelm C. W. Wien (1864–1928). As mentioned, for a detector to give a macroscopic response, it is necessary that one of its atoms receives a signal (like a photon) from the object system, and the cost of this quantum of information is the quantum of action:[71] *Planck's constant* tells us how much energy you have at a certain frequency and could be better defined as the quantum of *interaction*,[72] since it expresses the *cost of a quantum of information* (to be transmitted). On the other hand, *Boltzmann constant* tells us how much energy you have at a certain temperature or the quantity of information that is kept or stored (and can be extracted), that is, *how much information is potentially present* (Sect. 5.2.4). In brief, it is a quantum

[70] On this see Einstein (1949, pp. 687–88).

[71] Cohen–Tannoudji (1991, pp. 67–69).

[72] Cohen–Tannoudji (1991, p. 59).

of entropy information. With the words of French physicist Cohen–Tannoudji, the two constants express precise limitations showing that everything has a cost.[73] The latter limitation is one of the biggest insights of the German–American physicist Rolf Landauer (1927–1999): also according to the results of Sects 5.2.3–5.2.4, to acquire information is to select information and the latter is what costs energy. In the case of the erasure of a classical bit of information such energetic cost has been calculated and, according to Boltzmann's formula (3.79), corresponds to

$$E = k_{\mathrm{B}} T \ln 2, \tag{6.182}$$

where I recall that T the temperature of the system and ln the natural logarithm. Alternatively, erasing a single bit of information makes the entropy of the immediate surroundings increase by at least $k_{\mathrm{B}} \ln 2$. This result explains why quantum systems can process information in a reversible manner (like in quantum computation: Chap. 4) provided that no selection detection occurs.[74]

The previous examination allows us to build a new constant called *Brillouin* after the name of the French physicist that relates h and k_{B} (see also Sect. 3.2.4), which are themselves both related to energy[75]:

$$b := \frac{h}{k_{\mathrm{B}}} = \frac{1}{\lambda_T} \sqrt{\frac{2\pi m T}{k_{\mathrm{B}}}}, \tag{6.183}$$

which has the dimension of [Θ][T] (or of a temperature over frequency). This represents the *cost in action of information* or better a *quantum of cost*. It is the action needed for (transmitting or acquiring) a quantum of (classical) information. This could be perhaps formulated by saying that it expresses the action of a signal whose content is k_{B}.

Thus, energy can be defined as the *currency* of each local exchange. It is the basic physical quantity that is capable to convey and store information by propagating signals and compressing it in matter, respectively. The reason is that energy can show different degrees of order (or disorder) and thus is naturally associated with entropy. In other words, Boltzmann entropy, as a dynamical quantity, is the bridge between energy and information.

We can conclude that these three basic principles of physics: quantum-mechanical indeterminism (no source determinism but neither collapse determinism), the SR principle that any signal must be propagated at finite speed (and that of light represents the maximal speed attainable), and the thermodynamical second law (according to which all exchange processes go tendentially into the direction of entropy growing) are in fact deeply interrelated.[76]

---

[73] Cohen–Tannoudji (1991, p. 117).

[74] Landauer (1961, 1996), Bennett (1982), Bennett and Landauer (1985), Lloyd (2000).

[75] Cohen–Tannoudji (1991, pp. 71–72).

[76] Elitzur (1992).

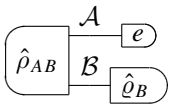
6.3 Networks

In the previous chapter, we have considered events mainly in their random nature. In the previous section, the stress was on the events as sources (of information). Now, we deepen this examination by considering events for the possible *effects* that they can give rise to (and receive from other events) through channels. These channels constitute a structure, namely a network, and, in agreement with the relational point of view supported here, "it is the structure of the connections that matters primarily".[77] In other words, we shall develop circuits that integrate quantum information and signal propagation. This means that we shall impose a causal structure on such networks.

6.3.1 D-Shaped Circuits

D-Shaped Circuits

The formalism of D-shaped kets and bras (D-kets and -bras) (Sect. 5.2.5) allows us to build D-shaped computational networks.[78] In fact, we can, for example represent Eq. (5.151) by means of the following circuit:



(6.184)

I recall that wires represent systems (e.g. $\mathcal{A}, \mathcal{B}$) freely evolving in time. I again remark that such networks represent mappings. In fact, the above input state $\hat{\rho}_{AB}$ is represented in the equation formalism as the D-shaped ket $\left|\hat{\rho}_{AB}\right)$.

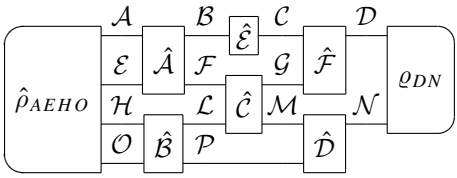
Circuits display the possible operations on systems. Both preparations and detections are events (or at least connected with events), the former without inputs (in fact, it is a determinative measurement that therefore starts a process *ex novo*), the latter without outputs (being the system absorbed in most cases or considered irrelevant for further operations). In other words, through preparation (Sect. 5.2.2) we give rise to a Markovian chain (Sect. 6.2.5). Note that events are here the only non-theoretic (and thus non-inferential) part of the scheme. Moreover, connections between the different steps are expressed in terms of probabilities.

Networks' Slices

Let us consider a relative complicated network, whose initial state is $\left|\hat{\rho}_{AEHO}\right)$ (for initial systems $\mathcal{A}, \mathcal{E}, \mathcal{H}, \mathcal{O}$):

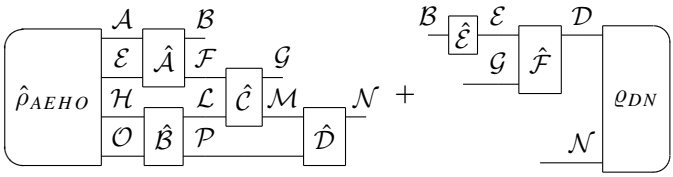
[77] Kuhlmann (2013).

[78] D'Ariano et al. (2017, Sects. 3.3 and 5.2).



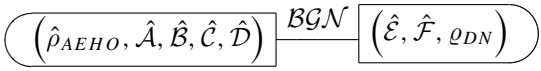
(6.185)

Note that after a transformation (represented by boxes containing each a superoperator) we assume that the system can be changed, so that any transformation can be understood as a mapping from e.g. system $\mathcal{A}$ into system $\mathcal{B}$ or vice versa. Moreover, each transformation can be considered as a channel (noisy if not deterministic) along one or more rows of the network (Sect. 6.2.1), and all multipartite inputs and gates establish correlations among systems following columns in the network. I recall that, according to Eqs. (5.70), states separate for effects and effects separate for states. These equations allow us to cut (slice) such a circuit into two parts. First, let us note that two systems are *independent* if they are not connected. A *slice* is then a set of independent systems that is complete, i.e. that would become no longer independent if we add any other system in the circuit. The two parts in which such a circuit can be sliced are a preparation test and an observation test:



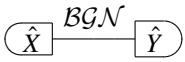
(6.186)

In fact, this is equivalent to the state-effect circuit



(6.187)

which can be further reduced to a single state-effect circuit:



(6.188)

In terms of equations, we have the preparation and observation

$$\left|\hat{X}\right) = \hat{\mathcal{D}}\hat{\mathcal{C}}\left(\hat{\mathcal{A}} \otimes \hat{\mathcal{B}}\right)|\hat{\rho}_{AEHO}) \,, \quad \left|\hat{Y}\right) = |\varrho_{DN}) = \hat{\mathcal{F}}\hat{\mathcal{E}}\left|\hat{X}\right) \,, \tag{6.189}$$

respectively, in analogy with Eq. (1.78). In fact, all possible transformations on some input systems can always be cast in a single transformation thanks to amplitude operators. This result shows that a state can only be defined for a subset of systems of a slice. Note that if all circuits are closed in this way (have some inputs and end with some effect operator), we can substitute conditional probabilities by joint probabilities.

6.3.2 *Error Correction*

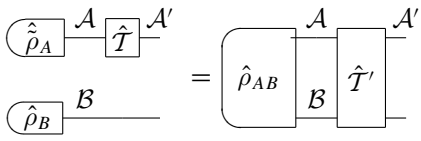
As seen in the previous section, a characteristic difference between classical and quantum information is that in the latter case we can hope to recover the information that is locally lost by enlarging the network of quantum systems, while the local dispersion process through subsequent classical (relativistic) events cannot be fully unmade. In other words, quantum-mechanically, on the basis of Sect. 6.2.1, we can have error correction as a way to drop the noise in a noisy channel (operation) as a part of a larger circuit.[79] The formalism of operations as expressed by superoperators (and related amplitude operators) becomes essential since they enable us to draw networks that embed both reversible and irreversible aspects of the quantum dynamics.

Identical Operations

The first issue is to correctly identify an operation.[80] Let us take a set of density matrices $\{|\hat{\rho}_j)\}$ spanning the whole state space of system $\mathcal{A}$ and a set of probabilities $\{\wp_j\}$. Consider now the mixed state

$$\left|\hat{\hat{\rho}}_A\right) = \sum_{j\in J} \wp_j \,|\hat{\rho}_j) \,, \tag{6.190}$$

and its purification $\left|\hat{\rho}_{AB}\right)$ for some system $\mathcal{B}$ (Sect. 1.4.3). Now, if the operations $\hat{\mathcal{T}}$ and $\hat{\mathcal{T}}'$ satisfy the equality between the following circuits:



(6.191)

it is evident that they must be equal, which means that the mapping

$$\hat{\mathcal{T}}_A \left|\hat{\hat{\rho}}_A\right) \longmapsto \left(\hat{\mathcal{T}}_A \otimes \hat{I}_B\right) |\hat{\rho}_{AB}) \tag{6.192}$$

is injective, with $\hat{\mathcal{T}}_A \otimes \hat{I}_B = \hat{\mathcal{T}}'_A$. In mathematics, we say that the mapping f is *injective* if, $\forall x \in X$ and $\forall x' \in X$, when $f(x) = f(x')$, then $x = x'$.

This sets on general ground three steps for identifying an unknown operation for any systems $\mathcal{A}, \mathcal{B}$:

1. Out of a mixture $\hat{\hat{\rho}}_A$ prepare the pure entangled state $\left|\hat{\rho}_{AB}\right)$,
2. Apply the unknown transformation $\hat{\mathcal{T}}$ on system $\mathcal{A}$,
3. Perform a complete measurement on the output state $(\hat{\mathcal{T}} \otimes \hat{I}_B)\left|\hat{\rho}_{AB}\right)$.

[79] See D'Ariano et al. (2017, Sect. 8.6).

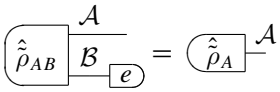
[80] See D'Ariano et al. (2017, Sects. 7.5–7.6, 8.1).

Steering

Let us now make use of the property called *steering*. First, let us note that the mixture (6.190) can be understood as a preparation in which we ignore the value of the index $j \in J$, and therefore the preparation output $\left|\hat{\rho}_j\right)$ (to which only a probability can be assigned). However, suppose that, instead of ignoring it, you encode it in another system, say $\mathcal{B}$, in such a way that everyone who measures the latter can find the value of j. This can be done if we prepare $\mathcal{B}$ in the random pure state $\left|\hat{\rho}'_j\right)$ so that the two systems are in the state

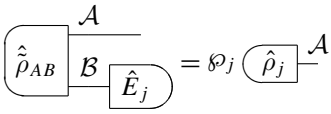
$$\left|\hat{\tilde{\rho}}_{AB}\right) = \sum_{j\in J} \wp_j \left|\hat{\rho}_j\right) \otimes \left|\hat{\rho}'_j\right) . \tag{6.193}$$

It is a kind of Lüder's mixture (see e.g. Eq. (3.12)). Such a state has two important properties. First, it is obviously an extension of the state (6.190), so that



(6.194)

Second, if one measures $\mathcal{B}$ with the test (like the choice of a polarisation filter) represented by the set of effects $\{\hat{E}_j\}$ that discriminates among the state $\{\hat{\rho}'_j | j \in J\}$, one can simulate the original preparation device of $\mathcal{A}$ for every j:



(6.195)

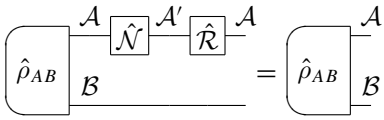
In other words, we can replace the preparation of a random pure state $\left|\hat{\rho}_j\right)$ (i.e. unknown to us) with the preparation of a single state of a larger system $\mathcal{AB}$ followed by an appropriate measurement on $\mathcal{B}$. Choosing different measurements on $\mathcal{B}$ we can 'steer' the ensemble decomposition of $\left|\hat{\tilde{\rho}}_A\right)$ deciding which particular ensemble we want to generate.

Note that this technique is strictly related to what is called indirect measurement (to measure a system for knowing the state of another system),[81] with the difference that here we deal with detection while steering is about determinative measurement.

Criterion for Error Correction

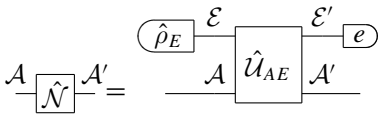
Let us now deal with error correction, having seen that each detection event induces local disturbance and noise. The condition (6.67), thanks to a recovery operation (6.65), for correcting the error represented by the noise, which is defined by transformation $\hat{\mathcal{N}}$ upon an input $\left|\hat{\varrho}_A\right)$ of system $\mathcal{A}$ (which can be very well a mixed state), can be met if we allow a purification $\left|\hat{\rho}_{AB}\right)$ of $\left|\hat{\varrho}_A\right)$ by considering an additional system $\mathcal{B}$:

[81] See Auletta et al. (2009, Sect. 9.11.1).



(6.196)

We can think of $\hat{\mathcal{N}}$ as the result of a reversible interaction between the input system and a sufficiently large environment $\mathcal{E}$ as

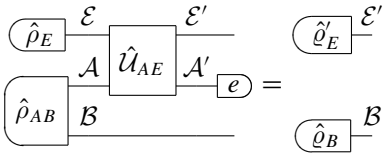


(6.197)

with $\hat{\rho}_E$ pure and

$$\hat{\mathcal{U}}_{AE}\left|\hat{\rho}_{AE}\right) := \hat{U}_{AE}\hat{\rho}_{AE}\hat{U}_{AE}^{\dagger} \tag{6.198}$$

being a trace-preserving operation (represented by a superoperator). On this basis, we get the following *criterion* for error correction[82]: the error represented by transformation $\hat{\mathcal{N}}$ is correctable upon input of $\left|\hat{\varrho}_A\right)$ iff the trace-preserving operation $\hat{\mathcal{U}}_{AE}$ does not generate correlations between the environment and the purifying system $\mathcal{B}$, i.e. iff

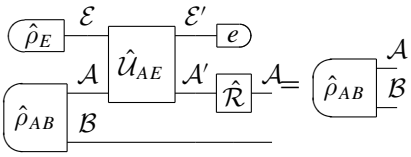


(6.199)

where $\hat{\varrho}'_E$, $\hat{\varrho}_B$ are (in general mixed) states of $\mathcal{E}'$ and $\mathcal{B}$, respectively.

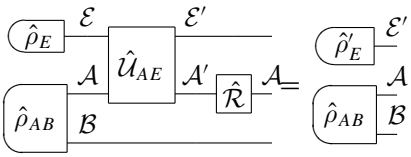
Sufficient Condition

In fact, from circuits (6.196)–(6.197) we obtain



(6.200)

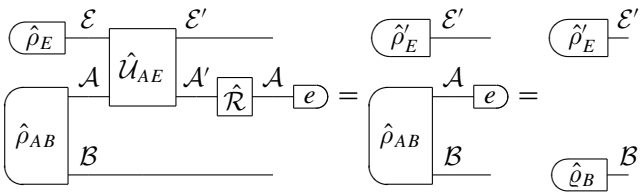
Since $\left|\hat{\rho}_{AB}\right)$ is pure, this implies that we have the product state



(6.201)

[82] D'Ariano et al. (2017, Sect. 8.6).

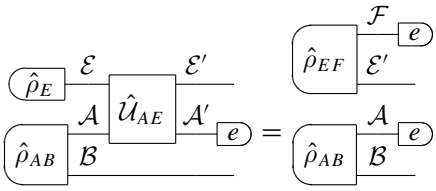
for some state $\left|\hat{\varrho}'_E\right)$ of the environment. Discarding system $\mathcal{A}$, and since $e_A\hat{\mathcal{R}}\mathcal{A}' = e_A$, we get the desired sufficient condition (if the operation $\hat{\mathcal{N}}$ is correctable, then there is no correlation between $\mathcal{B}$ and the environment):



(6.202)

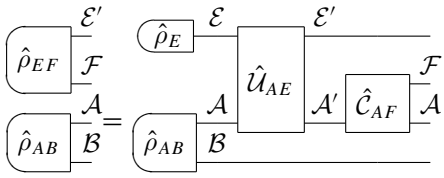
Necessary Condition

In order to prove the necessary condition, let us suppose that the circuit (6.199) holds (there is no correlation between $\mathcal{B}$ and the environment). Then, the state $\hat{\varrho}_B$ must be the marginal state of $\hat{\rho}_{AB}$ on system $\mathcal{B}$ (it suffices to discard the system $\mathcal{E}'$ on both sides). We can proceed in a similar way with the environment, by introducing an additional system $\mathcal{F}$ (that can be considered another piece of environment: Sect. 5.1.3). Then, we can rephrase this circuit as follows:



(6.203)

where $\hat{\rho}_{EF}$ is an arbitrary purification of $\hat{\rho}'_E$. Since any purification must be unique (as we shall prove later), then there must be some trace-preserving channel $\hat{\mathcal{C}}_{AF}$ between $\mathcal{A}$ and $\mathcal{F}$ and bringing $\mathcal{A}'$ into $\mathcal{F}\mathcal{A}$, such that the purification performed by adding $\mathcal{F}$ and that performed by adding $\mathcal{B}$ are their splits:

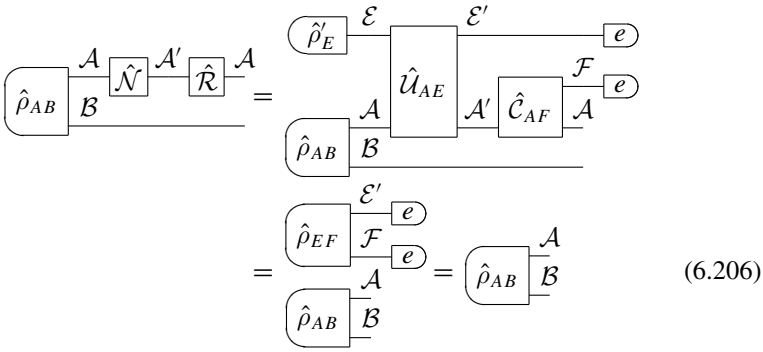


(6.204)

Defining the recovery operation

$$\hat{\mathcal{R}} := (e_F \otimes \hat{I}_A)\hat{\mathcal{C}}_{AF}, \tag{6.205}$$

and recalling definition (6.197), we get



(6.206)

In other words, if no information goes to the environment, then all the information must be in the output system $\mathcal{A}$, and therefore the error is correctable. It is to a certain extent the opposite of what happens with the loss of information in the environment during measurement but it is a splendid evidence for the correctness of that model.

Example

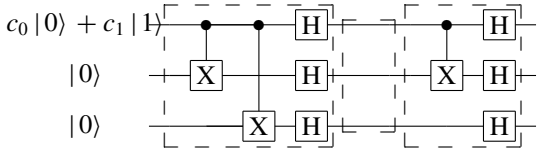
The correctness of the previous analysis can be nicely illustrated with an example (where I first use ordinary computational circuits for the sake of simplicity).[83] Let us define the *fidelity* F of a transmission as

$$F = \langle \psi \, | \hat{\rho} | \, \psi \rangle \ , \tag{6.207}$$

where $|\psi\rangle$ is the input state and $\hat{\rho}$ represents the output state, which, after decoherence, is generally a mixture. In other words, this equation represents the mean value of the output computed on the input state. It is clear that, when $F = 1$, we have perfect (deterministic) transmission. Then, we may encode an arbitrary one-qubit superposition input state $|\psi\rangle$ as follows:

$$\begin{aligned} |\psi\rangle\,|00\rangle = (c_0\,|0\rangle + c_1\,|1\rangle)\,|00\rangle &\overset{\text{CNOT}_{12}}{\longmapsto} c_0\,|000\rangle + c_1\,|110\rangle \\ &\overset{\text{CNOT}_{13}}{\longmapsto} c_0\,|000\rangle + c_1\,|111\rangle \\ &\overset{\text{H}\otimes\text{H}\otimes\text{H}}{\longmapsto} c_0\,|+++\rangle + c_1\,|---\rangle \ , \end{aligned} \tag{6.208}$$

where I recall that $|+\rangle\,, |-\rangle$ are defined in Eqs. (3.82). This enconding may be performed by the circuit in the left dashed square (compare with circuit (4.26a)):



(6.209)

[83] I follow here Auletta et al. (2009, Sect. 17.8.5). On computation and decoherence see Schlosshauer (2007, Chap. 7).

Since the first qubit already starts with a superposition of $|0\rangle$ and $|1\rangle$, after the first two CNOT gates we get a kind of (although not symmetric) GHSZ state (4.26b), while the three Hadamard gates contribute to generate an entangled state of the three qubits. Then, the latter is subject to decoherence in the central (the empty dashed square) area. Finally, the decoding area (the dashed rectangle on the right) is crucial in order to recover the initial encoding.

Now, using Eqs. (3.89), let us consider the decoherence effect on states $|+\rangle$ and $|-\rangle$ of a single qubit by adding the environment state $|\mathcal{E}\rangle$:

$$|+\rangle\,|\mathcal{E}\rangle \longmapsto |0\rangle\,|e_0\rangle + |1\rangle\,|e_1\rangle = (|+\rangle + |-\rangle)\,|e_0\rangle + (|+\rangle - |-\rangle)\,|e_1\rangle\,, \tag{6.210a}$$

$$|-\rangle\,|\mathcal{E}\rangle \longmapsto |0\rangle\,|e_0\rangle - |1\rangle\,|e_1\rangle = (|+\rangle + |-\rangle)\,|e_0\rangle - (|+\rangle - |-\rangle)\,|e_1\rangle\,, \tag{6.210b}$$

where $|e_0\rangle\,, |e_1\rangle$ are two states of the environment that are not necessarily orthogonal. Here, for brevity, I do not take normalisation into consideration. Making the substitutions

$$|e_+\rangle := |e_0\rangle + |e_1\rangle\,,\quad |e_-\rangle := |e_0\rangle - |e_1\rangle\quad, \tag{6.211}$$

we can rewrite the operations (6.210) in a way that makes decoherence manifest:

$$|+\rangle\,|\mathcal{E}\rangle \longmapsto |+\rangle\,|e_+\rangle + |-\rangle\,|e_-\rangle\,, \tag{6.212a}$$

$$|-\rangle\,|\mathcal{E}\rangle \longmapsto |-\rangle\,|e_+\rangle + |+\rangle\,|e_-\rangle\,. \tag{6.212b}$$

Thus, for the composite transformation for a single qubit we get

$$(c_0\,|+\rangle + c_1\,|-\rangle)\,|\mathcal{E}\rangle \longmapsto (c_0\,|+\rangle + c_1\,|-\rangle)\,|e_+\rangle + (c_0\,|-\rangle + c_1\,|+\rangle)\,|e_-\rangle\,. \tag{6.213}$$

Therefore, if only the first qubit of each encoded triple of qubits decoheres, starting with the encoding output state (6.208), we can write the whole transformation as

$$(c_0\,|+\,+\,+\rangle + c_1\,|-\,-\,-\rangle)\,|\mathcal{E}\rangle \longmapsto |\Psi\rangle\quad, \tag{6.214}$$

where the decohered output state is

$$|\Psi\rangle = (c_0\,|+\,+\,+\rangle + c_1\,|-\,+\,+\rangle)\,|e_+\rangle + (c_0\,|-\,-\,-\rangle + c_1\,|+\,-\,-\rangle)\,|e_-\rangle\quad. \tag{6.215}$$

From the fact that

$$|+\,+\,+\rangle = \frac{1}{2\sqrt{2}}\,(|000\rangle + |001\rangle + |010\rangle + |011\rangle + |100\rangle + |101\rangle + |110\rangle + |111\rangle)\,,$$

$$|-\,+\,+\rangle = \frac{1}{2\sqrt{2}}\,(|000\rangle + |001\rangle + |010\rangle + |011\rangle - |100\rangle - |101\rangle - |110\rangle - |111\rangle)\,,$$

$$|+--\rangle = \frac{1}{2\sqrt{2}}\left(|000\rangle - |001\rangle - |010\rangle + |011\rangle - |100\rangle + |101\rangle + |110\rangle - |111\rangle\right),$$
$$|---\rangle = \frac{1}{2\sqrt{2}}\left(|000\rangle - |001\rangle - |010\rangle + |011\rangle + |100\rangle - |101\rangle - |110\rangle + |111\rangle\right),$$

it follows that the first decoding operation (CNOT_{12}) acts on the compounds of state (6.215) as follows:

$$\mathrm{CNOT}_{12}(|+++\rangle) = |+++\rangle\,,\quad \mathrm{CNOT}_{12}(|-++\rangle) = |-++\rangle\,, \tag{6.217a}$$
$$\mathrm{CNOT}_{12}(|---\rangle) = |+--\rangle\,,\quad \mathrm{CNOT}_{12}(|+--\rangle) = |---\rangle\,, \tag{6.217b}$$

so that it changes the input state only when qubits 2 and 3 are in the state $|--\rangle$. Then, the decohered output state (6.215) maps to

$$(c_0\,|+++\rangle + c_1\,|-++\rangle)\,|e_+\rangle + (c_0\,|+--\rangle + c_1\,|---\rangle)\,|e_-\rangle\ . \tag{6.218}$$

Since $\mathrm{H}\,|+\rangle = |0\rangle$ and $\mathrm{H}\,|-\rangle = |1\rangle$, the three final Hadamard transformations send this state to

$$(c_0\,|0\rangle + c_1\,|1\rangle)\,|00\rangle\,|e_+\rangle + (c_0\,|0\rangle + c_1\,|1\rangle)\,|11\rangle\,|e_-\rangle\,. \tag{6.219}$$

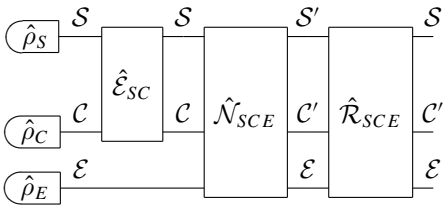
This can be generalised to

$$(c_0\,|0\rangle + c_1\,|1\rangle)\,|00\rangle\,|e_+\rangle + (c_0\,|0\rangle + c_1\,|1\rangle)\,|\cdot\cdot\rangle\,|e_-\rangle\ , \tag{6.220}$$

where $|\cdot\rangle$ is some place holder. In fact, in order to perform the error correction, it suffices to read the second and the third qubit: if one obtains $|00\rangle$, then one knows that the first qubit in the input state $|\psi\rangle$ (Eq. (6.208)) has been correctly decoded. Otherwise, we need to perform a unitary error correction. As said, we need to enlarge the systems in order to perform error correction. Note that, in such a way, no information goes to the environment, in agreement with the previous result. In fact, we have completed error correction iff we end with a product state of the input system and the environment:

$$(c_0\,|0\rangle + c_1\,|1\rangle)\,|00\rangle\,(|e_+\rangle + |e_-\rangle)\,. \tag{6.221}$$

This example can be nicely rephrased in terms of D-shaped networks:



(6.222)

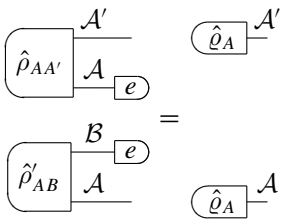
where $\mathcal{C}$ is the auxiliary encoding system and the output system $\mathcal{C}'$ is either $|00\rangle$ (in this case $\mathcal{C}' = \mathcal{C}$) or is irrelevant.

6.3.3 *Teleportation and Entanglement Swapping as Example of POVM Networks*

Let us consider in terms of networks some of the information protocols so far introduced (Sect. 3.3.6).

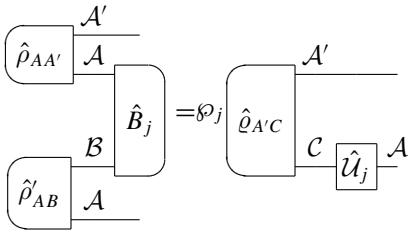
Teleportation with D-Circuits

Let us first consider teleportation by making use of the D-shaped circuits.[84] Note that, according to circuit (6.184), by definition we have



(6.223)

This means that $\left|\hat{\rho}_{AA'}\right) \otimes \left|\hat{\rho}'_{AB}\right)$ is a purification of $\left|\hat{\varrho}_A\right) \otimes \left|\hat{\varrho}_B\right)$. Now, thanks to the steering property (6.195), there must exist a (Bell) measurement such that (for $j = 1, 2, 3, 4$) we have



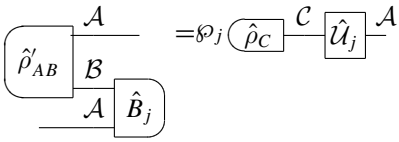
(6.224)

where $\hat{\mathcal{U}}_j$ represents the unitary operation performed by Bob and the two outputs refer to the rows that have not been "blocked". Since the mapping (6.192), which I reproduce here for a generic state $\left|\hat{\rho}_C\right)$

$$\hat{\mathcal{T}}_C \left|\hat{\rho}_C\right) \longmapsto \left(\hat{\mathcal{T}}_C \otimes \hat{I}_{A'}\right) \left|\hat{\rho}_{A'C}\right) , \tag{6.225}$$

[84] D'Ariano et al. (2017, Sect. 7.8).

is injective, we have



(6.226)

Teleportation with Amplitude Operators

The problem can be cast in a slightly different way by making use of amplitude operators. Such a formalism (Sects. 5.2.5 and 6.2.4) is able in fact to depict any kind of dynamics and in particular any kind of network so far introduced. We are accustomed to treat teleportation in terms of either a product between an entangled state for systems 23 and the state of system 1 for the input (Eq. (3.169)) or as a product state between one of the Bell basis for systems 12 and the transformed informational state for system 3 (a partial trace leaves only the latter component) (Eq. (3.171)). It is now advantageous to consider in both the input and the output the whole system 123 and see what are the transformations occurring for it. First, let us introduce a basis for the 3-qubit input. I shall make use of the basis (4.43) but with the introduction of the following nomenclature:

$$\begin{aligned} &|o_1\rangle := |0\rangle\,|00\rangle\ ,\ |o_2\rangle := |0\rangle\,|01\rangle\ ,\ |o_3\rangle := |0\rangle\,|10\rangle\ ,\ |o_4\rangle := |0\rangle\,|11\rangle\ , \\ &|o_5\rangle := |1\rangle\,|00\rangle\ ,\ |o_6\rangle := |1\rangle\,|01\rangle\ ,\ |o_7\rangle := |1\rangle\,|10\rangle\ ,\ |o_8\rangle := |1\rangle\,|11\rangle\ . \end{aligned} \tag{6.227}$$

The reason is that when we change from the expansion (3.169) into the expansion (3.171), we are in fact performing an unitary change of basis (Sect. 1.2.2). In other words, the state remains unchanged but what has evolved unitarily is the observable. In fact, we adopt the Heisenberg picture (Sect. 2.1.1), and change from a three-system observable $\hat{O}$, whose eigenbasis is represented by the eight states (6.227), to another three-system observable $\hat{O}'$, with a different eigenbasis that we need still to determine. First of all, according to the formalism (1.117a), let us write the state (3.169) of the compound system in terms of the above basis:

$$|\Psi(\{|o_k\rangle\})\rangle_{123} = \frac{1}{\sqrt{2}}\left[c_0\begin{pmatrix}0\\1\\0\\0\\0\\0\\0\\0\end{pmatrix} + c_1\begin{pmatrix}0\\0\\0\\0\\0\\1\\0\\0\end{pmatrix}\right] - \left[c_0\begin{pmatrix}0\\0\\1\\0\\0\\0\\0\\0\end{pmatrix} + c_1\begin{pmatrix}0\\0\\0\\0\\0\\0\\1\\0\end{pmatrix}\right]$$

$$= \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ c_0 \\ -c_0 \\ 0 \\ 0 \\ c_1 \\ -c_1 \\ 0 \end{pmatrix}$$

$$= \frac{1}{\sqrt{2}} \left(c_0 \left| o_2 \right\rangle - c_0 \left| o_3 \right\rangle + c_1 \left| o_6 \right\rangle - c_1 \left| o_7 \right\rangle \right) . \tag{6.228}$$

Now, under unitary transformation we rotate the basis in such a way that, in agreement with Eqs. (3.171) and (4.22) we get the new basis:

$$\left| o_1' \right\rangle = \left| \Psi^- \right\rangle |0\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \left| o_2' \right\rangle = \left| \Psi^+ \right\rangle |0\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \tag{6.229a}$$

$$\left| o_3' \right\rangle = \left| \Psi^- \right\rangle |1\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \quad \left| o_4' \right\rangle = \left| \Psi^+ \right\rangle |1\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \tag{6.229b}$$

$$\left| o_5' \right\rangle = \left| \Phi^- \right\rangle |0\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \quad \left| o_6' \right\rangle = \left| \Phi^+ \right\rangle |0\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \tag{6.229c}$$

$$\left|o_7'\right\rangle = \left|\Phi^-\right\rangle |1\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ -1 \end{pmatrix}, \quad \left|o_8'\right\rangle = \left|\Phi^+\right\rangle |1\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \tag{6.229d}$$

where I have made use of the expansions (3.170). As anticipated, these vectors represent the eigenbasis of the new (evolved) observable $\hat{O}'$. Taking advantage of the identities

$$|o_2\rangle = \sqrt{2}\left(\left|o_7'\right\rangle + \left|o_8'\right\rangle\right), \quad |o_3\rangle = \sqrt{2}\left(\left|o_1'\right\rangle + \left|o_2'\right\rangle\right), \tag{6.230}$$

$$|o_6\rangle = \sqrt{2}\left(\left|o_4'\right\rangle - \left|o_3'\right\rangle\right), \quad |o_7\rangle = \sqrt{2}\left(\left|o_6'\right\rangle - \left|o_5'\right\rangle\right), \tag{6.231}$$

in this new basis the same state $|\Psi\rangle_{123}$ is written as

$$\begin{aligned} \hat{U}_B \,|\Psi(\{|o_k\rangle\})\rangle_{123} &= \left|\Psi(\{\left|o_k'\right\rangle\})\right\rangle_{123} \\ &= -\frac{1}{2}\left[\left(c_0\left|o_1'\right\rangle + c_1\left|o_3'\right\rangle\right) + \left(c_0\left|o_2'\right\rangle - c_1\left|o_4'\right\rangle\right)\right. \\ &\quad \left. - \left(c_0\left|o_7'\right\rangle + c_1\left|o_5'\right\rangle\right) - \left(c_0\left|o_8'\right\rangle - c_1\left|o_6'\right\rangle\right)\right] \\ &= \frac{1}{2}\left(-c_0\left|o_1'\right\rangle - c_0\left|o_2'\right\rangle - c_1\left|o_3'\right\rangle + c_1\left|o_4'\right\rangle + c_1\left|o_5'\right\rangle - c_1\left|o_6'\right\rangle + c_0\left|o_7'\right\rangle + c_0\left|o_8'\right\rangle\right) \\ &= \frac{1}{2}\begin{pmatrix} -c_0 \\ -c_0 \\ -c_1 \\ c_1 \\ c_1 \\ -c_1 \\ c_0 \\ c_0 \end{pmatrix}. \end{aligned} \tag{6.232}$$

where the unitary operator $\hat{U}_B$ ruling the change of basis is built on the outline of the formula (1.115): we need to compute the scalar products between vectors of the second basis (the bras here) and vectors of the first basis (the kets here); this gives

$$\hat{U}_B = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad \hat{U}_B^\dagger = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 \end{bmatrix}.$$

It is easy to verify that the product of the two matrices (in whatever order) is the identity matrix. The reader can also verify that the application of matrix $\hat{U}_B$ to the state (6.228) gives the state (6.232) in the rotated basis. Now, we perform a measurement only on the Bell basis for systems 12 followed by Bob's unitary operation. First of all, let us rewrite the state (6.232) in the same way as (3.172)

$$\left|\Psi(\{\left|o'_k\right\rangle\})\right\rangle_{123} = -\frac{1}{2}\left(|B_1\rangle \otimes \hat{U}_1 + |B_2\rangle \otimes \hat{U}_2 - |B_3\rangle \otimes \hat{U}_3 + |B_4\rangle \otimes \hat{U}_4\right)_{12} \otimes |\psi\rangle_3 , \tag{6.233}$$

where $|\psi\rangle_3 = |\psi\rangle_1$ and the products between Bell's states $\left|B_j\right\rangle$ and corresponding Bob's unitary operators $\hat{U}_j$ are

$$\begin{aligned} |B_1\rangle \otimes \hat{U}_1 &= \left|\Psi^-\right\rangle_{12} \otimes \hat{I}_3, \quad |B_2\rangle \otimes \hat{U}_2 = \left|\Psi^+\right\rangle_{12} \otimes \mathrm{Z}_3, \\ |B_3\rangle \otimes \hat{U}_3 &= \left|\Phi^-\right\rangle_{12} \otimes \mathrm{X}_3, \quad |B_4\rangle \otimes \hat{U}_4 = \left|\Phi^+\right\rangle_{12} \otimes (-i\mathrm{Y}_3), \end{aligned} \tag{6.234}$$

in agreement with Table 3.6, Sect. 3.3.6. The input states $\left|B_j\right\rangle$'s can be any of the Bell states according to Eqs. (4.83). Finally, Alice's measurement projects the system 12 into one of the Bell states

$$\hat{P}_j \left|\Psi(\{\left|o'_k\right\rangle\})\right\rangle_{123} = \left|B_j\right\rangle_{12} \otimes \hat{U}_j |\psi\rangle_3 , \tag{6.235}$$

where $\hat{P}_j = \left|B_j\right\rangle_{12}\left\langle B_j\right|$, and Bob can apply the operation $\hat{U}_j$ to his qubit. We can put all these transformations together by making use of the amplitude operators:

$$\hat{\vartheta}_1 = \left\langle B_1 \left| \hat{I}\hat{P}_1\hat{U}_B \right| \Psi(\{|o_k\rangle\}) \right\rangle, \quad \hat{\vartheta}_2 = \left\langle B_2 \left| \mathrm{Z}\hat{P}_2\hat{U}_B \right| \Psi(\{|o_k\rangle\}) \right\rangle, \tag{6.236a}$$

$$\hat{\vartheta}_3 = \left\langle B_3 \left| \mathrm{X}\hat{P}_3\hat{U}_B \right| \Psi(\{|o_k\rangle\}) \right\rangle, \quad \hat{\vartheta}_4 = \left\langle B_4 \left| (-i\mathrm{Y})\hat{P}_4\hat{U}_B \right| \Psi(\{|o_k\rangle\}) \right\rangle, \tag{6.236b}$$

where I have dropped the index 123 for simplicity. Of course,

$$\hat{\vartheta}_j^\dagger = \left\langle \Psi(\{|o_k\rangle\}) \left| \hat{U}_B^\dagger \hat{P}_j \hat{U}_j \right| B_j \right\rangle, \text{ for } j = 1, 2, 3, 4. \tag{6.237}$$

Putting

$$\hat{\rho}_{123} = |\Psi(\{|o_k\rangle\})\rangle\ \langle\Psi(\{\langle o_k|\})|, \quad \hat{\rho}'_{123} = \left|\Psi(\{\left|o'_k\right\rangle\})\right\rangle \left\langle\Psi(\{\left\langle o'_k\right|\})\right|,$$

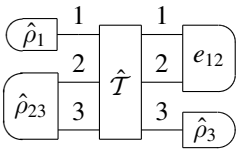
and $\hat{\rho}_3 = |\psi\rangle_3\langle\psi|$, we can write

$$\begin{Bmatrix} \hat{\vartheta}_1 \\ \hat{\vartheta}_2 \\ \hat{\vartheta}_3 \\ \hat{\vartheta}_4 \end{Bmatrix} \hat{\rho}_{123} \begin{Bmatrix} \hat{\vartheta}_1^\dagger \\ \hat{\vartheta}_2^\dagger \\ \hat{\vartheta}_3^\dagger \\ \hat{\vartheta}_4^\dagger \end{Bmatrix} = \left|B_j\right\rangle_{12}\left\langle B_j\right| \otimes \hat{\rho}_3, \tag{6.238}$$

apart from an overall phase factor. Resuming, we have

$$\begin{aligned}
\mathrm{Tr}_{12}\left(\hat{\vartheta}_j\hat{\rho}_{123}\hat{\vartheta}_j^{\dagger}\right) &= \mathrm{Tr}_{12}\left(\left\langle B_j\right|\hat{U}_j\hat{P}_j\hat{U}_B\left|\Psi(\{|o_k\rangle\})\right\rangle\hat{\rho}_{123}\left\langle\Psi(\{|o_k\rangle\})\right|\hat{U}_B^{\dagger}\hat{P}_j\hat{U}_j\left|B_j\right\rangle\right)\\
&= \mathrm{Tr}_{12}\left(\left\langle B_j\right|\hat{U}_j\hat{P}_j\hat{U}_B\hat{\rho}_{123}\hat{U}_B^{\dagger}\hat{P}_j\hat{U}_j\left|B_j\right\rangle\right)\\
&= \mathrm{Tr}_{12}\left(\hat{U}_j\left|B_j\right\rangle\left\langle B_j\right|\hat{\rho}'_{123}\left|B_j\right\rangle\left\langle B_j\right|\hat{U}_j\right)\\
&= \frac{1}{4}\mathrm{Tr}_{12}\left(\left|B_j\right\rangle\hat{U}_j\hat{U}_j\left|\psi\right\rangle_3\left\langle\psi\right|\hat{U}_j\hat{U}_j\left\langle B_j\right|\right)\\
&= \frac{1}{4}\hat{U}_j\hat{U}_j\left|\psi\right\rangle_3\left\langle\psi\right|\hat{U}_j\hat{U}_j\\
&= \frac{1}{4}\hat{\rho}_3,
\end{aligned} \tag{6.239}$$

where $\left|B_j\right\rangle$ can be any of the four Bell operator's eigenvectors and $j = 1, 2, 3, 4$. We can represent this by means of the computation circuit:



(6.240)

where, $\forall j = 1, 2, 3, 4$,

$$\hat{\mathcal{T}}\left|\hat{\rho}\right) = \sum_j \hat{\vartheta}_j\hat{\rho}\hat{\vartheta}_j^{\dagger}. \tag{6.241}$$

Entanglement Swapping with D-Circuits

We can proceed in a similar way with entanglement swapping. Using the formalism introduced here, we can resume entanglement swapping, after the initial preparation, as a two-step process. First, on the outline of Eq. (3.165), we apply a generalised change-of-basis unitary operator in this way

$$\hat{U}_B\left|\Psi\right\rangle = \frac{1}{2}\left[\pm\left|B_1B_1\right\rangle_{13,24} \pm \left|B_2B_2\right\rangle_{13,24} \pm \left|B_3B_3\right\rangle_{13,24} \pm \left|B_4B_4\right\rangle_{13,24}\right], \tag{6.242}$$

where the sign depends on which of the Bell bases the particles 12 and 34 initially are (see Eqs. (4.83)). Then, we apply a projector (on whichever couple, either 13 or 24) to get

$$\hat{P}_j\left(\hat{U}_B\left|\Psi\right\rangle\right) = \frac{1}{2}\left|B_jB_j\right\rangle_{13,24}, \tag{6.243}$$

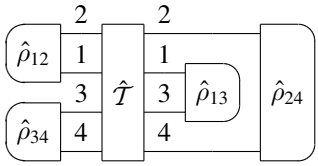
with $j = 1, 2, 3, 4$. Thanks to a Bell measurement on couples either 24 or 13 we get also the other couple entangled. Thus, putting $\hat{\rho}_{1234} = \left|\Psi\right\rangle\left\langle\Psi\right|$, we can introduce amplitude operators

$$\hat{\vartheta}_j = \left\langle B_j \left| \hat{P}_j \hat{U}_B \right| \Psi \right\rangle \text{ and } \hat{\vartheta}_j^\dagger = \left\langle \Psi \left| \hat{U}_B^\dagger \hat{P}_j \right| B_j \right\rangle, \tag{6.244}$$

where the $\left| B_j \right\rangle$'s can be either states of 13 or 24, thus allowing us to write

$$\left\langle B_j \left| \hat{P}_j \hat{U}_B \right| \Psi \right\rangle \hat{\rho}_{1234} \left\langle \Psi \left| \hat{U}_B^\dagger \hat{P}_j \right| B_j \right\rangle = \frac{1}{2} \left| B_j \right\rangle_{13} \langle B_j \right| \otimes \left| B_j \right\rangle_{24} \langle B_j \right|. \tag{6.245}$$

We can schematically summarise this in the following circuit:



(6.246)

with

$$\hat{\mathcal{T}} \left| \hat{\rho} \right) = \sum_j \hat{\vartheta}_j \hat{\rho} \hat{\vartheta}_j^\dagger, \tag{6.247}$$

for $j = 1, 2, 3, 4$ ranging on Bell states.

Bohm's Measurement Channels

As a comment to the previous formalism, it is opportune to recall here Bohm's idea of the channels of measurement.[85] He assumed that during measurement the different components of the quantum state of the object system are canalised along different 'paths'. The previous formalism makes such an idea precise, showing that a single operation is composed of channels (Sect. 6.2.1) leading in fact to the different possible measurement results. Moreover, this is in perfect agreement with the idea that each possible measurement result is coupled with a different fragment of the universe (Sect. 5.1.3) and of the local environment (Sect. 5.2.6). However, I stress that the whole measurement process is not a deterministic but a probabilistic function (nor source, neither collapse, determinism: Sect. 6.2.5), what does not allow us to consider these trajectories as classical ones (Sects. 5.2.5 and 6.2.4): they overlap and are finally discriminated entirely only at the irreversible step of detection. Bohm tried to interpret the split of channels in analogy with the bifurcation point of chaotic systems, and mentioned the non-linear factors that would be a consequence of the quantum potential (Sect. 3.3.4). However, the process here is essentially linear apart from the step of detection expressed by a projection.

Triadic Structure of Teleportation and Entanglement Swapping

In conclusion, it is interesting to consider that both teleportation and entanglement swapping schemes have a triadic structure. In the case of teleportation, we have

[85] See e.g. Bohm and Hiley (1993, Sects. 5.1–5.2).

(1) Alice *measures* the input information together with another system,
(2) The outcome is communicated to Bob (and therefore this piece of (classical) information, in addition to the quantum ebit, is *shared* with him),
(3) Bob performs a *correction* through unitary processing on the possible output.

In the case of entanglement swapping, we have

(1) A quadriparticle system is *prepared* in the state (3.165),
(2) A Bell *detection* on an unentagled couple is performed, which in this way becomes entangled,
(3) The other unentangled couple *becomes also entangled*, and thus they share information.

These schemes are not by chance but regard the fundamental characters of quantum information, as we shall see in the next subsection.

6.3.4 *Information Selection, Information Sharing, Information Acquiring*

Process, Share, Select

At the beginning of Sect. 6.3.2, it has been recalled that we identify a transformation for systems $\mathcal{A}, \mathcal{B}$ thanks to (i) the preparation of a pure entangled state $\hat{\rho}_{AB}$, (ii) application of the unknown transformation $\hat{\mathcal{T}}$ on system $\mathcal{A}$, and (iii) performing a complete measurement on the output state $(\hat{\mathcal{T}} \otimes \hat{I}_B) \left| \hat{\rho}_{AB} \right)$ Eq. (6.192). This crucial triad can be operationally generalised as follows[86]:

- Prepare a system $\mathcal{B}$ in a pure state $\hat{\rho}_B$.
- Evolve system $\mathcal{A}$ and $\mathcal{B}$ together by a reversible transformation that brings them into coupled systems.
- Discard system $\mathcal{B}$.

These three aspects correspond to what I have called elsewhere[87]

- To *process* information.
- To be correlated somehow with the source process, i.e. *sharing* information with the latter.
- To locally *select* one among different possible options.

As we have seen, when measuring, once we have observed or acquired data, we must perform an information extrapolation that allows us to have an 'informed guess' (of Bayesian kind) about the value of the external parameter representing the event on which our data conditionally depend (Sect. 6.2.5). This is the process of information *selection*. However, from our examination it follows that this information

[86] See D'Ariano et al. (2017, Sect. 7.1).
[87] See Auletta (2005b, 2006a).

selection would be totally ineffective without some pre-existing *correlation* between our settings and the detectors, since no codes could be otherwise coupled. In fact, any information-acquisition process necessarily deals with signals (out of events) telling us something about a target system. Finally, in order to get information at all this must have been *processed* at the start for introducing the necessary variability. Therefore, to *acquire* information comprehends the above three different aspects. There are reasons to assume that no other way to deal with information is possible apart from the above three aspects. Information processing should be understood as an *internal dealing with information*, while the other two aspects correspond to what is called the *linear transfer of information* from one system to another (e.g. from the object system to the apparatus) and to *non-linear selection* of a specific option, respectively (Sect. 5.1.3).

So far, we recover the operations that I have summarised for measurement, where preparing a system is an information processing (specifically, a preparation of a source: Sect. 5.2.2), coupling is information sharing (Sect. 5.2.3), and detecting is information selection (Sect. 5.2.4). However, information (in all its forms) is a much more general concept than the measurement steps. In fact, information is a second-order equivalence class. It can be understood as

- An equivalence class of states, which are equivalence classes of preparations (Sect. 5.2.2.). In fact, preparation deals with some specific physical details of the the object system. For instance, we can prepare a system in an eigenstate of the spin component $\hat{\sigma}_z$. All the different ways to prepare such a state determine it. Information is not in those details but in the fact that all 'dual' states are equivalently coded as inputs (Sect. 3.2.3).
- As an equivalence class of observables or correlations, which are equivalence classes of premeasurements (Sect. 5.2.3). In fact, any premeasurement deals again with some physical details of the object system. For instance, we deal with a spin correlation like a singlet. This defines an observable (spin in this case). At the opposite, the equivalence class of correlations deals with code–alphabet coupling. Thus, we can understand correlations as having two different meanings: correlations can be understood in a specific physical sense (like spin correlation) or in an abstract (informational) sense as what all these kind of correlations have in common, as it is evident with the Bell states (3.164).
- As a an equivalence class of properties, which are equivalence classes of detections (Sect. 5.2.4). In fact, detections always deal with the physical details of the system: to detect the polarisation of a photon is not the same as to detect the spin of an electron. All different detections that allow us to attribute a property to the system determine it. Instead, information consists here in the fact that a set of such properties represents equivalent code selections, for instance a detector clicking or not clicking.

It may sound strange that we take information to be the basic physical quantity and we consider it to be a second-order equivalence class. However, we need to distinguish here between the global point of view, according to which what is primary is the

combinatorics of possibilities, and the local point of view, according to which what is primary are information-selection events.

Measurement, Teleportation, Entanglement Swapping in a Single Scheme

This allows us to understand in new ways some of the information protocols we are dealing with in this book (which in turn highlight in a general and abstract way the crucial elements of quantum theory). In fact, according to the previous analysis, the three operations of teleportation correspond to: select information, share it, and process it, while those of entanglement swapping to process a system, measure it, share information. If this is correct, then teleportation represents a kind of 'reversed' measurement (it exchanges determinative and selective measurements) while entanglement swapping exchanges the last two steps of measurement. In fact, in the case of teleportation, it is like the output system $|\psi\rangle_3$ would have been prepared by Bob, as displayed by the RHS of Eq. (6.226). Actually, it is not truly reversed since the whole process goes according to the general rules of quantum information acquisition (and signal exchanging), the only specificity of this protocol being represented by the Bell measurement and the (spacelike) distance between the partners. In fact,

- In the ordinary process of a *measurement* we have that we process an object system $\mathcal{C}$ (Step P, for "processing"), couple it with the apparatus $\mathcal{A}$ in such a way that also detectors $\mathcal{B}$ are connected with $\mathcal{A}$ (Step S, for"sharing"), and perform a detection such that we have a mapping from $\mathcal{B}$ to $\mathcal{C}$ (Step Σ, for "selecting").
- In the *teleportation* scheme, we have Alice's measurement of systems $\mathcal{A}$–$\mathcal{C}$ (Step Σ), where $\mathcal{A}$ is entangled with some system $\mathcal{B}$. Then, Alice transfers to Bob the result of the outcome of measuring $\mathcal{A}$–$\mathcal{C}$, whilst another piece of information is shared quantum-mechanically (Step S). Finally, Bob processes $\mathcal{B}$ in such a way to get $\mathcal{C}$ (Step P).
- In the *entanglement swapping* scheme, we have the initial preparation of the entangled couples $\mathcal{A}$–$\mathcal{B}$ and $\mathcal{C}$–$\mathcal{D}$ (Step P). Then, a Bell measurement is performed on systems $\mathcal{A}$–$\mathcal{C}$ (Step Σ) so that systems $\mathcal{B}$ and $\mathcal{D}$ share now information (Step S).

Thus we have the following sequences: measurement: PSΣ, teleportation: ΣSP, entanglement swapping: PΣS. It may be recalled that we already agreed that detection can always be displaced to different steps during the information-acquisition process (Sects. 4.4 and 5.2.1). According to this examination, it seems that this is also true for the other steps of information acquisition. Moreover, I recall that coupling and detecting are inverse operations (Sect. 4.2), so that it appears in principle possible to exchange them. On the other hand, determinative measurement (preparation) and detection have in fact the same structure of information selection, the only difference being that in the first case we are interested in the *output state* while in the former case in the *input state* (after preparation), and so they could also be considered as inverse of each other. Then, they can be also exchanged. It seems that the only exchange that makes no sense is between preparation and coupling, for the simple reason that to prepare a system after a coupling would in most cases nullify coupling and so gives simply rise to a new sequence (according to the Markovian chain when quantum events happen: see Sect. 6.2.5). This analysis displays the conceptual ground of why

delayed-choice experiments cannot alter the results of the corresponding ordinary ones (Sects. 2.4.3).

As mentioned, what is important is that the general rules of signal exchanging are satisfied (Sect. 6.2.5). In fact, in the teleportation protocol, thanks to information sharing, Bob uses Alice's result for preparing a system in a certain state. It can be considered the quantum way to receive instructions for preparation. For entanglement swapping, after an initial preparation, a detection event establishes a correlation, what is quite usual as such apart from the fact that such an event sends us a signal that the other couple is also entangled.

No Information Addition

There are specific reasons for this structure of information acquisition. First, I recall that quantum systems contain an infinite amount of information as far as they did not undergo local procedures (Sect. 3.2.3). However, such an infinite amount of information is only possible and, as mentioned in Sects. 3.4.3 and 6.2.5, it is the experimental context to make it potentially available in a distributive sense (not as a whole but in the sense that any of the eigenstates of the chosen observable could be the outcome of the experiment). So nothing can, and need, to be added to these systems from the exterior, since it would contradict the unitary evolution of these systems (Sect. 6.1.1). So, when they exchange information, this cannot be done through some information addiction. I also recall that to create information is strictly impossible from a general thermodynamic point of view, because this would mean to create order out of nothing. In fact, to select information (and acquire it) is quite the opposite: it is to (locally) break some kind of symmetry (such that all outcomes are possible) and therefore to increase disorder and not order. In other words, to acquire information does not mean to add something, a piece of information that we do not had before. Quite the opposite: it means to *select something already actionable* (potentially available) thanks to a previous correlation, and therefore to *throw away* something else: it is a selection act. This is in agreement with the fact that to select information has an energetic cost (Sect. 6.2.5).

No Mechanical Transfer of Information

This examination gives us the opportunity to correct a big misunderstanding when dealing with information acquisition: very often one assumes that there is something that is transferred from the source event (or process) to the receptor, which would bring us back to a pure mechanical understanding of the causal effect of information (Sects. 3.3.6, 3.4.1–3.4.2, and 6.2.5). However, properly speaking, there is nothing that happens in this way. It is clear that something needs to physically interact with the system that works as a receptor or detector (for instance, a quantum impinging on the latter, but also the human eye works as a quite efficient photodetector), what implies somewhere a quantum event (Sect. 6.2.2) and therefore local exchange of the quantum of interaction (energy) and efficient mechanical kind of causation (Sect. 6.2.5). Nevertheless, the final selection is operated among informational options that are *inside* the receptor system (and are locally codified by it: I recall that, in the most elementary situation, also a detector's clicking or not clicking is a code alphabet),

and this selection tells something about the source event *only because* there is a correlation (which we have assumed to be a kind of causal constraint and not an efficient form of causation). Indeed, the linear transfer of information rather consists in the sole fact that the informational resources of one system are now coupled with the informational resources of the other system, and since we have shown in Sects. 3.3.6, 5.2.3 and 5.2.6, that systems can become correlated without any local interaction, we must definitely exclude a mechanical transfer of something from one system to the other. What I am saying here applies also to the classical case.

Refinement of the Definition of Physical System

On the basis of this examination, we can postulate that any kind of operation can be finally led to one of these informational operations (or to a combination of them). In Sect. 5.2.6 I have defined a physical system as what can undergo the different experimental procedures. Now that we have established a robust connection between these operations and the three informational operations, we can generalise the (whether classical or quantum-mechanical) notion of physical system by disentangling it from the classical notion of object. In fact, we can overcome the difficulties raised previously (Sects. 2.1.1, 2.1.3, and 3.3.4) and say that we consider a *physical system* whatever could, at least in principle, be dealt with by one of the informational operations summarised here. Then, a physical system is defined as *something* (a type–token: Sects. 3.4.2 and 5.1.2) that will give some (not necessarily deterministic) *output* when undergoes one of these *operations*.

The Three Operations in Details

Now, I shall show that the whole process of information acquiring consists of the mentioned three aspects. To this purpose, let us first consider the three operations in general terms (for both classical and quantum information)[88]:

- Viewing the whole from the point of view of the involved systems, we have that the object (measured) system which represents, as I have stressed, encoded information evolving reversibly in time if let free, is anyway capable to process information and so to deliver an output during preparation; thus, it can be considered as an information *processor* (Chap. 4). This time evolution can be considered as a (deterministic or, for quantum-mechanical systems, both deterministic and probabilistic) change in some informational elements.
- The measuring device that is coupled to the object system is a *regulator*, while the final operation of detection is performed through a *decider*. The regulator owes its name to the fact that choosing a certain apparatus with a certain experimental setup and with a certain pointer indeed contributes to determine the (formal and therefore potential) conditions in which information is acquired. Since its role consists in a coupling of code alphabets, and for this reason cannot provide the necessary variety of information (which is guaranteed at the source by the object system as processor), this determination is rather a tuning of the measurement process.

[88]For what follows see Auletta (2011, Sect. 2.3). See also Auletta and Wang (2014, Sect. 12.5).

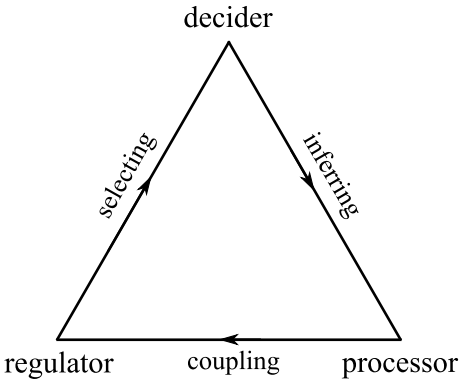


Fig. 6.10 The fundamental information triad is consists of three elements (processor, regulator and decider) and three relational operations (coupling, selecting, and inferring). Adapted from Auletta (2011, p. 48)

- The final selection that happens at the detectors can be random or not (in the quantum-mechanical case it is in fact random). However, there is always some sort of incertitude affecting the final selection act (and therefore also information acquisition as a whole) due to the non-linearity of this selection. This means that also the encoded activity in the final detector is crucial here. In many classical situations, the information is linearly coded only by the final receptors while the signal can be, for example a chaotic perturbation of some medium. In QM is not so but the detector code is also crucial. Obviously, the decider can provide a selection that, thanks to the indirect connection with the processor through the regulator, will finally consist in an option within a set of alternative possibilities.

Information Triad

The whole process can then be represented in Fig. 6.10 as the fundamental information triad. The relation established between the regulator and the processor is coupling, which allows information to be subsequently acquired. The relation between the decider and the regulator is information selection. Finally, the decider can acquire information about the processor (or the event resulting from processing) by performing in this way the analogue of inferring (the experimenter performs any way precisely such an inference: see Sect. 6.2.5), on the outline of the subjectivist interpretation without obviously supporting their pure epistemic view of QM (Sect. 3.1.1). In other words, when a decider (even randomly) selects and eventually stores some information which through an appropriate coupling reveals something about another system, we have a result that at a pure physical level bears some structural relation to what is a true inference. It is also important to realise that this inferring could be considered as part of a further preparation procedure in which a processor is determined. For instance, we may decide to measure our system again (provided that it is not annihilated in the detection event) starting from its output state. The reason could be that we are not sure of our inference because of some doubts about the apparatus reliability. In this case, the outcome of the first measurement can be considered as a

new preparation (do not forget that preparation is a determinative measurement).[89] In this way, the whole process presents a certain circularity.

Shannon Information

It has been pointed out that if one begins with the core Shannon's notions of source, channel and coding, the generalisation from classical to quantum information is straightforward and unproblematic.[90] This is true but with some provisos, as recalled in Sect. 3.2.4. In fact, Shannon, the father of classical information theory, pointed out that information is concerned with a reduction of incertitude (a choice among alternative possibilities) and that, in order to have an exchange of information, we need both variety and interdependence.[91] Note that variability as such does not imply disorder. In fact, also isolated quantum systems display variability although they are maximally ordered. It is only information acquisition that both requires and generates disorder. So far, there is good agreement but the provisos are

- Shannon's approach mainly dealt with engineering problems of communications and therefore made use of a statistical notion of information (Shannon entropy) that results not fully adequate for dealing with the intrinsic variety of the quantum information also when von Neumann entropy is zero (Sect. 6.1).
- The second proviso is the following. For the problems dealt with by Shannon, the task is to increase the matching between an input and an output in controlled situations. In this case, the reduction of incertitude already happens at the source (by the sender who chooses a certain message among many possible ones). Thus, the subsequent process has only the purpose to be faithful to this choice: the main parameter is represented here by *fidelity*, in fact the degree of matching between output and input. This is crucial for information in general and is also used in QM where it is defined as (6.207). It is clear that, when $F = 1$, we have perfect transmission (compare with what said in Sect. 5.2.4). This is the case for all deterministic (even quantum deterministic) transformations. However, when dealing with signal exchanges, when information is in fact acquired, the possibility to compare input and output in such a way represents rather a limiting case even in many classical situations (Sects. 6.2.1 and. 6.3.2).

The problem was that Shannon's followers took this specific model as a general model of dealing with information, although Shannon himself warned about a too schematic interpretation of his own contribution[92]:

> A thorough understanding of the mathematical foundation [of information theory] and its communication applications is surely a prerequisite to other applications. I personally believe that many of the concepts of information theory will prove useful in these other fields—and, indeed, some results are already quite promising—but the establishing of such applications is not the trivial matter of translating words to a new domain, but rather the slow tedious process of hypothesis and experimental verification.

[89] D'Ariano et al. (2017, Sect. 2.8.1).

[90] See e.g. Timpson (2013, Sect. 3.6).

[91] Shannon (1948).

[92] Quoted in Battail (2014, p. 29).

The worry consists in the fact that, in the most general (also classical) case, the reduction of incertitude needs to be at the *output* and not at the input. This is evident for quantum systems since input (or prepared) qubits do not represent (classical) selected messages. We can, however, make the initial Shannon's idea of the message selection at the input broader and speak, in the quantum-mechanical case, of the *selection of a qubit*. In that case, the classical and the quantum case become similar and what we presuppose is in fact a *preparation* at the start of the communication process. Obviously, we need to replace the identification of source variety with statistical entropy by a deterministic–probabilistic (quantum) information processing able to generate variability (and thus randomness of the possible outcomes) independently of any statistical consideration (see also Sect. 6.1).

Any Event is an Ultimate Fact

The selection operated by the receiver *is* the received message (it is the physical event, the token, that has *in fact happened*). This justifies why information selection is the only irreversible happening here. In fact, as mentioned, the receivers may also try to verify again whether their understanding is correct. However, this understanding concerns the inference about the input message, not the *act of selection* at the output, i.e. it does not concern the event itself by which a reduction of incertitude at the output has been produced forever. This, at least as far as we can testify it, is an ultimate fact and we cannot circumvent it (while the inference is its interpretation). This recalls us the difference between events and property ascription (Sect. 5.2.4), and explains that events are the only non-inferential and therefore non-theoretical part of any network, as anticipated in Sect. 6.3.1.[93] This is the main difference between what is only possible and what is physically actual in our (or presumably any other) universe. In other words, we find in information selection the theoretical and formal ground for the ontological definition of quantum event given in Sect. 5.1.1.

6.4 Quantum Mechanics Out of Six Informational Principles

We have considered the basic elements of quantum information. It is now time to address its foundational relevance. We shall follow the research of the Italian physicist Mauro D'Ariano and his co-workers, Giulio Chiribella, and Paolo Perinotti, trying to go to the foundations of QM as a continuation of Wheeler's work (Sect. 3.2.2). In fact, it is a new derivation of quantum theory from purely information-theoretic assumptions. This turns out what appeared to be a weakness of information (its adimensionality) in its strength: its absolute generality. What is most important is that such an approach is *operational*, an issue that will be explored further in the following.

[93] Geroch (1978, p. 3) tells us that events need to be considered as a part of the world in which we live and not as a theoretical construct.

In the previous sections, I have pointed out the main commonalities between classical and quantum information. Here, we shall ground this examination on the basis of six general principles. In fact, all of them apart from the last one are common to both classical and quantum theories. The six principles can be formulated both in the framework of quantum physics as also part of a general operational probabilistic theory (OPT).[94]

Although I prefer a more pragmatic approach to an axiomatic one (as it will be evident in the last chapter of the book), this research is extremely important since it shows that quantum information is necessary and sufficient for accounting for the whole of QM.

6.4.1 Principle of Causality

Preparation and Detection

The principle of causality grounds the way in which signals are emitted out of events and propagated through some network. As anticipated in Sects. 5.2.2 and 6.2.5, the probability of preparation is independent of the probability of observation performed at later steps, where *later* needs to be understood in terms of computational steps in a network.[95] In fact, assume that we prepare some system in some state, where such a preparation is achieved probabilistically (note that here and in the following we shall deal with systems that are in general finite-dimensional). That is, we have only a probability $\wp_i$ to prepare the system in the state $\left|\hat{\rho}_i\right)$. The set of all these possible outcomes of preparation is a preparation test $\{\left|\hat{\rho}_i\right)\}$, so that we assume that the system undergoing preparation is in some initial state

$$\left|\hat{\rho}\right) = \sum_i \wp_i \left|\hat{\rho}_i\right) , \tag{6.248}$$

where the $\left|\hat{\rho}_i\right)$s are pure. The density matrix $\left|\hat{\rho}\right)$ can be thought of to represent an ensemble of systems. This state is in most cases a mixture but it is not necessarily so (if one of the $\wp_i$ is 1 and the other probabilities 0, then it is a pure state). For the sake of simplicity, assume that the $\left|\hat{\rho}_i\right)$'s represent a maximal orthogonal basis for the state space of the system and $\left|\hat{\rho}\right)$ is a maximally mixed state, so that, according to Eq. (1.368) (Fig. 1.28, Sect. 1.4.1), for n dimensions we have

$$\left|\hat{\rho}\right) = \frac{1}{n}\sum_i \left|\hat{\rho}_i\right) = \frac{1}{n}\hat{I}. \tag{6.249}$$

[94] D'Ariano (2010).

[95] D'Ariano et al. (2017, Sect. 2.4 and Chap. 5).

Let us now write the joint probability to have prepared the system in the state $\left|\hat{\rho}_i\right)$ and have observed the event represented by effect $\hat{E}_j$:

$$\wp(i,j) = \left(\hat{E}_j \left|\hat{\rho}_i\right.\right). \tag{6.250}$$

Thus, the marginal probability (see circuit (6.184))

$$\wp(i) = \sum_j \wp(i,j) = \sum_j \left(\hat{E}_j \left|\hat{\rho}_i\right.\right) = \left(e \left|\hat{\rho}_i\right.\right) = \mathrm{Tr}\left(\hat{\rho}_i\right) \tag{6.251}$$

of preparing the ith state of the ensemble does not depend on the choice of the *observation test*. That is, $\{\hat{E}_j\}$ is independent of the choice of a particular set of effect operators spanning the space of the initial state of the system.

Causality

Let us put the previous considerations in general terms. Taking advantage of networks (6.185)–(6.186), we can say that *causality* means that the probability distribution of a test within a circuit is not conditioned by tests not connected to its inputs. In such a case, having preparation $\hat{\mathcal{X}}$, whose output is $\left|\hat{X}\right)$, and observation $\hat{\mathcal{Y}}$, whose output is $\left|\hat{Y}\right)$, we say the probability $\wp(i|\hat{X},\hat{Y})$ to prepare a system in a state $\hat{\rho}_i$ is independent of $\hat{\mathcal{Y}}$, or that, given another observation test $\hat{\mathcal{Z}}$, whose output is $\left|\hat{Z}\right)$, we have

$$\wp(i|\hat{X},\hat{Y}) = \wp(i|\hat{X},\hat{Z}) = \wp(i|\hat{X}). \tag{6.252}$$

We can also say, in general terms, that a test $\hat{\mathcal{T}}$ *precedes* another test $\hat{\mathcal{T}}'$ if some output systems of transformation $\hat{\mathcal{T}}$ are connected to some input systems of transformation $\hat{\mathcal{T}}'$. Recall that the connectivity among events relies on both causal constraints (correlations among possible events) and causal connections.

An alternative way to consider the problem is to say that a theory is causal iff for every system there is a unique deterministic operation (Eqs. (5.67) and (5.149)). In other words, in a theory that is not causal, we have multiple deterministic effects for the same system, what implies that preparations do exist that are achieved with different probability depending either on the observation test or on correlations with other systems. This is why entanglement cannot alter the local probability distributions. We can also formulate this by saying that a probabilistic theory is causal iff every state is proportional to a deterministic one, and we get a deterministic form from a probabilistic one through normalisation.[96] Then, if a theory is not causal there must exist probabilistic states that cannot be prepared deterministically. In other words, the failure of causality would correspond to a *limitation of preparations* or to local operations in general.

[96] D'Ariano et al. (2017, Sect. 5.4).

The existent plurality of causes could represent a threat to a well-defined notion of causality.[97] However, both the connectivity of events and the fact that networks can be sliced, as recalled in Sects. 5.2.5 and 6.3.1, allows us to build up foliations over the operational circuits that lead to the equivalence with Einstein's notion of causality (Sects. 2.3 and 3.3).[98]

Causality and Tests

Expanding the previous statements, causality is equivalent to the following statements about *tests*:

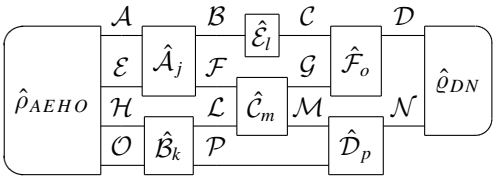
1. Completeness of the observation (or preparation) tests: for any system $\mathcal{A}$ and any observation (or preparation) test $\{\hat{E}_j\}_{j\in J}$ we have

$$\sum_{j\in J} \hat{E}_j = \hat{I}_A. \tag{6.253}$$

2. Completeness of tests for subsystems: for any systems $\mathcal{A}, \mathcal{B}$ and any test $\{\hat{E}_j\}$ from $\mathcal{A}$ to $\mathcal{B}$ one has

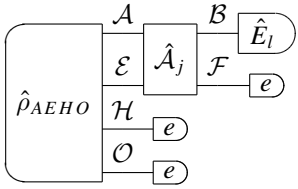
$$\sum_j (e_B \,|\, \hat{E}_j = (e_A \,|\,. \tag{6.254}$$

3. Domination of transformations, i.e. Kraus condition (5.134). When we have equality, $\hat{\mathcal{T}}$ is a deterministic channel.
4. Domination of effects: effects are dominated by a unique deterministic effect operator. This means that we can substitute a network with another in which all tests following the event of interest are substituted by an arbitrary deterministic test. Consider again the network (6.185) but with a slight modification, so that it reads now



(6.255)

Now, assume that we are interested to evaluate the probability $\wp(j, l|\hat{\rho}_{AEHO})$ of events $\hat{\mathcal{A}}_j$ and $\hat{\mathcal{E}}_l$. Then, we can replace it by the following circuit:

$\wp(\hat{E}_l, \hat{\mathcal{A}}_j) =$


(6.256)

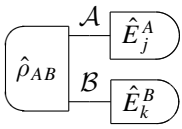
[97] Poincaré (1902, p. 45 and 51–52).

[98] D'Ariano and Tosini (2013) D'Ariano et al. (2017, Sect. 5.1).

In other words, the deterministic effect works as a wastebasket since it erases all unwanted information.

Local Tests and Correlations as Insufficient

Since it is impossible to signal without interaction (Sect. 6.2), in a causal theory it is impossible to send signals by performing only local tests in presence of some quantum correlations without specifying which is the sequence of operations. Suppose Alice and Bob performing some local tests on their systems, so that the joint probability $\wp(j,k)$ can be represented by the circuit

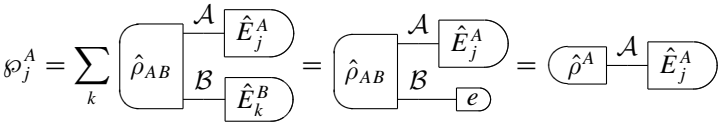


(6.257)

The state $\left|\hat{\rho}_{AB}\right)$ is an entangled state and the two marginal probabilities are given by

$$\wp_j^A = \sum_k \wp_{jk}, \quad \wp_k^B = \sum_j \wp_{jk}, \tag{6.258}$$

in agreement with Eqs. (3.179). Now, Alice's marginal does not depend on the choice of Bob's test since, according to circuit (6.194), we have



(6.259)

This shows that both space and time (space-time) are involved in causal relations, in the sense that there must be indeed some local interaction in presence of correlations but we also need a *specific sequence* of operations in the propagation of the signal.

Four Remarks

I have four remarks about this first principle:

1. In any consideration of causality we need to involve the notion of premeasurement (Sect. 5.2.3).
2. Premeasurement necessarily involves the notion of quantum features and correlations in general. Then, it is necessary to take into account causal constraints, that is, restrictions of the space of the possible events determined by the correlations that arise through a coupling process (Sect. 6.3.4). In other words, in the causality ruling, for example the measurement process also causal constraints play a (weaker) causal role, but it is crucial since causal constraints contribute to establish causal dependency (Sects. 3.4 and 5.1.2). This means that QM can represent the most general theory of causality if also causal constraints are considered to be fundamental part of the causal processes occurring in our world.

3. The proponents like to keep the assumption of causality distinct of, and preliminary to, any notion of locality. If we understand with "locality" any notion connected with space-time, this is right for the reasons that have been already mentioned in Sects. 5.1.5 and 6.2.5. However, in order to speak of an event as something irreducible to any theoretical notion, we need a kind of (ontological) *determination* and this is provided by the notion of information selection (for instance, by selecting a particular state during preparation), which is therefore basic also to the notion of causality. I think that the"it from bit" of Wheeler (Sect. 3.2.2) went in such a direction.
4. The principle, rather than about quantum information as such, it is rather about the effects of quantum events (or quantum sources), that is, it is about the way signals are generated and exchanged.

6.4.2 *Atomicity of the Composition*

Atomicity of the composition tells us that the composition of two atomic operations is atomic.[99] A world in which this principle be violated, Alice would perform a local operation $\hat{\mathcal{A}}_j$ through which she could precisely know what is her outcome, and similarly Bob with his system through the operation $\hat{\mathcal{B}}_k$, and nevertheless the combined resulting operations, i.e. $\hat{\mathcal{B}}_k\hat{\mathcal{A}}_j$, could be simulated by a third party by performing a test $\{\hat{\mathcal{C}}_x\}$ such that

$$\mathcal{A} \boxed{\hat{\mathcal{A}}_j} \mathcal{A}' \boxed{\hat{\mathcal{B}}_k} \mathcal{A}'' = \sum_x \mathcal{A} \boxed{\hat{\mathcal{C}}_x} \mathcal{A}'' \tag{6.260}$$

Now, what is such extra information? Since Alice already had maximal knowledge about the first step and Bob about the second step, the outcome x should specify how these two steps interacted together, what is either a third (intermediate) step, and so the composed process is again atomic, or a piece of information that is non-local in time. However, this could only be a correlation (Sect. 5.1.5), but a correlation alone cannot provide us with additional information about whatsoever happened or happening (Sects. 3.2.2 and 3.4.2). This proves that several atomic operations can always be formulated as a single atomic transformation, as, for example in circuit (6.240).

In simple words, this principle establishes that "maximal knowledge of the episodes implies maximal knowledge of the full history".[100] A physical theory where this assumption did not hold would be highly pathological, because the mere composition of two processes in which events are involved, which considered by themselves are specified with the maximum degree of accuracy possible, would generate some global information that cannot be accessed on a step-by-step basis.

[99] D'Ariano et al. (2017, Sect. 2.8.3 and Sect. 4.1).

[100] D'Ariano et al. (2017, Sect. 5.3).

6.4.3 Perfect Discriminability

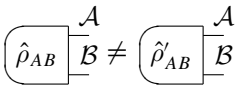
The third assumption is called of *perfect discriminability*.[101] First, let us note that, for preparation processes, the coarse-grained processes are called mixed states and fine-grained processes are called pure states, since the former represent loss of information relative to the latter (Sect. 3.1.1). In fact, we have that the Shannon entropy of the system is $\geq$ the von Neumann entropy, with sign of equality only when the states are orthogonal, what implies that by mixing non-orthogonal states there is a loss of information (Sect. 3.2.2). Thus, perfect discriminability should be sufficient condition of orthogonality (Sect. 6.1.1).

Perfect discriminability means that there is an observation test with probabilities 0 and 1 that is capable to assess with certainty a non-trivial property of the system. The assumption tells us that every state that is not completely mixed (Eq. (6.249)) can be perfectly distinguished from some other state. In more operational terms, if a state is not compatible with some preparation, then it is perfectly distinguishable from some other state that is the result of such preparation procedure: see Sect. 5.2.2. This means that if we have some definite information about a state, we are also able to experimentally verify/falsify it with certainty.

6.4.4 Local Discriminability

Local discriminability is a consequence of the fact that effects separate states and states separate effects (Sect. 5.2.4). In fact, the latter means that local effects that are factorised into single-system effects exist and they separate joint multipartite states as well as local states separate multipartite effects. This can be proved by considering the direct product between Hilbert spaces of subsystems constituting a compound system. Such a consequence is in turn sufficient condition of local discriminability: we can discriminate any pair of joint states of multipartite systems by using only *local* measurements.[102] In other words, *quantum-mechanical holism* is compatible with *local reductionism*. Thus, we are again confronted with a basic (but not to be understood in sharp terms) dichotomy between global and local.

Thus, local discriminability means that, if two bipartite states are different, then they give different probabilities for at least one product experiment. In other words, $\forall \left| \hat{\rho}_{AB} \right), \left| \hat{\rho}'_{AB} \right)$, there are local effects $\hat{E}^A$, $\hat{E}^B$ such that the circuit

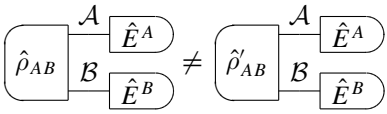


(6.261)

is sufficient condition of the circuit

[101] D'Ariano et al. (2017, Sect. 4.2 and Chap. 10).

[102] D'Ariano et al. (2017, Sect. 2.4.1, Sect. 4.2, Chap. 6).



(6.262)

This is also called *local tomography* (see Sect. 6.2.3), since it means that the state of a composite system can be completely determined by the statistics of local measurements performed over the components or subsystems. In other words, for any state of a composite system, the joint probabilities of all local effects completely identify the state. I also recall that, if local discriminability holds, for any two transformations $\hat{\mathcal{T}}$, $\hat{\mathcal{T}}'$, the condition $\hat{\mathcal{T}} \left| \hat{\rho} \right) = \hat{\mathcal{T}}' \left| \hat{\rho} \right) , \forall \left| \hat{\rho} \right)$ implies that $\hat{\mathcal{T}} = \hat{\mathcal{T}}'$, as displayed in circuit (6.191).

It is important to note that what is required here is that the error probability in discrimination is strictly smaller than $1/2$.[103] The case of perfect local discriminability would correspond to zero error, but it is not required by the principle.

It may be noted that, mathematically, this principle is the key reason for the choice of *complex* (instead of real) *Hilbert spaces* (Sects. 1.2.1–1.2.2): in real Hilbert space quantum theory, there are some bipartite states that can be distinguished perfectly with global measurements but give the same statistics for all possible local measurements (working with real Hilbert spaces satisfies the local tomography only if we restrict our considerations to pure states). In other words, a quantum theory with real Hilbert spaces satisfies the atomicity of composition but not local tomography.

Note that from local discriminability follows that the dimension of the state space of any composite system is the product of the dimensions of component systems, as displayed, for example by Eq. (1.381). In other words, the basic structure of the tensor product for composite systems is a consequence of local discriminability. And this is clearly the classical component of entanglement as far as the product structure of states (or Hilbert) spaces of composite systems allows us to characterise a transformation by computationally running it only on the input system *without considering entangled states with any other additional system*.

6.4.5 Ideal Compression

The fifth assumption is called *ideal compression*: for every state there exists an ideal compression scheme, in agreement with Shannon's classical information theory (Sects. 3.2.4 and 6.3.4). The difference of the quantum-information approach relative to Shannon's theory of information is that the latter deals with noise (and irreversible loss of information) and statistics, while here we deal with a noiseless single use of the source (Sects. 3.2.4 and 6.1), what allows us to speak of ideal compression.

Making use of the results of Sects. 6.2.1 and 6.3.2, we can put the things on a general and abstract ground.[104] We can say that a compression scheme that is lossless

[103] D'Ariano et al. (2017, Sect. 3.3).

[104] D'Ariano et al. (2017, Sects. 4.3 and 8.3).

and maximally efficient is ideal, and the fifth assumption states that ideal compression of information is *always possible*. In other words, ideal compression requires that every state $\left|\hat{\rho}\right)$ can be ideally encoded into a suitable physical system. Moreover, in every ideal compression protocol, the decoding is an atomic transformation. Ideal compression allows us to identify the subset $F_{\hat{\rho}}$ of the state space of the system, and this identification is essentially unique. This allows the following crucial result that is the basis of any quantum information transfer (e.g. between system and apparatus): if two systems $\mathcal{A}$ and $\mathcal{A}'$ allow for ideal compression of a state $\left|\hat{\rho}\right)$ of a third system, then they are operationally equivalent. In fact, suppose that $(\mathcal{A}, \hat{\mathcal{E}}, \hat{\mathcal{D}})$ and $(\mathcal{A}', \hat{\mathcal{E}}', \hat{\mathcal{D}}')$ are two ideal compression protocols for state $\left|\hat{\rho}\right)$, where I recall that $\hat{\mathcal{E}}$, $\hat{\mathcal{E}}'$ and $\hat{\mathcal{D}}$, $\hat{\mathcal{D}}'$ denote encoding and decoding, respectively. Let us define two transformations

$$\hat{\mathcal{U}} = \hat{\mathcal{E}}'\hat{\mathcal{D}} \quad \text{and} \quad \hat{\mathcal{U}}' = \hat{\mathcal{E}}\hat{\mathcal{D}}' \tag{6.263}$$

from system $\mathcal{A}$ to system $\mathcal{A}'$ and from system $\mathcal{A}'$ to system $\mathcal{A}$, respectively, which are clearly reversible and are the inverse of each other. Note that $\hat{\mathcal{U}}$ has the same general form displayed in Eq. (6.198). In fact, for every state $\left|\hat{\varrho}\right)$ of system $\mathcal{A}$ one has

$$\begin{aligned}
\hat{\mathcal{U}}'\hat{\mathcal{U}}\left|\hat{\varrho}\right) &= \hat{\mathcal{E}}\left(\hat{\mathcal{D}}'\hat{\mathcal{E}}'\right)\hat{\mathcal{D}}\left|\hat{\varrho}\right) \\
&= \hat{\mathcal{E}}\left(\hat{\mathcal{D}}'\hat{\mathcal{E}}'\right)\left|\hat{\sigma}\right) \\
&= \hat{\mathcal{E}}\left|\hat{\sigma}\right) \\
&= \hat{\mathcal{E}}\hat{\mathcal{D}}\left|\hat{\varrho}\right) \\
&= \left|\hat{\varrho}\right) ,
\end{aligned} \tag{6.264}$$

where I have used Eqs. (6.60) for $\mathcal{A}$ and $\mathcal{A}'$, and defined $\left|\hat{\sigma}\right) = \hat{\mathcal{D}}\left|\hat{\varrho}\right)$ as a state of $\mathcal{A}'$ in analogy with Eq. (6.63). From this, it follows that

$$\left|\hat{\varrho}\right) = \hat{\mathcal{D}}'\hat{\mathcal{E}}'\left|\hat{\varrho}\right) = \hat{\mathcal{D}}\hat{\mathcal{E}}\left|\hat{\varrho}\right) . \tag{6.265}$$

Since $\hat{\mathcal{U}}'\hat{\mathcal{U}}$ is equal to identity on every input state, by local tomography (or local discriminability) it is the identity on $\mathcal{A}$. A similar argument proves that $\hat{\mathcal{U}}\hat{\mathcal{U}}'$ is the identity on $\mathcal{A}'$ (Eq. (6.60)). In fact,

$$\hat{\mathcal{U}}\hat{\mathcal{U}}'\left|\hat{\sigma}\right) = \hat{\mathcal{E}}'\hat{\mathcal{D}}'\left|\hat{\sigma}\right) = \left|\hat{\sigma}\right) . \tag{6.266}$$

Then, the two systems are operationally equivalent. We can say in conclusion that to impose a code is to perform a physical operation on a system for generating sequences that can be mapped to that of another system, for instance the apparatus.

Note that ideal compression is equivalent to the existence of a minimal purification, i.e. a purification where the marginal state of the purifying system is completely mixed.

6.4.6 *Purification*

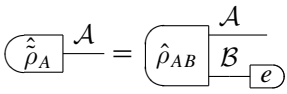
The Principle

Now we come the final assumption (*purification*), the only pure quantum mechanical requirement.[105] I recall that a purification of a state of a certain system $\mathcal{A}$ is a pure state of some composite system $\mathcal{AB}$, with the property that the former is the marginal (the result of a partial trace) of such a pure state (see e.g. Sects. 6.3.2 and 6.4.1). The purification assumption states that every state can be obtained as the marginal of a pure (bipartite) state in an essentially unique way and can be expressed as: *every state can be purified*. Informally speaking, such a postulate states that the ignorance about a part is always compatible with a maximal knowledge of the whole.

The existence of pure bipartite states with mixed marginal was already recognised by Schrödinger as the characteristic trait of quantum theory since it is rooted in quantum features (Sects. 3.1.2, 3.2.1, 3.3.3, and 5.1.2). In particular, we can purify a mixed state thanks to (i) the fact that the purifying system shows features and (ii) the terms of such a system are classically correlated with the terms of the system in mixed state, as it is evident from Eqs. (1.407)–(1.409).

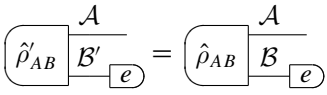
Proof of Purification Uniqueness

Let us show that purification is unique. For every system $\mathcal{A}$ and for every state $\left|\hat{\rho}_A\right)$, there exists a system $\mathcal{B}$ and a pure state $\left|\hat{\rho}_{AB}\right)$ such that (see also circuit (6.194))



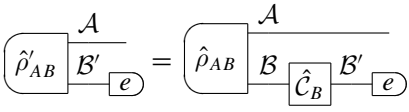
(6.267)

Now, for a fixed purifying system, every two purifications of the same state are connected by a reversible transformation on the purifying system. Indeed, suppose that $\mathcal{B}$ and $\mathcal{B}'$ are both purifications of $\mathcal{A}$:



(6.268)

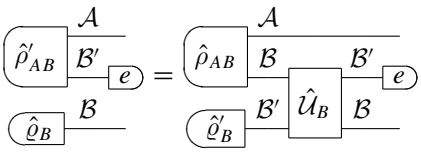
Then, there exists a deterministic transformation $\hat{\mathcal{C}}_B$, acting only on $\mathcal{B}$, such that



(6.269)

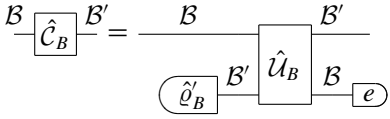
Thus, every state has an essentially unique purification: a purification of a state also purifies any of its refinements. In fact, the uniqueness of purification is sufficient condition of

[105] D'Ariano et al. (2017, Sect. 2.10 and Chap. 7).



(6.270)

for some reversible operation $\hat{\mathcal{U}}_B$ (that could be the swap gate (4.28)). As a consequence, the transformation $\hat{\mathcal{C}}_B$ is some deterministic transformation on $\mathcal{B}$ defined by



(6.271)

This makes of quantum theory the only *physical* theory of information as far as the ignorance of an observer is compatible with the maximal knowledge of Nature as postulated by science. This is another way to say that science does not reduce to observation (Sect. 5.1.2).

State-Transformation Isomorphism

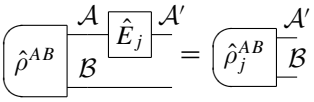
If the pure state $\left|\hat{\rho}_{AB}\right)$ satisfies the relation[106] (see Eq. (5.151) and circuit (6.184))

$$\left(e_A|\hat{\rho}_{AB}\right) = \left|\hat{\varrho}_B\right) , \tag{6.272}$$

where $\left|\hat{\varrho}_B\right)$ is the marginal of $\left|\hat{\rho}_{AB}\right)$, then the latter is an extension of the former. The reciprocal is also true: all the bipartite states that are extensions of $\left|\hat{\varrho}_B\right)$ are of the form $\left|\hat{\rho}_{AB}\right)$ such that there is a transformation having the form of the LHS of Eq. (6.272). Then, there is state-transformation isomorphism.

No Information Acquisition, no Disturbance

Thanks to purification, we can invert the kind of question of Sect. 6.2.2, and show that every measurement that does not extract information about the states that are compatible with a given state, that is in the set $F_{\hat{\rho}}$ (Sect. 6.2.1), can be implemented without disturbing those input states.[107] First, let us prove that every measurement can be implemented in a test where each transformation is pure. Thanks to the state–transformation isomorphisms there must exist a test $\{\hat{E}_j\}$ for the system $\mathcal{A}$ being in the state $\left|\hat{\varrho}_A\right)$, such that



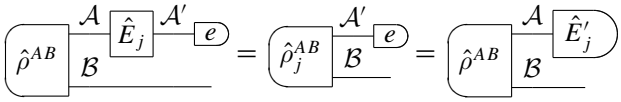
(6.273)

where $\left|\hat{\rho}_{AB}\right)$ is some purification of $\left|\hat{\varrho}_A\right)$. Since each $\left|\hat{\rho}^{AB}_j\right)$ is pure and the state $\left|\hat{\rho}^{AB}\right)$ is faithful, each of these transformations is also pure. Moreover, applying the

[106] D'Ariano et al. (2017, Sect. 7.11).

[107] D'Ariano et al. (2017, Sects. 10.2–10.3).

deterministic effect on $\mathcal{A}'$, we obtain

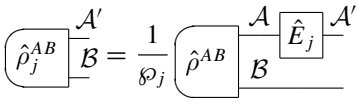


(6.274)

or $\left(\hat{E}'_j \,|\hat{\rho}^{AB}\right) = (e_{A'} \,|\, \hat{E}_j \,|\, \hat{\rho}^{AB})$. Suppose now that the measurement $\{\hat{E}'_j\}$ extract no information upon the input state $|\,\hat{\varrho}_A)$, i.e.

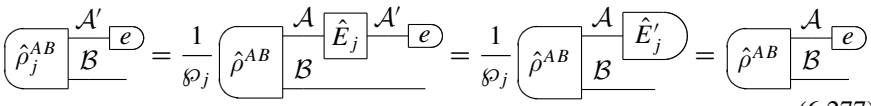
$$\left(\hat{E}'_j \,|\hat{\varrho}_A\right) =_{\hat{\rho}} \wp_j \left(e_A \,|\hat{\varrho}_A\right) \tag{6.275}$$

for all outcomes j. Now, for each j define the state



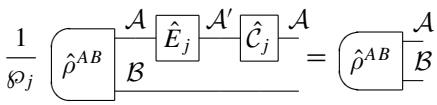
(6.276)

The state $\left|\hat{\rho}_j^{AB}\right)$ is pure and its marginal on $\mathcal{B}$ is equal to its marginal taken on $|\,\hat{\rho}^{AB})$. Indeed,



(6.277)

In the last step I have used Eq. (6.275). Then, due to the uniqueness of purification, there must exist a deterministic transformation $\hat{\mathcal{C}}_j$ such that

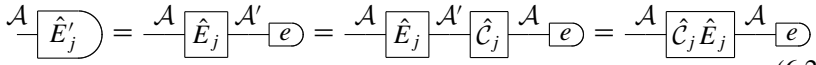


(6.278)

which means

$$\hat{\mathcal{C}}_j \hat{E}_j =_{\hat{\rho}} \wp_j \hat{I}_A, \tag{6.279}$$

which in turn tells us that the test defined by $\hat{\mathcal{C}}_j \hat{E}_j$ is not disturbing upon input $|\,\hat{\varrho}_A)$. This provides the desired realisation of the test $\{\hat{E}'_j\}$:



(6.280)

Quantum Theory and Information

The issue at stake here is not how to distinguish quantum theory from theories that we already know, but to distinguish it from any possible theory. While in the framework of deterministic classical theory, there is no space for information at the fundamental

level, purification establishes the link between physics and information theory. In fact, information theory would make no sense without the notion of probability and mixed states (corresponding to ignorance or incertitude). And it is such a link that allows rigorous calculations even in presence of types and equivalence classes. Now, purification tells us that "ignorance is physical". When we discard a system by performing a partial trace, the remaining system represents a"physical token of our ignorance". Note that possible theories that satisfy purification but have no entanglement, are such that no mixed state can exist, and therefore no random event can occur since random events can be used to generate mixed states (Sect. 5.2.4). In other words, it is purification with entanglement to make here the difference and to constitute QM as a specific and irreducible theory.

Final Considerations

We have shown that quantum theory, differently from any other chapter of physics, can be well axiomatised, with purely mathematical axioms containing no physical primitive. So, in a sense, it is as valid as a piece of pure mathematics. This must be contrasted with the mechanical part of the theory, with the old axiomatic, like the so-called "quantisation rules" (Sect. 1.2.2), which are in fact extrapolated and generalised starting from the heuristic argument of the Ehrenfest theorem (see footnote 136, Sect. 1.3.3), which in turn is based on the superseded theory of CM.

6.5 Summary

- Four fundamental no-go theorems of quantum information have been presented: no-cloning, no-deleting, no-programming, no-bit commitment. The fundamental result is that information is conserved in our universe and code alphabets cannot be shared quantum-mechanically.
- Quantum conditional entropies can be negative.
- Although the dynamics of the whole universe or of sufficiently large cluster of systems, is reversible, in fact its evolution is not. Indeed, there is a single transformation that would lead it back while there are potentially infinite ones that bring it to other states.
- There is no information acquisition without disturbance and therefore without involving some kind of event.
- The formalism of path integrals and quantum jumps allows us to describe propagation processes like signals exchanging by integrating multipath dynamics and quantum events.
- Only unsharp observables and quasi-probability distributions allow for informational completeness. From here, the quantum tomography method for extracting information from a quantum state.
- Information acquisition demands signals that determine an accumulation of noise through local dissipation processes.

- Information sources are not causes: in order to information be 'delivered', additional causal constraints are necessary.
- We acquire information always through data and never have access to the source event. Due to the delayed effect of any event, we can reconcile QM with relativity. We need to consider the whole process through which a certain effect reaches us, so that also the context plays a role and we cannot strictly apply a classical deterministic model.
- Any kind of operation can be represented as a channel: they are either deterministic or probabilistic. In the latter case, noise is involved. Noise can be corrected when and only when there is no information delivered to the environment.
- The only allowed operations on information are processing, sharing and selecting information. They can be somehow displaced in the different information protocols. All operations reduce to one of them or to some combination of them.
- This allows us to define a system whatever could, at least in principle, be dealt with by one of these informational operations.
- Information acquisition is a kind of inverse Bayesian inference.
- Selection of information is the only operation that costs energy and is therefore irreversible. Selection comes always at the end of any process of information exchange and its acquisition is an ultimate fact.
- An information triad has been proposed: processor–regulator–decider as well coupling–selecting–inferring. The process displays circularity.
- The whole of QM can be derived out of six informational principles: causality, atomicity of the composition, perfect discriminability, local discriminability, ideal compression, and the specific quantum assumption of purification.

References

Afriat, A., and G. Tarozzi. 2006. Can Nothing Cause Nonlocal Quantum Jumps?. In *Quantum Mechanics: Are There Quantum Jumps?*, A. Bassi et al. (eds.) 3–7. New York: American Institute of Physics, Melville.

Armstrong, David M. 1978. *A Theory of Universals*. Cambridge: Cambridge University Press.

Auletta, Gennaro. 2000. *Foundations and Interpretation of Quantum Mechanics: In the Light of a Critical-Historical Analysis of the Problems and of a Synthesis of the Results*. Singapore: World Scientific; rev. ed. 2001.

Auletta, Gennaro. 2005a. Quantum Information and Inferential Reasoning. *Foundations of Physics* 35: 155–169.

Auletta, Gennaro. 2005b. Quantum Information as a General Paradigm. *Foundations of Physics* 35: 787–815.

Auletta, Gennaro. 2006a. The Problem of Information, in [Auletta 2006b, 109–127].

Auletta, Gennaro (ed.). 2006b. *The Controversial Relations Between Science and Philosophy: A Critical Assessment*. Libreria Editrice Vaticana: Vatican City.

Auletta, Gennaro. 2011. *Cognitive Biology: Dealing with Information from Bacteria to Minds*. Oxford: University Press.

Auletta, G., and S.Y. Wang. 2014. *Quantum Mechanics for Thinkers*. Singapore: Pan Stanford Publishing. https://doi.org/10.4032/9789814411721.

Auletta, G., M. Fortunato, and G. Parisi. 2009. *Quantum Mechanics*. Cambridge: University Press, rev. paperback ed. 2013.
Barnum, H., C.M. Caves, C.A. Fuchs, R. Jozsa, and B. Schumacher. 1996. Noncommuting Mixed States Cannot Be Broadcast. *Physical Review Letters* 76: 2818–2821.
Battail, Gérard. 2014. *Information and Life*. Berlin: Springer.
Bayes, Thomas. 1763. An Essay Toward Solving a Problem in the Doctrine of Chances. *Philosophical Transactions of the Royal Society of London* 53: 370–418.
Beltrametti, E., and M.J. Maczyński. 1991. On a Characterization of Classical and Nonclassical Probabilities. *Journal of Mathematical Physics* 32: 1280–1286.
Bennett, Charles H. 1982. The Thermodynamics of Computation. A Review. *International Journal of Theoretical Physics* 21: 905–940.
Bennett, C.H., and R. Landauer. 1985. The Fundamental Physical Limits of Computation. *Scientific American* 2531: 48–56. https://doi.org/10.1038/scientificamerican0785-48.
Bjorken, J.D., and S.D. Drell. 1964. *Relativistic Quantum Mechanics*. New York: McGraw-Hill.
Bohm, D., and B.J. Hiley. 1993. *The Undivided Universe: An Ontological Interpretation of Quantum Theory*. London: Routledge.
Bokulich, A., and G. Jaeger (eds.). 2010. *Philosophy of Quantum Information and Entanglement*. Cambridge: Cambridge University Press.
Boltzmann, Ludwig. 1896. *Vorlesungen über Gastheorie I*. Leipzig, J. A: Barth.
Byron Jr., F.W., and R.W. Fuller. 1969–70. *Mathematics of Classical and Quantum Physics, 1969–1970*. New York: Dover, 1992.
Caves, C.M., C.A. Fuchs, and R. Schack. 2002. Quantum Probabilities as Bayesian Probabilities. *Physical Review A* 65: 022305.
Cerf, N.J., and C. Adami. 1997. Negative Entropy and Information in Quantum Mechanics. *Physical Review Letters* 79: 5194–5197.
Chibbaro, S., L. Rondoni, and A. Vulpiani. 2014. *Reductionism, Emergence and Levels of Reality: The Importance of Being Borderline*. Heidelberg: Springer. https://doi.org/10.1007/978-3-319-06361-4.
Clayden, J., N. Greeves, S. Warren, and P. Wothers. 2001. Organic Chemistry. Oxford: University Press.
Clifton, R., J. Bub, and H. Halvorson. 2003. Characterizing Quantum Theory in Terms of Information-Theoretic Constraints. *Foundations of Physics* 33: 1561–1591.
Cohen-Tannoudji, Gilles. 1991. *Les Constantes Universelles*. Paris: Hachette.
Collins, Harry. 2010. *Tacit and Explicit Knowledge*. Chicago: University of Chicago Press.
D'Ariano, G. Mauro. 2010. Probabilistic theories: What is special about Quantum Mechanics?. in [Bokulich/Jaeger 2010, 85–126].
D'Ariano, G.M., and A. Tosini. 2013. Emergence of Space-Time from Topologically Homogeneous Causal Networks. *Studies in History and Philosophy of Modern Physics* 44: 294–299. https://doi.org/10.1016/j.shpsb.2013.04.003.
D'Ariano, G.M., and H.P. Yuen. 1996. Impossibility of Measuring the Wave Function of a Single Quantum System. *Physical Review Letters* 76: 2832–2835.
D'Ariano, M., G. Chiribella, and P. Perinotti. 2017. *Quantum Theory from First Principles: An Informational Approach*. Cambridge: University Press.
Deutsch, David. 2011. *The Beginning of Infinity: Explanations that Transform the World*. London: Penguin Books.
Deutsch, D., and P. Hayden. 2000. Information Flow in Entangled Quantum Systems. *Proceedings of the Royal Society* A456: 17759–17774. https://doi.org/10.1098/rspa.2000.0585.
Dieks, Dennis. 1982. Communication by EPR Devices. *Physics Letters* 92A: 271–272.
Dirac, Paul A.M. 1930. *The Principles of Quantum Mechanics*, 4th ed., Oxford: Clarendon, 1958, 1993.
Einstein, Albert. 1949. Remarks Concerning the Essays Brought Together in This Co-operative Volume, in [Schilpp 1949, 665–688].

Elitzur, Avshalom C. 1992. Locality and Indeterminism Preserve the Second Law. *Physics Letters A* 167: 335–340. https://doi.org/10.1016/0375-9601(92)90268-q.
Feynman, R.P., and A.R. Hibbs. 1965. *Quantum Mechanics and Path Integrals, New York: McGraw-Hill, 1965; emended edition by Daniel F, 2005*. New York, Dover: Styer.
Feynman, Richard P. 1948. Space-Time Approach to Non-Relativistic Quantum Mechanics. *Review of Modern Physics* 20: 367–387.
Fuchs, C.A., and A. Peres. 1996. Quantum-State Disturbance versus Information Gain: Uncertainty Relations for Quantum Information. *Physical Review* A53: 2038–2045.
Geroch, Robert. 1978. *General Relativity from A to B*. Chicago: University of Chicago Press.
Gnedenko, Boris V. 1969. *The Theory of Probability*. Moscow: Mir.
Helstrom, Carl W. 1976. *Quantum Detection and Estimation Theory*. New York: Academic.
Imamo$\bar{g}$lu, A. 1993. Logical Reversibility in Quantum-Nondemolition Measurements. *Physical Review* A47: R4577–R4580.
Ionicioiu, R., and D.R. Terno. 2011. Proposal for a Quantum Delayed-Choice Experiment. *Physical Review Letters* 107: 230406.
Kant, A. 1968. *Kants Werke. Akademie Textausgabe*. Berlin: W. de Gruyter.
Kant, Immanuel. 1787. *Kritik der reinen Vernunft*, 2nd ed. III v. of [Kant A].
Kuhlmann, Meinard. 2013. What is Real? *Scientific American* 8: 32–39.
Landauer, Rolf. 1961. Irreversibility and Heat Generation in the Computing Process. *IBM Journal of Research and Development* 5: 183–191.
Landauer, Rolf. 1996. Minimal Energy Requirements in Communication. *Science* 272: 1914–1919.
Lloyd, Seth. 2000. Ultimate Physical Limits to Computation. *Nature* 406: 1047–1054. https://doi.org/10.1038/35023282.
Lloyd, S., and Y.J. Ng. 2004. Black Hole Computers. *Scientific American* 291 (5): 52–61. https://doi.org/10.1038/scientificamerican1104-52.
Lo, H.-K., and H. Chau. 1997. Is Quantum Bit Commitment Really Possible? *Physical Review Letters* 78: 3410–3413. https://doi.org/10.1103/PhysRevLett.78.3410.
Mabuchi, H., and P. Zoller. 1996. Inversion of Quantum Jumps in Quantum Optical Systems Under Continuous Observation. *Physical Review Letters* 76: 3108–3111.
Mayers, Dominic. 1997. Unconditionally Secure Quantum Bit Commitment is Impossible. *Physical Review Letters* 78: 3414–3417. https://doi.org/10.1103/PhysRevLett.78.3414.
Mensky, Michael B. 1996. A Note on Reversibility of Quantum Jumps. *Physics Letters* 222A: 137–140.
Nielsen, M.A., and C.M. Caves. 1997. Reversible Quantum Operations and Their Application to Teleportation. *Physical Review* A55: 2547–2556.
Nielsen, M.A., and I.L. Chuang. 1997. Programmable Quantum Gate Arrays. *Physical Review Letters* 79: 321–324. https://doi.org/10.1103/PhysRevLett.79.321.
Pati, A.K., and S.L. Braunstein. 2000. Impossibility of Deleting an Unknown Quantum State. *Nature* 404: 164–165.
Pauli, Wolfgang. 1980. *General Principles of Quantum Mechanics*. Berlin: Springer, engl. trans. of [Pauli 1933].
Peruzzo, A., P. Shadbolt, N. Brunner, S. Popescu, and J.L. O'Brien. 2012. *A Quantum Delayed Choice Experiment. Science* 338: 634–637.
Pfister, C., and S. Wehner. 2013. An information-Theoretic Principle implies that any Discrete Physical Theory is Classical. *Nature Communications* 41851: https://doi.org/10.1038/ncomms2821.
Poincaré, O. 1928. *OEuvres*. Paris: Gauthier-Villars.
Poincaré, Henri. 1890. Sur le problème des trois corps et les équations de la dynamique, *Acta Mathematica* 13: 1–270; rep. in [Poincaré O, VII, 262–490].
Poincaré, Henri. 1902. *La science et l'hypothèse, Paris*. Flammarion.
Prugovečcki, Eduard,. 1977. Information-Theoretical Aspects of Quantum Measurement. *International Journal of Theoretical Physics* 16: 321–333.
Prugovečki, Eduard. 1971. *Quantum Mechanics in Hilbert Space*. 2nd ed.: New York: Academic Press, 1981.

Royer, Antoine. 1989. Measurement of Quantum States and the Wigner Function. *Foundations of Physics* 19: 3–30.
Schilpp, Arthur (ed.). 1949. *Albert Einstein. Philosopher-Scientist*. La Salle (Illinois): Open Court, 1949, 3d ed. 1988.
Schlosshauer, Maximilian. 2007. *Decoherence and the Quantum-to-Classical Transition*. Berlin: Springer.
Schumacher, Benjamin W. 1990. Information from Quantum Measurements, in [Zurek 1990, 29–37].
Schumacher, Benjamin W. 1991. Information and Quantum Nonseparability. *Physical Review* A44: 7047–7052.
Shannon, Claude E. 1948. A Mathematical Theory of Communication. *Bell System Technical Journal* 27 (379–423): 623–656.
Singh, P., S.D. Joshi, R.K. Patney, and K. Saha. 2017. The Fourier Decomposition Method for Nonlinear and Nonstationary Time Series Analysis. *Proceedings of the Rpoytal Society A* 473: 20160871. https://doi.org/10.1098/rspa.2016.0871.
Timpson, Christopher G. 2013. *Quantum Information Theory and the Foundations of Quantum Mechanics*. Oxford: University Press.
Ueda, M., and M. Kitagawa. 1992. Reversibility in Quantum Measurement Processes. *Physical Review Letters* 68: 3424–3427.
Wheeler, John A. 1983. Law Without Law, in [Wheeler/Zurek 1983, 182–213].
Wheeler, J.A., and W. Zurek (eds.). 1983. *Quantum Theory and Measurement*. University Press: Princeton.
Wiener, Norbert. 1948. *Cybernetics, or Control and Communication in the Animal and in the Machine*. 2nd ed. Cambridge: MIT Press, 1961, 1965.
Wigner, Eugene P. 1932. On the Quantum Correction for Thermodynamic Equilibrium. *Physical Review* 40: 749–759.
Wootters, W.K., and W.H. Zurek. 1982. A Single Quantum Cannot Be Cloned. *Nature* 299: 802–803.
Yuen, Horace P. 1976. Two-Photon Coherent States of the Radiation Field. *Physical Review* A13: 2226–2243.
Zurek (ed.). 1990. *Complexity, Entropy and the Physics of Information*. Redwood City: Addison-Wesley.

Chapter 7
Quantum Mechanics, Fields and Geometry



Although information is able to explain the fundamental character of QM, its quantumness, it remains also true that, when we deal with the physics of the real world, there are several notions like mass, energy, electric charge and so on that cannot be directly derived from those principles. And, although huge progress have been done in this field (and further ones are expected), it remains true that most of those notions have a classical origin (Sect. 1.3.2). Already Poincaré developed at his time a seminal analysis on the lack of foundations of physical sciences and the conventionality of their basic notions.[1] Most of our assumptions are even disguised definitions, according to the great French scholar. We can say that QM has even added additional difficulties to these already difficult problems.

On the other hand, if QM is the basic physical theory, and as a consequence, the framework developed in the previous chapter is rigorous, then this theoretical construct must be sufficiently robust to integrate neighbouring physical theories. I shall pursue here two main goals. First, I shall show very schematically that QM is naturally extended to fields (for both force carriers and matter particles) and to basic matter structures. Second, that general relativity (that I shall present in a very sketchy form) and QM can be made compatible. This will demand a certain theoretical effort.

The chapter is divided into three major subjects: (1) fields (classical and quantum mechanical), with special focus on electromagnetic field, and basic constituents of matter, (2) gravity (introduction to both general relativity and quantum gravity), (3) structure of matter (atoms and molecules) and statistical mechanics.

[1] Poincaré (1893, 1897), Poincaré (1902, Chaps. 6–8). The same in Einstein (1949, p. 12).

© Springer Nature Switzerland AG 2019

G. Auletta, *The Quantum Mechanics Conundrum*,
https://doi.org/10.1007/978-3-030-16649-6_7

7.1 Fields, Matters, Geometry

7.1.1 Classical Electromagnetic Field

The classical electromagnetic theory was built essentially during the 19th century, thanks to the contributions of the French physicist Charles-Augustin de Coulomb (1736–1806), the British physicist and chemist Michael Faraday (1791–1867) and the already mentioned James C. Maxwell. However, its roots go far back and the first contributions to the theory of the magnetic field and to the electric force are due to two jesuits of the *Collegio Romano*, the German polymath Athanasius Kircher (1602–1680) and the already mentioned Rogerius Boscovich.[2]

Preliminary Mathematical Notions

The Green's theorem for the plane, after the name of the mathematician George Green, tells us that a double integral on a (plane) region R for functions P, Q of variables x, y, is given by

$$\iint_R \left(\frac{\partial P}{\partial x} - \frac{\partial Q}{\partial y}\right) dxdy = \oint_C Pdx + Qdy, \tag{7.1}$$

where C is the closed curve representing the contour of R that is followed by CCW. Note that the contour integral is the same whatever deformation we perform. In fact, each 'positive' (outward) deviation from a perfect circle is compensated by a corresponding 'negative' (inward) deviation so that we finally get a circumference. Note also that, for any $z \in \mathbb{C}$, we have[3]

$$\oint \frac{1}{z} dz = 2\pi i, \tag{7.2}$$

which in fact correspond to an integral of the complete circle in Fig. 4.5, Sect. 4.2.3. Assuming a vector field

$$\mathbf{F}(x, y, z) = P(x, y, z)\mathbf{e}_x + Q(x, y, z)\mathbf{e}_y + R(x, y, z)\mathbf{e}_z, \tag{7.3}$$

its *divergence* is given by the scalar product (defining a scalar field) between the nabla vector (1.155) and $\mathbf{F}$

$$\boldsymbol{\nabla} \cdot \mathbf{F} = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}, \tag{7.4}$$

[2] Kircher (1643, 1667), Boscovich (1745, 1763). For a reconstruction of the latter contributions see Auletta (2015a).

[3] Byron and Fuller (1969–70, II, p. 325).

in agreement with rule (1.63).[4] The divergence expresses the magnitude of a vector field's source or sink at a given point, or also the volume density of the outward flux of a vector field through an infinitesimal surface per unit volume. We have already met the divergence of a gradient (1.159). On the other hand, the *curl* of the vector field (7.3) is a new vector field expressed by the cross product

$$\boldsymbol{\nabla} \times \mathbf{F} = \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} \right) \mathbf{e}_x + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x} \right) \mathbf{e}_y + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) \mathbf{e}_z, \tag{7.5}$$

in accordance with the rule (1.220). The curl is a vector operator that describes the infinitesimal rotation of the 3D vector field **F**. At every point in the field, the curl of that point is represented by a vector and the length and direction of this vector characterise the rotation at that point. Helmholtz's theorem, after the name of Hermann von Helmholtz, also known as the fundamental theorem of vector calculus, states that any sufficiently smooth, rapidly decaying vector field **F** in three dimensions can be resolved into the sum of an irrotational (curl free) vector field and a solenoidal (divergence free) vector field. Since an irrotational vector field has a scalar potential φ and a solenoidal vector field has a vector potential **A**, in appropriate conditions, we have

$$\mathbf{F} = -\nabla\varphi + \boldsymbol{\nabla} \times \mathbf{A}. \tag{7.6}$$

This is also known as the *Helmholtz decomposition*. Thus, the curl of the gradient (1.156) of a scalar field $\varphi(x, y, z)$ is zero:

$$\boldsymbol{\nabla} \times (\boldsymbol{\nabla}\varphi) = 0. \tag{7.7a}$$

Reciprocally, the divergence of a curl is also zero:

$$\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \times \mathbf{A}) = 0. \tag{7.7b}$$

Very important in such a context is the *Stokes' theorem*, after the name of the Irish–British mathematician, physicist and theologian George G. Stokes (1819–1903): if the closed line l represents a CCW circuit being the contour of an open surface S and $\mathbf{F}(x, y, z)$ some vector field, we have that

$$\oint_l \mathbf{F} \cdot d\mathbf{l} = \int_S \boldsymbol{\nabla} \times \mathbf{F} \cdot dS, \tag{7.8}$$

where $dS = \mathbf{n}da$ for some vector **n** and small surface element da. The previous equation makes clear that the curl of **F** does not depend on the reference frame but is a property intrinsic to the vectorial field **F**. If we take an infinitesimal element dS of the surface and the vector **n** perpendicular to it (and oriented according to the right-hand rule), the Stokes' theorem applied to dS reads

[4] Byron and Fuller (1969–70, I, Sect. 1.7).

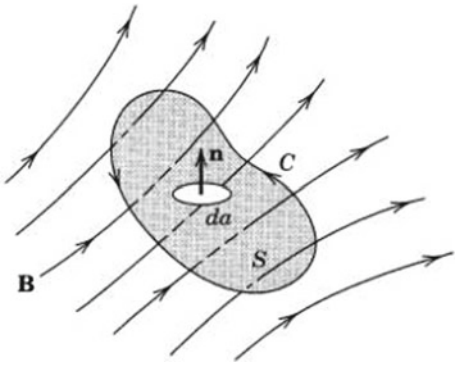


Fig. 7.1 Magnetic flux through an open surface **S**. Adapted from Jackson (1962, p. 210)

$$\nabla \times \mathbf{F} \cdot \mathbf{n} = \lim_{dS \to 0} \frac{1}{dS} \oint_l \mathbf{F} \cdot d\mathbf{l}. \tag{7.9}$$

Since the scalar product of two vectors (here $\nabla \times \mathbf{F}$ and $\mathbf{n}$) is invariant relative to the reference frame, the component of $\mathbf{F} \cdot d\mathbf{l}$ in the $\mathbf{n}$ direction is also invariant.

Maxwell's Equations

Classically, the *electric vector field* $\mathbf{E}$, in the first approximation, is a function of the negative gradient $-\nabla\varphi$ of a scalar potential φ, in agreement with Eq. (7.6). Assuming for simplicity absence of sources, for the divergence of the *magnetic field* $\mathbf{B}$, we have

$$\nabla \cdot \mathbf{B} = 0, \tag{7.10}$$

which is the *first Maxwell equation*, after the name of James C. Maxwell.[5] From this and Eq. (7.7b), we infer that $\mathbf{B}$ must be the curl of some vector field $\mathbf{A}(\mathbf{r})$, the *vector potential*,[6]

$$\mathbf{B}(\mathbf{r}) = \nabla \times \mathbf{A}(\mathbf{r}). \tag{7.11}$$

Similarly, our assumption of absence of sources for the electric field gives

$$\nabla \cdot \mathbf{E} = 0, \tag{7.12}$$

which is the *second Maxwell equation* (also called Coulomb law, after the name of the French physicist C. A. de Coulomb).

Let us now consider Fig. 7.1. The magnetic flux through a surface S is given by

[5] Maxwell (1873).

[6] See Jackson (1962, Sect. 5.4). Traditionally, one used to distinguish between the magnetic field $\mathbf{H}$ and the magnetic induction $\mathbf{B} = \mu_0\mathbf{H}$, where μ_0 is the magnetic permeability. Modern approaches consider $\mathbf{B}$ as the basic quantity, sometimes called magnetic field. I follow this convention.

$$F_M = \int_S \mathbf{B} \cdot \mathbf{n}\, da, \tag{7.13}$$

where da is a small element of S and $\mathbf{n}$ is some direction, while the *electromotive force* (i.e. the voltage developed by any source of electrical energy) produced around the circuit is given by

$$\mathcal{E} = \oint_C \mathbf{E} \cdot d\mathbf{l}, \tag{7.14}$$

where $\mathbf{E}$ is the classical electric field at the element $d\mathbf{l}$ of the circuit.[7] In 1831, the British physicist and chemist Michael Faraday inferred from his observations that

$$\mathcal{E} = -k \frac{dF_M}{dt}, \tag{7.15}$$

where the constant k depends on the chosen units for involved quantities. Using Eq. (7.13), we can write in general way the *Faraday's law* (also known as Faraday's law of induction) as

$$\oint_C \mathbf{E} \cdot d\mathbf{l} = -k \frac{d}{dt} \int_S \mathbf{B} \cdot \mathbf{n}\, da. \tag{7.16}$$

Thus, the induced electromotive force in any closed circuit is proportional to the negative of the total time derivative of the magnetic flux enclosed by the circuit. Galilean invariance (Sect. 2.3.2) requires that the constant k be equal to c^{-1}, i.e. the inverse of the vacuum speed of light, so that, from Eq. (7.16), we get

$$\oint_C \mathbf{E} \cdot d\mathbf{l} = -\frac{1}{c} \frac{d}{dt} \int_S \mathbf{B} \cdot \mathbf{n}\, da. \tag{7.17}$$

Moreover, we have

$$c = \frac{1}{\sqrt{\epsilon_0 \mu_0}}, \tag{7.18}$$

where ϵ_0 and μ_0 are the vacuum electric permittivity (a measure of how an electric field affects, and is in turn affected by, a dielectric medium (insulator)) and magnetic permeability (the degree of magnetisation of a material in response to a magnetic field), respectively.

If the circuit is moving with a velocity $\mathbf{v}$ in some direction, the flux through the circuit may change because (i) the flux changes with time at a point or (ii) the translation of the circuit changes the location of the boundary. Then, the RHS of Eq. (7.17) is composed of the following two terms:

$$\frac{d}{dt} \int_S \mathbf{B} \cdot \mathbf{n}\, da = \int_S \frac{\partial \mathbf{B}}{\partial t} \cdot \mathbf{n}\, da + \oint_C (\mathbf{B} \times \mathbf{v}) \cdot d\mathbf{l}, \tag{7.19}$$

[7] See Jackson (1962, Sects. 6.1–6.3).

which allows us to rewrite Eq. (7.17) as

$$\oint_C \left[\mathbf{E} + \frac{1}{c}(\mathbf{B} \times \mathbf{v})\right] \cdot d\mathbf{l} = -\frac{1}{c}\int_S \frac{\partial \mathbf{B}}{\partial t} \cdot \mathbf{n}\, da. \tag{7.20}$$

This is an equivalent statement of Faraday's law applied to a moving circuit C, so that we can think of the circuit C and the surface S in Eq. (7.17) as being at a certain position in space instantaneously. Using the Stokes' theorem (7.8) and provided that the circuit is held fixed in the chosen reference frame (in order to have $\mathbf{E}$ and $\mathbf{B}$ defined in the same frame), the transformation of the electromotive force (7.17) in a surface integral leads us to

$$\int_S \left(\nabla \times \mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{B}}{\partial t}\right) \cdot \mathbf{n}\, da = 0, \tag{7.21}$$

and the differential form that the Faraday's law (7.16) takes in vacuum under the simplified assumption of no charges and no currents is

$$\nabla \times \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t}. \tag{7.22}$$

This is the *third Maxwell equation*. In a similar way, we can derive

$$\nabla \times \mathbf{B} = \frac{1}{c^2}\frac{\partial \mathbf{E}}{\partial t}, \tag{7.23}$$

which is the *fourth Maxwell equation*.

Gauge Symmetry

The properties of the electric and magnetic fields $\mathbf{E}$ and $\mathbf{B}$ can be extracted from the vector potential $\mathbf{A}$ and the scalar potential φ thanks to the relations[8]

$$\mathbf{E} = -\nabla\varphi - \frac{\partial}{\partial t}\mathbf{A}\,, \tag{7.24a}$$

$$\mathbf{B} = \nabla \times \mathbf{A}\,, \tag{7.24b}$$

where the first equation is derived from Eq. (7.6), while the second one is Eq. (7.11). Since, thanks to the (7.7a),

$$\nabla \times \mathbf{E} = \nabla \times \left(-\frac{\partial}{\partial t}\mathbf{A}\right), \tag{7.25}$$

[8]For the remainder of the present Sect. I follow (Auletta et al. 2009, Sect. 13.1.1).

and

$$\frac{\partial}{\partial t}\frac{\partial}{\partial j}f = \frac{\partial}{\partial j}\frac{\partial}{\partial t}f, \tag{7.26}$$

for $j = x, y, z$, where f is a smooth function (having derivative of all orders, like the multivalued exponential and trigonometric functions, which is the case for $\mathbf{A}$, as we shall see), then Eqs. (7.10) and (7.22) are automatically satisfied, whereas Eqs. (7.12) and (7.23) become

$$\nabla \cdot \left(-\nabla\varphi - \frac{\partial}{\partial t}\mathbf{A}\right) = 0\,, \tag{7.27a}$$

$$\nabla \times (\nabla \times \mathbf{A}) = \frac{1}{c^2}\frac{\partial}{\partial t}\left(-\nabla\varphi - \frac{\partial}{\partial t}\mathbf{A}\right)\,, \tag{7.27b}$$

respectively. This change of representation from the fields to the potentials presents a difficulty, since different scalar and vector potentials may lead to the same fields. In particular, since for any scalar function, we have Eq. (7.7a) and $\nabla\partial/\partial t = 0$ by Eq. (1.155), the combined *gauge* transformations for any scalar function $f(\mathbf{r}, t)$

$$\mathbf{A} \longmapsto \mathbf{A} + \nabla f\,, \tag{7.28a}$$

$$\varphi \longmapsto \varphi - \frac{\partial}{\partial t}f \tag{7.28b}$$

do not alter the values of $\mathbf{E}$ and $\mathbf{B}$. Since the latter represent the measurable properties of the electromagnetic field, different potentials that lead to the same fields would describe the same physical situation. Note that, gauge symmetry is not a symmetry of the kind seen in Sect. 1.2.5. In fact, "by a symmetry in the strict sense, we understand a mapping of the space of physical states of a system which respects the dynamics of the system, but maps states to physically distinct states". At the opposite, gauge symmetry is such that "all configurations which are related by symmetry transformations are to be regarded as corresponding to a single physical state. In other words, this kind of symmetry is part of the definition of what one really understands by physical states. A symmetry transformation of this type is thus not a mapping from a physical state to another physical state; rather, it realises the redundancy in the description of the configurations of the system."[9] Recall that electromagnetic waves (light) proceed transversal to the propagation direction. However, quantum fluctuations will explore all possible field configurations, including longitudinal waves. Then, we should ensure that the latter do not represent a real physical effect. Adding and subtracting the latter so that the equations of the electromagnetic field remains unchanged is the business of gauge symmetry. In fact, from the point of view of an observer at the speed of light, the electromagnetic wave appears stationary, and here there would be no distinction between transverse and longitudinal waves (Einstein's original thought about relativity has some connections with this).

[9] Fuchs and Schweigert (1997, pp. 12–13).

Then, if wish to translate Eqs. (7.12) and (7.23) in a univocal way, we have to add an extra constraint to Eqs. (7.24). For our purposes, it is convenient to adopt the so-called Coulomb or radiation gauge, after the name of C. A. de Coulomb, for which

$$\nabla \cdot \mathbf{A} = 0 \, . \tag{7.29}$$

Moreover, the scalar potential φ is a function of the spatial charge distribution and in the case where there are no sources, can be eliminated from the problem with the choice

$$\varphi = 0. \tag{7.30}$$

Then, Eq. (7.24a) becomes

$$\mathbf{E}\,(\mathbf{r}, t) = -\frac{\partial}{\partial t}\mathbf{A}\,(\mathbf{r}, t) \ , \tag{7.31}$$

and Eq. (7.12) is automatically satisfied. On the other hand, Eq. (7.27b) finally reads

$$\nabla^2 \mathbf{A}\,(\mathbf{r}, t) = \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\mathbf{A}\,(\mathbf{r}, t) \ , \tag{7.32}$$

where I have used the mathematical identity

$$\nabla \times (\nabla \times \mathbf{A}) = \nabla\,(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A} \tag{7.33}$$

and Eqs. (7.29)–(7.30). Equation (7.32) shows the important result that the vector potential satisfies the classical wave equation (2.14). It is known that the classical solution of Eq. (7.32) is given by the continuous Fourier expansion (see Eq. (1.182))

$$\mathbf{A} = \frac{1}{2\pi}\int d\mathbf{k} d\omega e^{i\omega t} e^{i\mathbf{k}\cdot\mathbf{r}} \delta(k^2 - c^2\omega^2) \ , \tag{7.34}$$

where $\mathbf{k} = 2\pi/\lambda$ is the (classical) propagation vector (see Eqs. (1.6) and (1.166)) and ω is the angular frequency (5.87). Note that, we have expressed the vector potential as an integral ranging on two components that show a spatial ($\mathbf{k}$) and a temporal (ω) dependency, respectively. The previous expression shows that $\mathbf{A}$ is indeed a smooth function.

7.1.2 Second Quantisation

The background of the second quantisation is in the work on Bose–Einstein statistics (Sect. 1.3.1) and the subsequent development of quantum optics (Sects. 6.2.2 and 6.2.3). Important was also Dirac's contribution,[10] as well as of the German physicist

[10] Dirac (1927).

Pascual Jordan (1902–1980) and the Swedish physicist Oskar Benjamin Klein (1894–1977) for bosons[11]; moreover, I mention the contributions of P. Jordan and E. Wigner for fermions (especially, relevant for quantum-field theory).[12]

Quantised Fields

In order to accomplish the quantisation of the electromagnetic field, it is now required that we replace the classical vector potential $\mathbf{A}$ by a quantum-mechanical operator $\hat{\mathbf{A}}$.[13] In analogy with the classical case (7.34), we then expand the vector potential operator in a discrete Fourier series Eq. (1.182) as

$$\hat{\mathbf{A}}(\mathbf{r},t) = \sum_{\mathbf{k}} c_{\mathbf{k}} \left[\hat{a}_{\mathbf{k}} \mathbf{u}_{\mathbf{k}}(\mathbf{r}) e^{-i\omega_{\mathbf{k}} t} + \hat{a}^{\dagger}_{\mathbf{k}} \mathbf{u}^{*}_{\mathbf{k}}(\mathbf{r}) e^{i\omega_{\mathbf{k}} t} \right] , \tag{7.35}$$

where the $c_{\mathbf{k}}$'s are constants to be determined, the dimensionless amplitudes are now quantum-mechanical operators $\hat{a}_{\mathbf{k}}$, $\hat{a}^{\dagger}_{\mathbf{k}}$ and $\mathbf{u}_{\mathbf{k}}(\mathbf{r})$, which is proportional to $e^{i\mathbf{k}\cdot\mathbf{r}}$, is a discrete set of orthogonal mode functions (Fourier components of the field) that, due to Eqs. (7.32) and (7.29), have to satisfy

$$\left(\nabla^2 + \frac{\omega^2_{\mathbf{k}}}{c^2} \right) \mathbf{u}_{\mathbf{k}}(\mathbf{r}) = \left(\nabla^2 + \mathbf{k}^2 \right) \mathbf{u}_{\mathbf{k}}(\mathbf{r}) = 0 , \tag{7.36a}$$

$$\boldsymbol{\nabla} \cdot \mathbf{u}_{\mathbf{k}}(\mathbf{r}) = 0 , \tag{7.36b}$$

where for simplicity I have assumed here that $\mathbf{u}_{\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}$ and I have made use of the fact that, from Eqs. (1.6) and (5.87), we have $\omega_{\mathbf{k}} = c|\mathbf{k}|$. If, for simplicity, we choose to quantise the free field inside a cube of side l (the so-called *cavity* or electromagnetic resonator) with periodic boundary conditions, the wave (propagation) vector $\mathbf{k}$ has components

$$k_x = \frac{2\pi n_x}{\mathrm{l}} , \quad k_y = \frac{2\pi n_y}{\mathrm{l}} , \quad k_z = \frac{2\pi n_z}{\mathrm{l}} , \tag{7.37}$$

where n_x, n_y and n_z are integers. Notice that in many applications, l will tend to infinity. Moreover, the mode functions may be expressed as plane waves, i.e.

$$\mathbf{u}_{\mathbf{k}}(\mathbf{r}) = \frac{\mathbf{e}}{\mathrm{l}^{\frac{3}{2}}} e^{i\mathbf{k}\cdot\mathbf{r}} , \tag{7.38}$$

[11] Jordan and Klein (1927).

[12] Jordan and Wigner (1928).

[13] For this Sect. I follow (Auletta et al. 2009, Sect. 13.1.2–13.1.3). For mathematical aspects, see Byron and Fuller (1969–70, I, Sect. 1.7).

where the polarisation vector $\mathbf{e}$ must satisfy the condition $\mathbf{e} \cdot \mathbf{k} = 0$ as a consequence of the gauge condition (7.36b):

$$\boldsymbol{\nabla} \cdot \mathbf{u}_{\mathbf{k}}(\mathbf{r}) = \frac{i}{1^{\frac{3}{2}}} e^{i\mathbf{k}\cdot\mathbf{r}}\, \mathbf{e} \cdot \mathbf{k} = 0. \tag{7.39}$$

This implies, as expected, that $\mathbf{e}$ is always orthogonal to the propagation direction, and therefore the polarisation vector has two independent directions. In other words, this confirms the transverse nature of the electromagnetic waves (Fig. 1.2, Sect. 1.1), so that we have, in fact, the unit polarisation vectors $\mathbf{e}_\lambda$ ($\lambda = 1, 2$) for which

$$\mathbf{k} \cdot \mathbf{e}_\lambda = 0\,, \tag{7.40a}$$

$$\mathbf{e}_\lambda \cdot \mathbf{e}_{\lambda'} = \delta_{\lambda\lambda'}\,. \tag{7.40b}$$

Then, taking into account Eqs. (7.40), we rewrite the mode functions as

$$\mathbf{u}_{\mathbf{k},\lambda}(\mathbf{r}) = \frac{\mathbf{e}_\lambda}{1^{\frac{3}{2}}} e^{i\mathbf{k}\cdot\mathbf{r}}\,, \tag{7.41}$$

as well as the operator $\hat{\mathbf{A}}$ as

$$\hat{\mathbf{A}}(\mathbf{r}, t) = \sum_{\mathbf{k},\lambda} c_{\mathbf{k}} \left[\hat{a}_{\mathbf{k},\lambda} \mathbf{u}_{\mathbf{k},\lambda}(\mathbf{r}) e^{-i\omega_{\mathbf{k}} t} + \hat{a}^\dagger_{\mathbf{k},\lambda} \mathbf{u}^*_{\mathbf{k},\lambda}(\mathbf{r}) e^{i\omega_{\mathbf{k}} t} \right]. \tag{7.42}$$

A complete calculation shows that the normalisation factors have to be written as

$$c_{\mathbf{k}} = \sqrt{\frac{\hbar}{2\omega_{\mathbf{k}} \epsilon_0}}\,, \tag{7.43}$$

where I recall that ϵ_0 is the electric permitittivity; this, together with Eqs. (7.24b), (7.31), (7.32), and (7.42), allows us to express the electric and magnetic fields as the Hermitian operators[14]

$$\hat{\mathbf{E}}(\mathbf{r}, t) = -\frac{\partial}{\partial t}\hat{\mathbf{A}} = i \sum_{\mathbf{k},\lambda} \left(\frac{\hbar\omega_{\mathbf{k}}}{2 1^3 \epsilon_0} \right)^{\frac{1}{2}} \left[\hat{a}_{\mathbf{k},\lambda} e^{i(\mathbf{k}\cdot\mathbf{r} - \omega_{\mathbf{k}} t)} - \hat{a}^\dagger_{\mathbf{k},\lambda} e^{-i(\mathbf{k}\cdot\mathbf{r} - \omega_{\mathbf{k}} t)} \right] \mathbf{e}_\lambda, \tag{7.44a}$$

$$\hat{\mathbf{B}}(\mathbf{r}, t) = \boldsymbol{\nabla} \times \hat{\mathbf{A}} = i \sum_{\mathbf{k},\lambda} \left(\frac{\hbar k}{2 c 1^3 \epsilon_0} \right)^{\frac{1}{2}} \left[\hat{a}_{\mathbf{k},\lambda} e^{i(\mathbf{k}\cdot\mathbf{r} - \omega_{\mathbf{k}} t)} - \hat{a}^\dagger_{\mathbf{k},\lambda} e^{-i(\mathbf{k}\cdot\mathbf{r} - \omega_{\mathbf{k}} t)} \right] \mathbf{b}_\lambda, \tag{7.44b}$$

respectively, thanks to the four components (Eq. (7.5))

[14] See Auletta et al. (2009, Problem 13.4) for details.

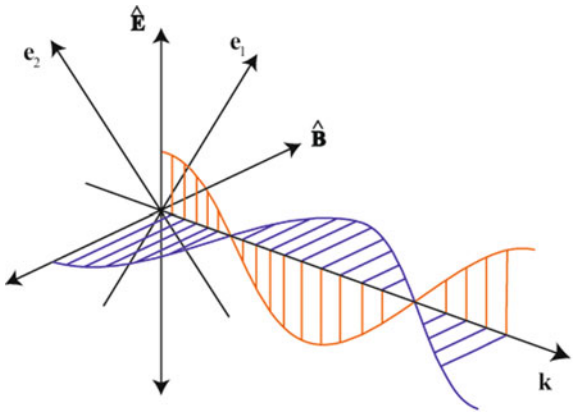


Fig. 7.2 The three principal directions of the electromagnetic field (the electric and magnetic fields oscillate along orthogonal directions): for simplicity, I assume that $\mathbf{e}_1$ is parallel to the x- direction, $\mathbf{e}_2$ is parallel to the y-direction, $\mathbf{k}$ is in the z-direction ($k_x = k_y = 0$), and $\mathbf{b}_1 = \mathbf{e}_2$, $\mathbf{b}_2 = -\mathbf{e}_1$. Adapted from Auletta et al. (2009, p. 460)

$$\begin{aligned}\boldsymbol{\nabla} \times \left(e^{\pm i\mathbf{k}\cdot\mathbf{r}}\mathbf{e}_1\right) &= \pm i\left(k_z e^{\pm i\mathbf{k}\cdot\mathbf{r}}\mathbf{e}_y - k_y e^{\pm i\mathbf{k}\cdot\mathbf{r}}\mathbf{e}_z\right), \\ \boldsymbol{\nabla} \times \left(e^{\pm i\mathbf{k}\cdot\mathbf{r}}\mathbf{e}_2\right) &= \mp i\left(k_z e^{\pm i\mathbf{k}\cdot\mathbf{r}}\mathbf{e}_x - k_x e^{\pm i\mathbf{k}\cdot\mathbf{r}}\mathbf{e}_z\right),\end{aligned} \tag{7.45}$$

where

$$\mathbf{b}_\lambda = \mathbf{k} \times \mathbf{e}_\lambda \tag{7.46}$$

is a unit vector whose direction is orthogonal both to $\mathbf{k}$ and to $\mathbf{e}_\lambda$, in agreement with the fact that electric and magnetic fields oscillate along orthogonal directions, both remaining orthogonal to the wave vector (Fig. 7.2). Thus, in classical electrodynamics $\mathbf{E}(\mathbf{r}, t)$ and $\mathbf{B}(\mathbf{r}, t)$, representing the measurable properties of the electromagnetic field, are real vectors, while their quantum counterparts (7.44), as expected, become Hermitian operators.

Creation and Annihilation Operators

As said, quantisation requires that $\hat{a}_\mathbf{k}$ and $\hat{a}^\dagger_\mathbf{k}$ be interpreted as quantum-mechanical operators. In order to correctly identify them, we resort to the classical expression for the energy of the electromagnetic field in a resonator, that is,[15]

$$H = \frac{1}{2}\int_{\mathbb{I}^3} d\mathbf{r}\left(\epsilon_0 \mathbf{E}^2 + \frac{\mathbf{B}^2}{\mu_0}\right), \tag{7.47}$$

where I recall that μ_0 is the magnetic permeability. Substituting the quantum expressions (7.44) in the latter equation, we can obtain

[15] Jackson (1962, p. 237).

$$\hat{H} = \sum_{\mathbf{k},\lambda} \hbar\omega_{\mathbf{k}} \left(\hat{a}^{\dagger}_{\mathbf{k},\lambda} \hat{a}_{\mathbf{k},\lambda} + \frac{1}{2} \right) , \tag{7.48}$$

in agreement with Eq. (6.70), so that, if we interpret $\hat{a}_{\mathbf{k},\lambda}$ and $\hat{a}^{\dagger}_{\mathbf{k},\lambda}$ as the annihilation and creation operators (see Eqs. (6.68) and comments), the Hamiltonian of the electromagnetic field is that of a system of independent harmonic oscillators (see Sect. 6.2.2), one for each mode of the field and polarisation direction.

Thus, quantising the electromagnetic field amounts to quantise each of the harmonic oscillators. This is called *second quantisation*. In other words, while first quantisation (ordinary QM) refers to the discrete modes (Fourier components of the field), for example of the harmonic oscillator, second quantisation refers to the integer numbers of excitations of each of these modes: first, we interpret these modes as having energy $E = \hbar\omega$ (see Fig. 6.3, Sect. 6.2.1), and then quantise each mode as a harmonic oscillator: in other words, field quantisation is accomplished by interpreting $\hat{a}_{\mathbf{k},\lambda}$ and $\hat{a}^{\dagger}_{\mathbf{k},\lambda}$ as the annihilation and creation operators of the $\mathbf{k}$th field mode with polarisation direction λ, respectively. Thus, if a scalar field could be represented as a vector according to first quantisation, in the second quantisation it becomes an operator. In agreement with Eq. (6.78a), the annihilation and creation operators of the field obey the commutation relations

$$\left[\hat{a}_{\mathbf{k},\lambda}, \hat{a}_{\mathbf{j},\lambda'}\right] = \left[\hat{a}^{\dagger}_{\mathbf{k},\lambda}, \hat{a}^{\dagger}_{\mathbf{j},\lambda'}\right] = 0 \, , \quad \left[\hat{a}_{\mathbf{k},\lambda}, \hat{a}^{\dagger}_{\mathbf{j},\lambda'}\right] = \delta_{\mathbf{kj}}\delta_{\lambda\lambda'} \, . \tag{7.49}$$

As we know, the creation (annihilation) operator $\hat{a}^{\dagger}_{\mathbf{k},\lambda}$ ($\hat{a}_{\mathbf{k},\lambda}$) applied to a state of definite energy 'adds' ('removes') a quantum of energy $\hbar\omega_{\mathbf{k}}$ to (from) the mode $\mathbf{k}$, λ of the cavity field. These energy quanta need to be interpreted as light quanta or photons, whose wave vector is precisely $\mathbf{k}$. The number of photons in each mode of the cavity, in agreement with the Bose–Einstein statistics (Sect. 1.3.1), is given by the eigenvalue $n_{\mathbf{k},\lambda}$ of the corresponding number operator Eq. (6.69)

$$\hat{N}_{\mathbf{k},\lambda} = \hat{a}^{\dagger}_{\mathbf{k},\lambda} \hat{a}_{\mathbf{k},\lambda} \, . \tag{7.50}$$

In order to specify the total field in the cavity, it is then necessary to indicate the number of photons (or *occupation number*) for each mode: a generic state of the total field can be written as

$$|\{n_{\mathbf{k}}\}\rangle = \left|n_{\mathbf{k}_1}, n_{\mathbf{k}_2}, n_{\mathbf{k}_3}, \ldots\right\rangle = \left|n_{\mathbf{k}_1}\right\rangle \otimes \left|n_{\mathbf{k}_2}\right\rangle \otimes \left|n_{\mathbf{k}_3}\right\rangle \otimes \ldots \, , \tag{7.51}$$

where for the sake of simplicity, I have omitted the λ dependence, and the equality sign is a consequence of the fact that different cavity modes are independent.

Energy and Vacuum State

Finally, I note that the total energy of a cavity field in the state (7.51) is given by

$$E = \sum_{\mathbf{k}} \hbar\omega_{\mathbf{k}} \left(n_{\mathbf{k}} + \frac{1}{2} \right) . \tag{7.52}$$

It should be noted that, according to Eq. (7.52), even in the case when no excitations are present (i.e. when no photon is present in each of the modes $\mathbf{k}$, so that $n_{\mathbf{k}} = 0, \forall \mathbf{k}$), the total energy does not vanish. In fact, the energy of this state, called the *vacuum state*, is

$$E_0 = \frac{\hbar}{2} \sum_{\mathbf{k}} \omega_{\mathbf{k}} , \tag{7.53}$$

and since the sum is infinite and there is no upper bound for the allowed frequencies $\omega_{\mathbf{k}}$, it is also infinite. This energy is called the zero-point energy and constitutes a puzzling feature of quantum electrodynamics. However, if we neglect gravity (since energy has gravitational effects), this feature does not raise particular difficulties for calculations aiming at comparing theoretical results with experiments: practical experiments are only sensitive to *changes* in the total energy of the electromagnetic field for which, of course, the zero-point energy cancels out. Nevertheless, although not affecting the results of the present subsection, the zero-point energy represent a conceptual difficulty in quantum-field theory.

Fock Space

Note that, in the second quantisation, due to Eq. (7.51), the Hilbert space is promoted to a Fock space (after the name of V. Fock), which is defined at each time as the direct sum

$$\mathcal{F} = \oplus_n \mathcal{H}_n \tag{7.54}$$

of the Hilbert space $\mathcal{H}_n$ of n-particle states. If there are different particle types, the Fock space is the direct sum of the Hilbert spaces associated with each particle. The Fock space is the same at all times, by time-translation invariance, and in any frame, by Lorentz invariance.

Invariances

Our universe presents a number of symmetries and the question arises of which ones quantum-field theory respects.[16] As mentioned, the first one is space-time isotropy. Thus, our theory should be translation invariant: if we take all our fields $\varphi(x)$ and

[16] Schwartz (2014, Sect. 8.1).

replace them by $\varphi(x + a)$ for any constant 4 vector a^ν, the observables should look the same. Another symmetry is Lorentz invariance (Sect. 2.3.2): physics should look the same whether we point our measurement apparatus to the left or to the right, or put it on a lab moving at constant velocity. As mentioned, the group of both space and time translations and Lorentz transformations is called the *Poincaré group* (after the name of H. Poincaré) ISO(1, 3): it is the isometry group of Minkowski space. Our universe also presents different types of particles, with some particles having mass and characterised by other quantum numbers. They have momentum and the value of spin projected on some axis. If we rotate or boost to change the frame (Eqs. (2.95)–(2.99)), only the momenta and the spin projection change, as determined by the Poincaré group, but the other quantum numbers do not. In particular, particles transform under irreducible unitary representations of the Poincaré group.

Nevertheless, there is a potential conflict between unitarity and Lorentz invariance. The problem is that, for example in the case of a 2D particle, the boost matrix

$$\Lambda = \begin{bmatrix} \cosh\beta & \sinh\beta \\ \sinh\beta & \cosh\beta \end{bmatrix} \tag{7.55}$$

is not unitary. In particular, there is a conflict between having a Hilbert space with a positive norm, which is a physical requirement leading to the $\delta^{\mu\nu}$ inner product (1.65b) preserved under unitary transformations, and the requirement of Lorentz invariance, which needs the $g^{\mu\nu}$ inner product (2.91) preserved under Lorentz transformations.

7.1.3 Quantum-Field Theory

Quantum-field theory (QFT), the merging of QM with SR, started with an early contribution of the Swedish physicist Oskar Benjamin Klein and the German physicist Walter Gordon (1893–1939) and a series of papers of Paul Dirac,[17] followed by many other contributions, like that of R. Feynman. It is an important chapter of physics, since it is not only a construction out of QM and SR but it is also the basis of the Standard model of particles. Although the second quantisation of the electromagnetic field was a crucial step towards QFT, its treatment is essentially non-relativistic. It may be further noted that A. Einstein was not involved in the building of QFT likely due to his scepticism towards QM (Sects. 2.3 and 3.3).

Covariant and Contravariant Derivatives

The field theory is said to be local: there is no "action at a distance", in agreement with Einstein's *Nahewirkung*.. This will be an important feature of the Standard model of

[17] Dirac (1928a, b, 1930a, b).

particles. Let us consider a scalar field.[18] For a scalar field $\varphi(x) = \varphi(x^0, \mathbf{x})$ that is function of the space-time quadrivector (2.81) infinitesimally moving from position x to $x + dx$, we have

$$\frac{d\varphi}{dx^\mu} = \frac{\partial\varphi}{\partial x^\mu}, \tag{7.56}$$

which is invariant under Lorentz transformation (2.84), and I recall the reader that here and in the following I use Einstein's convention (2.83). This Lorentz invariance is true for scalar fields but not necessarily for other quantities. Note that, the spatial coordinates $x^j(t)$ for particles are not the same as the x^j for fields: while the former are time-dependent, the latter are not (are independent variables) and denote the fixed space location upon which the values of the field (and other quantities like energy density) depend. In other words, particle's position coordinates are its generalised spatial coordinates and its momentum components the conjugate momenta, while for fields, the field is itself a generalised coordinate and each field has its own conjugate momentum (density), and this is different from the physical momentum (density) that the field possesses. The set $\partial\varphi/\partial x^\mu$ must make up a covariant vector so that we can write

$$\partial_\mu\varphi := \left(\frac{\partial\varphi}{\partial t}, \boldsymbol{\nabla}\varphi\right), \tag{7.57}$$

where the four derivatives are

$$\partial_\mu := \frac{\partial}{\partial x^\mu} = (\partial_t, \partial_x, \partial_y, \partial_z), \tag{7.58}$$

and according to Eq. (2.90), for the relative contravariant expression, we have

$$g^{\mu\nu}\partial_\nu\varphi = \partial^\mu\varphi := \frac{\partial\varphi}{\partial x_\mu} = \left(\frac{\partial\varphi}{\partial t}, -\boldsymbol{\nabla}\varphi\right), \tag{7.59}$$

where making use of the rule (2.91) for scalar product and recalling Einstein's convention (2.83), we have that the products

$$\partial_\mu\varphi\partial^\mu\varphi = \left(\frac{\partial\varphi}{\partial t}\right)^2 - (\boldsymbol{\nabla}\varphi)^2, \text{ and } \partial_\mu\partial^\mu\varphi = \frac{\partial^2\varphi}{\partial t^2} - \boldsymbol{\nabla}^2\varphi \tag{7.60}$$

are invariant under Lorentz transformations. To use covariant derivatives is, especially, relevant when there are more fields and we like to preserve the gauge condition.

[18] Klauber (2013, Chaps. 2–3), Cottingham and Greenwood (2007, Sect. 2.3, Chap. 3), Schwartz (2014, Chap. 3). I recommend the first textbook for students who know little about the subject and the latter textbook for a complete and updated treatment of quantum-field theory.

Lagrangian and Hamiltonian Densities

The classical Langrangian L (1.25) or (1.51) and Hamiltonian H (1.43) can be expressed in terms of the Lagrangian density $\mathcal{L}$ and energy density $\mathcal{H}$ (the energy per unit volume), respectively, as follows:

$$L = T - U = \int \mathcal{L} dx^3, \quad H = T + V = \int \mathcal{H} dx^3, \tag{7.61}$$

with densities

$$\mathcal{L} = \mathcal{T} - \mathcal{V} \text{ and } \mathcal{H} = \mathcal{T} + \mathcal{V}, \tag{7.62}$$

where it should be noted that kinetic terms are bilinear while interaction terms have three or more fields. It is customary to use Lagrangian densities for the sake of simplicity, and I shall adopt this convention here and in the following.

In a quantum-field context, the quadrivectors (2.81) and (2.85) read

$$x^\mu = (t, \hat{\mathbf{x}}) \text{ and } p^\mu = (E, \hat{\mathbf{p}}), \tag{7.63}$$

respectively, which both have a scalar component. We pass from classical to quantised fields according to the following map from Poisson brackets (1.214) to commutators (1.210):

$$\left\{\varphi^j(\mathbf{x}, t), \pi_k(\mathbf{y}, t)\right\} = \delta^j_k \delta(\mathbf{x} - \mathbf{y}) \longmapsto \left[\varphi^j(\mathbf{x}, t), \pi_k(\mathbf{y}, t)\right] = i\hbar \delta^j_k \delta(\mathbf{x} - \mathbf{y}), \tag{7.64}$$

where π_k is the conjugate momentum density of the field and $j, k = 1, 2, \ldots, n$.[19]

In analogy with Eqs. (1.35) and (1.42), we have

$$\mathcal{H} = \frac{\partial \mathcal{L}}{\partial \dot{\varphi}^k} \dot{\varphi}^k - \mathcal{L}, \text{ with } \frac{\partial \mathcal{L}}{\partial \dot{\varphi}^k} = \pi_k, \tag{7.65}$$

where here, as anticipated, the generalised position q is replaced by the (quantised) field φ. Thus, the action in 3D spatial dimensions takes the form

$$S = \int L dt = \int \mathcal{L}(x) dx^4. \tag{7.66}$$

Since for a scalar field $\varphi(x) = \varphi(x^0, \mathbf{x})$, the Lagrangian density takes the Lorentz-invariant form

$$\mathcal{L} := \mathcal{L}(\varphi, \partial_\mu \varphi), \tag{7.67}$$

[19] Although quantised, and therefore represented by operators, I follow here the standard convention for fields and their conjugate momentum densities.

it makes the action (7.66) also Lorentz invariant: at any point x in space-time, this Lagrangian density depends only on the field and its first derivatives at that point. The field equation is easily derived from the condition $\delta S = 0$ (see Eq. (1.26)), together with the condition that the field vanishes at large distances, so that, generalising to space-time Eq. (1.33), we find the Euler–Lagrange equation for the relativistic Lagrangian density:

$$\frac{\partial \mathcal{L}}{\partial \varphi} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \varphi)} \right) = 0. \tag{7.68}$$

Note that, the partial derivative of the term between parentheses is on the 'variable' $\partial_\mu \varphi$. Although the Lagrangian possesses the important property of being Lorentz invariant, differently from the Hamiltonian (or the number operator), is not observable.

Derivation of the Energy–Momentum Tensor

Space-time translation invariance tells us that physics at a given point is the same at another point. Let us consider a uniform infinitesimal space-time displacement

$$x^\mu \to x^\mu + \delta x^\nu, \tag{7.69}$$

where δx^ν does not depend on x^μ. Making use of the rules (1.27), the corresponding change in the field is

$$\frac{\delta \varphi}{\delta x^\nu} = \partial_\nu \varphi. \tag{7.70}$$

This also applies to the Lagrangian density itself, which is a scalar:

$$\frac{\delta \mathcal{L}}{\delta x^\nu} = \partial_\nu \mathcal{L}. \tag{7.71}$$

From Eq. (7.68) for $\partial \mathcal{L}/\partial \varphi$ and on the outline of Eq. (1.38), we split the change of the Lagrangian into two parts:

$$\delta \mathcal{L} = \frac{\partial \mathcal{L}}{\partial \varphi} \delta \varphi + \frac{\partial \mathcal{L}}{\partial (\partial_\mu \varphi)} \delta (\partial_\mu \varphi). \tag{7.72}$$

Using the fact that

$$\delta (\partial_\mu \varphi) = \partial_\mu (\delta \varphi), \tag{7.73}$$

and Leibniz product rule (1.206), we rewrite Eq. (7.72) as

$$\delta \mathcal{L} = \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \varphi)} \delta \varphi \right), \tag{7.74}$$

which using Eqs. (7.70)–(7.71), takes the final form

$$\partial_\nu \mathcal{L} = \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)} \partial_\nu \varphi \right), \tag{7.75}$$

from which, thanks to the transformation $\partial_\nu \mathcal{L} = \partial_\mu (g_{\mu\nu} \mathcal{L})$, it follows that

$$\partial_\mu \hat{\mathbf{T}}_{\mu\nu} = 0, \tag{7.76}$$

where

$$\hat{\mathbf{T}}_{\mu\nu} := \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)} \partial_\nu \varphi - g_{\mu\nu} \mathcal{L} \tag{7.77}$$

is the energy–momentum tensor or *stress–energy tensor*, describing the density and flux of energy and momentum in space-time: it is an attribute of matter, radiation and non-gravitational force fields. Note that, the component $\hat{\mathbf{T}}_{00}$ corresponds to the energy density of the field, which turns out to be the Hamiltonian density (7.65).

Derivation of the Klein–Gordon Equation

Let us derive the *Klein–Gordon* equation, after the names of the already mentioned Oskar Klein and the German physicist Walter Gordon: it is a relativistic version of the Schrödinger equation (1.10) or (1.158), where the field φ replaces the state vector or the wave function and we use relativistic quadrivectors. It describes the equations of motion for a free scalar field. The two physicists took as the dynamical quantity ruling time translations not the Hamiltonian (as for the ordinary Schrödinger equation) but $\hat{H}^2$, so that we have

$$\left(i\hbar \frac{\partial}{\partial t} \right) \left(i\hbar \frac{\partial}{\partial t} \right) \varphi = \hat{H}^2 \varphi. \tag{7.78}$$

Starting from the classical expression (2.78) for energy, we can write

$$\hat{H} = \sqrt{m_0^2 c^4 + \hbar^2 c^2 \partial_j \partial_j}, \tag{7.79}$$

so that we reformulate the previous equation as

$$-\frac{\partial}{\partial x^0} \frac{\partial}{\partial x_0} \varphi = \left(\frac{\partial}{\partial x^j} \frac{\partial}{\partial x_j} + \frac{m_0^2 c^2}{\hbar^2} \right) \varphi. \tag{7.80}$$

This allows us to finally derive the Klein–Gordon equation

$$\left(\partial_\mu \partial^\mu + \mu^2 \right) \varphi = 0, \text{ with } \mu^2 = \frac{m_0^2 c^2}{\hbar^2}. \tag{7.81}$$

Thus, the essential difference between this equation and Schrödinger equation (1.10) is that in the latter the time derivative is first order while here is second order.

For the sake of notation, in the following examination, I shall put $\hbar = c = 1$, what gives all quantities dimensions of mass to some power and allows us to use m instead of μ.

Spin 0 Particles

Let us now consider some classes of particles in a very schematic form. For spin 0, we have just a degree of freedom (corresponding to total angular momentum $j = 0$: see Sect. 1.3.1) and the Lagrangian density (7.62) with its kinetic and potential components, takes the form[20]

$$\mathcal{L} = \frac{1}{2}\left(\partial_\mu\varphi\partial_\mu\varphi - m^2\varphi^2\right). \tag{7.82}$$

The Lagrangian density (7.82) allows us to compute the energy density (7.65) as

$$\frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\dot{\varphi} - \mathcal{L} = \frac{1}{2}\left[(\partial_t\varphi)^2 + (\boldsymbol{\nabla}\varphi)^2 + m^2\varphi^2\right], \tag{7.83}$$

and using Eqs. (7.60), to write the Klein–Gordon equation (7.81) for spin 0 particles as

$$\left(-\frac{\partial^2}{\partial t^2} + \boldsymbol{\Delta} - m^2\right)\varphi = 0, \tag{7.84a}$$

where I have used the Laplacian (1.159) and dropped the index for rest mass. In more compact form, using the Dalembertian (2.80), which can be rewritten as

$$\Box := \pm\frac{\partial}{\partial x_\mu}\frac{\partial}{\partial x^\mu}, \tag{7.84b}$$

it takes the form

$$\left(\Box + m^2\right)\varphi = 0. \tag{7.84c}$$

Spin 1 Particles

For spin 1, due to the presence of the magnetic field (7.44b), we no longer deal with scalar fields only and there are three degrees of freedom if $m > 0$ (massive spin-1 particles), which need to be embedded in a vector field (potential)

$$A^\mu := (A^0, A^1, A^2, A^3) = (\varphi, \hat{\mathbf{A}}), \tag{7.85}$$

[20] Schwartz (2014, Chap. 8).

being a contravariant quadrivector field (whose four components one is the usual scalar field) and satisfies the so-called Lorentz gauge condition $\partial_\mu A^\mu = 0$, which is in fact a partial gauge fixing of the electromagnetic potential. In fact,

$$A_\mu(x) \to A_\mu(x) + \partial_\mu \alpha(x) \tag{7.86}$$

for any function $\alpha(x)$. In other words, two fields A_μ that differ by the derivative of a scalar are physically equivalent. The vector field is the direct sum of 3D (spin 1) and 1D (spin 0) representations of the rotation group SO, i.e. we have SO(1,3).

Allowing arbitrary coefficients a, b for the different terms, the most general free Lagrangian for massive spin 1 particles is

$$\mathcal{L} = \frac{1}{2}\left(a A_\mu \,\Box\, A_\mu + b A_\mu \partial_\mu \partial_\nu A_\nu + m^2 A_\mu^2\right), \tag{7.87}$$

where, relative to the previous case, we have one more Lorentz-invariant two-derivative kinetic term $A_\mu \partial_\mu \partial_\nu A_\nu$. Dividing by A_μ, the equations of motions (7.81) become

$$a \,\Box\, A_\mu + b \partial_\mu \partial_\nu A_\nu + m^2 A_\mu = 0, \tag{7.88a}$$

or, taking the derivative (7.58) of the latter equation,

$$\left[(a+b)\Box + m^2\right]\left(\partial_\mu A_\mu\right) = 0. \tag{7.88b}$$

Let us introduce the 2-form *Maxwell or electromagnetic tensor* (in fact, $\hat{F}_{\mu\nu} = -\hat{F}_{\nu\mu}$: see Sect. 1.2.3), given Eqs. (7.24b) and (7.31), which is the zero-trace matrix (the index μ for the rows and the index ν for the columns)

$$\begin{aligned} \hat{F}_{\mu\nu} &:= \partial_\mu A_\nu - \partial_\nu A_\mu \\ &= \begin{bmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -B_z & B_y \\ -E_y & B_z & 0 & -B_x \\ -E_z & -B_y & B_x & 0 \end{bmatrix}, \end{aligned} \tag{7.89}$$

with matrix elements

$$F_{\mu=\nu} = 0,\ F_{0j} = -F_{j0} = \hat{E}_j,\ F_{ij} = -F_{ji} = -\epsilon_{ijk}\hat{B}_k, \tag{7.90}$$

for $\mu, \nu = 0, 1, 2, 3$; $i, j, k = x, y, z$, and where I have used Eq. (7.5) for the components of the magnetic field. The contravariant form is

$$\hat{F}^{\mu\nu} = \begin{bmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{bmatrix}, \tag{7.91}$$

Assuming $a = 1, b = -1$, Eq. (7.87) reads

$$\begin{aligned}\mathcal{L} &= -\frac{1}{4}\hat{F}_{\mu\nu}\hat{F}^{\mu\nu} + \frac{1}{2}m^2 A_\mu^2 \\ &= -\frac{1}{4}\hat{F}_{\mu\nu}^2 + \frac{1}{2}m^2 A_\mu^2.\end{aligned} \tag{7.92}$$

Equation (7.92) is called the *Proca Lagrangian*, after the name of the Romanian physicist Alexandru Proca (1897–1955). In other words, the classical electric and magnetic 3 vectors give rise to a single (both classical and quantum mechanical) electromagnetic field 4 tensor.[21] This means that $\hat{\mathbf{B}}$ is a contravariant pseudo 3 vector. Note also that we have

$$\frac{1}{2}\epsilon_{\alpha\beta\gamma\delta}\hat{F}^{\alpha\beta}\hat{F}^{\gamma\delta} = -4\hat{\mathbf{B}}\cdot\hat{\mathbf{E}}. \tag{7.93}$$

The energy–momentum tensor (7.77) for Lagrangian (7.92) is given by

$$\begin{aligned}\hat{\mathbf{T}}_{\mu\nu} &= \frac{\partial\mathcal{L}}{\partial(\partial_\mu A_\alpha)}\partial_\nu A_\alpha - g_{\mu\nu}\mathcal{L} \\ &= -\hat{F}_{\mu\alpha}\partial_\nu A_\alpha + g_{\mu\nu}\left(\frac{1}{4}\hat{F}_{\alpha\beta}^2 - \frac{1}{2}m^2 A_\alpha^2\right),\end{aligned} \tag{7.94}$$

where the indexes α, β range over the four space-time coordinates. Using Eqs. (7.44) and (7.93), the energy density is given by

$$\hat{\mathbf{T}}_{00} = -\left(\partial_t A_\alpha - \partial_\alpha\varphi\right)\partial_t A_\alpha - \frac{1}{2}\left(\hat{\mathbf{B}}^2\hat{\mathbf{E}}^2 + m^2 A_\alpha A_\alpha\right), \tag{7.95}$$

since

$$\hat{F}_{\mu\alpha}\partial_\nu A_\alpha = \left(\partial_\mu A_\alpha - \partial_\alpha A_\mu\right)\partial_\nu A_\alpha, \tag{7.96}$$

and thus $\hat{F}_{0\alpha}\partial_t A_\alpha$ is equal to the first term of Eq. (7.95).

Note that, for massless spin 1 particles, the Lagrangian (7.92) reduces to

$$\mathcal{L} = -\frac{1}{4}\hat{F}_{\mu\nu}^2, \tag{7.97}$$

which is gauge invariant. Given the gauge condition (7.86), the equations of motion become

$$\Box A_\mu - \partial_\mu(\partial_\nu A_\nu) = 0. \tag{7.98}$$

[21] Rindler (2001, p. 130).

Note that, matter coupled to massless spin 1 particles automatically involves the existence of antiparticles.[22]

Spinors

Many particles have a non-zero spin. In particular, let us consider electrons, and more specifically, a superposition $|\psi\rangle$ of spin states (1.322).[23] In the non-relativistic limit, the time evolution of this ket is ruled by the so-called *Schrödinger–Pauli* equation, after the names of E. Schrödinger and W. Pauli,[24]

$$\begin{aligned} i\partial_t |\psi\rangle &= \hat{H} |\psi\rangle = \left\{ \left[\frac{1}{2m_\mathrm{e}} \left(i\nabla - \mathrm{e}\hat{\mathbf{A}} \right)^2 - \mathrm{e}\varphi \right] \hat{I} + \mu_B \hat{\mathbf{B}} \cdot \hat{\boldsymbol{\sigma}} \right\} |\psi\rangle \\ &= \left\{ \left[\frac{1}{2m_\mathrm{e}} \left(i\nabla - \mathrm{e}\hat{\mathbf{A}} \right)^2 - \mathrm{e}\varphi \right] \hat{I} + \mu_B \begin{bmatrix} \hat{B}_z & \hat{B}_x - i\hat{B}_y \\ \hat{B}_x + i\hat{B}_y & -\hat{B}_z \end{bmatrix} \right\} |\psi\rangle \,, \end{aligned} \tag{7.99}$$

where e is the charge of the electron, m_e its mass, and

$$\mu_B = \frac{\mathrm{e}}{2m_\mathrm{e}} \tag{7.100}$$

is the *Bohr magneton*, after the name of N. Bohr, denoting the strength of the electron's magnetic dipole moment in Eq. (7.99). I have expanded the Hamiltonian, the term between square brackets on the first row having the form of the usual Schrödinger equation, while $\hat{\mathbf{B}} \cdot \hat{\boldsymbol{\sigma}}$ is called the Stern–Gerlach term describing the spin interaction with the magnetic field. Moreover, note that the kinetic term is connected by a (minimal) coupling factor $\mathrm{e}\hat{\mathbf{A}}$ and $\mathrm{e}\varphi$ is the electric potential term. Finally, in the second row, I have made use of the expansion (1.326) for $\hat{\mathbf{B}}$. Of course, we are interested in deriving a relativistic counterpart of this equation.

In order to satisfy the Lorentz transformation (2.102), let us now define the linear combinations

$$\hat{J}_n^+ := \frac{1}{2} \left(\hat{J}_n + i\hat{K}_n \right) \text{ and } \hat{J}_n^- := \frac{1}{2} \left(\hat{J}_n - i\hat{K}_n \right), \tag{7.101}$$

which satisfy the commutation relations

$$\left[\hat{J}_l^+, \hat{J}_m^+ \right] = i\epsilon_{lmn} \hat{J}_n^+, \left[\hat{J}_l^-, \hat{J}_m^- \right] = i\epsilon_{lmn} \hat{J}_n^-, \left[\hat{J}_l^+, \hat{J}_m^- \right] = 0. \tag{7.102}$$

These commutation relations indicate that the Lie algebra for the Lorentz group has two commuting (Abelian) subalgebras. The algebra generated by $\hat{J}_n^+$ (or $\hat{J}_n^-$) is the

[22] Schwartz (2014, p. 141).

[23] On what follows see Schwartz (2014, Chap. 10).

[24] Pauli (1927).

3D rotation algebra SO(3) = SU(2). Thus, it has been shown that

$$\mathrm{SO}(1,3) = \mathrm{SU}(2) \oplus \mathrm{SU}(2). \tag{7.103}$$

Now, it turns out that Pauli spin matrices satisfy the commutation relations (1.325c), which, thanks to the Levi-Civita tensor, can be reformulated as (1.340) or

$$\left[\hat{\sigma}_i, \hat{\sigma}_j\right] = 2i\epsilon_{ijk}\hat{\sigma}_k. \tag{7.104}$$

Rescaling, we obtain

$$\left[\frac{\hat{\sigma}_i}{2}, \frac{\hat{\sigma}_j}{2}\right] = i\epsilon_{ijk}\frac{\hat{\sigma}_k}{2}, \tag{7.105}$$

which is in fact SO(3) algebra (see Fig. 4.4, Sect. 4.2.1). There exist two complex $\hat{J} = 1/2$ representations:

$$\begin{pmatrix} \frac{1}{2} \\ 0 \end{pmatrix} \text{ and } \begin{pmatrix} 0 \\ \frac{1}{2} \end{pmatrix}. \tag{7.106}$$

Thus, we can set

$$\hat{J}_n^- = \frac{1}{2}\hat{\sigma}_n, \tag{7.107}$$

which generates the "1/2" in the first (7.106). Then, the easiest thing is to put

$$\hat{J}_n^+ = 0 \tag{7.108}$$

for the "0" element. For the second (7.106), we put

$$\hat{J}_n^- = 0 \text{ and } \hat{J}_n^+ = \frac{1}{2}\hat{\sigma}_n. \tag{7.109}$$

Now, in agreement with Eq. (2.102), rotations and boosts generators (2.101) are described as

$$\hat{J} = \hat{J}^- + \hat{J}^+, \quad \hat{K} = i\left(\hat{J}^- - \hat{J}^+\right), \tag{7.110}$$

respectively, so that we finally get the group generators in the $(1/2, 0)$ and $(0, 1/2)$ representations. Since the Pauli matrices are Hermitian, the rotations are Hermitian and the boosts are anti-Hermitian ($\hat{K}^\dagger = -\hat{K}$). Using the latter equations, we have the so-called *left* and *right spinors* (see also Eq. (1.329)), which are adjoints of each other:

$$|\psi\rangle_L := \begin{pmatrix} \frac{1}{2} \\ 0 \end{pmatrix} : \ \hat{J} = \frac{1}{2}\hat{\boldsymbol{\sigma}}, \quad \hat{K} = \frac{i}{2}\hat{\boldsymbol{\sigma}}, \tag{7.111a}$$

$$|\psi\rangle_R := \begin{pmatrix} 0 \\ \frac{1}{2} \end{pmatrix} : \ \hat{J} = \frac{1}{2}\hat{\boldsymbol{\sigma}}, \quad \hat{K} = -\frac{i}{2}\hat{\boldsymbol{\sigma}}, \tag{7.111b}$$

and are known as Weyl spinors, after the name of H. Weyl. As mentioned in Sect. 1.3.1, a spinor is a very particular object that after a 2π rotation, instead of going back to its initial configuration, transforms to its negative and thus needs an additional 2π rotation for that.

Dirac Matrices

I shall introduce now a relativistic equation for the *free electron* in order to reformulate the Schrödinger–Pauli equation for space-time. This is the *Dirac equation*. In order to derive the latter, let us come back to the basic Schrödinger's equation (1.10).[25] In order that this equation secures a symmetry between space and time, Dirac postulated that the Hamiltonian (the so-called *Dirac Hamiltonian*) for a free electron takes the form

$$\hat{H}_D = \frac{\hbar c}{i}\hat{\boldsymbol{\alpha}} \cdot \hat{\mathbf{p}} + \hat{\beta} m c^2, \tag{7.112}$$

where m and $\hat{\mathbf{p}}$ are mass and momentum of the electron, while $\hat{\boldsymbol{\alpha}} = (\hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3)$ and $\hat{\beta}$ constitute the *chiral representation*[26]

$$\hat{\alpha}^j = \begin{bmatrix} -\hat{\sigma}^j & \hat{\mathbf{0}} \\ \hat{\mathbf{0}} & \hat{\sigma}^j \end{bmatrix} \text{ and } \hat{\beta} = \begin{bmatrix} \hat{\mathbf{0}} & \hat{\sigma}^0 \\ \hat{\sigma}^0 & \hat{\mathbf{0}} \end{bmatrix}, \tag{7.113}$$

with $\hat{\sigma}^0$ as the 2D identity matrix, $\hat{\sigma}^j$, $j = 1, 2, 3$ as the Pauli matrices (1.325), and

$$\hat{\mathbf{0}} := \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}. \tag{7.114}$$

In this context, I have reintroduced the constants $\hbar$, c for comparative reasons. Then, the *Dirac equation* is

$$\left[\frac{\hbar c}{i}\left(\hat{\alpha}_1 \frac{\partial}{\partial x^1} + \hat{\alpha}_2 \frac{\partial}{\partial x^2} + \hat{\alpha}_3 \frac{\partial}{\partial x^3}\right) + \hat{\beta} m c^2\right] |\psi\rangle = i\hbar \frac{\partial}{\partial t} |\psi\rangle \ . \tag{7.115}$$

The reason for Dirac's choice is that the square modulus of the wave function of a relativistic particle should provide the probability density to find the particle in a certain small space-time region in analogy with Born's probability rule (Sects. 1.1.3

[25] Dirac (1928a, b). See also Cottingham and Greenwood (2007, Chap. 5).

[26] Schwartz (2014, Chap. 10) Cottingham and Greenwood (2007, Chaps. 6–8).

and 1.2.3). If this equation describes the motion of a free particle, it should be in agreement with the Klein–Gordon equation (7.84c). This can be in fact proved. However, the advantage of the Dirac equation is that it uses first-order time derivatives (as the Schrödinger one).

Since the $\hat{\alpha}_j$ and $\hat{\beta}$ are 4×4 matrices, the Dirac wave function or state vector is a four-component column field that can be written by using the left-and right-handed components (7.111):

$$|\psi\rangle = \begin{pmatrix} |\psi\rangle_L \\ |\psi\rangle_R \end{pmatrix} = \begin{pmatrix} |\psi\rangle_L \\ \mathbf{0} \end{pmatrix} + \begin{pmatrix} \mathbf{0} \\ |\psi\rangle_R \end{pmatrix}, \quad \left\langle \overline{\psi} \right| = \begin{pmatrix} \psi_R^\dagger & \psi_L^\dagger \end{pmatrix}, \tag{7.116}$$

in agreement with Eq. (1.329), where

$$\mathbf{0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \tag{7.117}$$

and

$$\left\langle \overline{\psi} \right| := |\psi\rangle^\dagger \hat{\gamma}^0 = \langle \psi | \hat{\gamma}^0. \tag{7.118}$$

The spinors

$$\hat{\gamma}^0 = \begin{bmatrix} \hat{\mathbf{0}} & \hat{\sigma}^0 \\ \hat{\sigma}^0 & \hat{\mathbf{0}} \end{bmatrix} \text{ and } \hat{\gamma}^j = \begin{bmatrix} \hat{\mathbf{0}} & \hat{\sigma}^j \\ -\hat{\sigma}^j & \hat{\mathbf{0}} \end{bmatrix}, \tag{7.119}$$

known as *Dirac matrices*, are obtained from matrices (7.113): $\hat{\gamma}_0 = \hat{\beta}$ and $\hat{\gamma}_j = \hat{\beta}\hat{\alpha}^j$. From the properties of $\hat{\beta}$, $\hat{\alpha}^j$ matrices it follows that

$$(\hat{\gamma}^0)^2 = \hat{I}, \ (\hat{\gamma}^j)^2 = -\hat{I}, \ \left[\hat{\gamma}^\mu, \hat{\gamma}^\nu\right]_+ = 2g^{\mu\nu}, \tag{7.120}$$

for $j = 1, 2, 3$ and $\mu \neq \nu$. This is an instance of Clifford algebra (Sect. 4.4), although in that case, we would have a signature $(++++)$, and is called *Dirac algebra*.[27] Since a spinor can be thought as an object upon which the elements of the Dirac algebra act as operators, then $|\psi\rangle$ is itself a spinor, and its additional (internal, i.e. not dependent on position and time) degrees of freedom describe precisely the electron spin, again in agreement with Eq. (1.329), but with four components instead of two.[28]

Going back to natural units, the Lagrangian density for the Dirac field becomes in such a context[29]:

[27] Penrose (2004, p. 619).

[28] Penrose (2004, Sect. 11.5 and p. 621).

[29] Schwartz (2014, Sect. 10.3).

$$\begin{aligned}\mathcal{L} &= \left\langle \overline{\psi} \right| (i\hat{\gamma}^\mu \partial_\mu - m) \left| \psi \right\rangle \\ &= \langle \psi | \, \hat{\gamma}^0 (i\hat{\gamma}^\mu \partial_\mu - m) \left| \psi \right\rangle \,, \end{aligned} \tag{7.121}$$

and the Dirac equation (7.115) takes the symmetrical form

$$(i\hat{\gamma}^\mu \partial_\mu - m) \left| \psi \right\rangle = (i \not{\partial} - m) \left| \psi \right\rangle = 0, \tag{7.122}$$

where I have used the simplified (Feynman's) notation

$$\not{\partial} := \hat{\gamma}_0 \frac{\partial}{\partial t} - \hat{\boldsymbol{\gamma}} \cdot \boldsymbol{\nabla} = \hat{\gamma}^\mu \partial_\mu, \tag{7.123}$$

where $\hat{\boldsymbol{\gamma}} = (\hat{\gamma}^1, \hat{\gamma}^2, \hat{\gamma}^3)$, from which

$$\not{\partial}^2 = \Box \tag{7.124}$$

follows. This allows to split the Klein–Gordon equation (7.84c) into two components:

$$(\not{\partial} - im)(\not{\partial} + im) \left| \psi \right\rangle = -\left(\not{\partial}^2 + m^2\right) \left| \psi \right\rangle = 0. \tag{7.125}$$

Note that, the equation of motion for $\overline{\psi}$ takes the form

$$-i\partial_\mu \left\langle \overline{\psi} \right| \hat{\gamma}^\mu - m \left\langle \overline{\psi} \right| = 0. \tag{7.126}$$

Coupling to the Photon

Under a gauge transformation $\left| \psi \right\rangle$ transforms just like a scalar, and for a spinor with electric charge -1 (like the electron), we have[30]

$$\left| \psi \right\rangle \longmapsto e^{-i\alpha} \left| \psi \right\rangle \,. \tag{7.127}$$

This means that we can use the covariant derivative $\partial_\mu + ieA_\mu$ as for a scalar (see Eqs. (7.84)), with an additional term represented by A_μ. Then, the Dirac equation (7.115) in natural units becomes

$$(i \not{\partial} - e \not{A} - m) \left| \psi \right\rangle = 0, \tag{7.128}$$

where

$$\not{A} := \hat{\gamma}^\mu A_\mu = \hat{\gamma}_\mu A^\mu. \tag{7.129}$$

[30] Schwartz (2014, Sect. 10.4).

Let us compare the Dirac equation to the Klein–Gordon equation (7.125) for a scalar field φ coupled to A_μ, which, under the previous assumption, takes the form

$$\left[(i\partial_\mu - \mathrm{e}A_\mu)^2 - m^2\right]\varphi = 0, \tag{7.130}$$

and in order to get the correct relativistic form of the Schrödinger–Pauli equation, multiply Eq. (7.128) by $(i\not{\partial} - \mathrm{e}\not{A} - m)$, as to obtain

$$\begin{aligned} 0 &= (i\not{\partial} - \mathrm{e}\not{A} - m)(i\not{\partial} - \mathrm{e}\not{A} - m)\,|\,\psi\rangle \\ &= \left[(i\partial_\mu - \mathrm{e}A_\mu)(i\partial_\nu - \mathrm{e}A_\nu)\hat{\gamma}^\mu\hat{\gamma}^\nu - m^2\right]|\,\psi\rangle \\ &= \frac{1}{4}\left\{\left[i\partial_\mu - \mathrm{e}A_\mu, i\partial_\nu - \mathrm{e}A_\nu\right]_+\left[\hat{\gamma}^\mu, \hat{\gamma}^\nu\right]_+ + \left[i\partial_\mu - \mathrm{e}A_\mu, i\partial_\nu - \mathrm{e}A_\nu\right]\left[\hat{\gamma}^\mu, \hat{\gamma}^\nu\right] - m^2\right\}|\,\psi\rangle\,, \end{aligned} \tag{7.131}$$

where it may be noted that the first two terms on the last row are the products between two anticommutators and between two commutators, respectively. Since the first commutator gives

$$\left[i\partial_\mu - \mathrm{e}A_\mu, i\partial_\nu - \mathrm{e}A_\nu\right] = -\mathrm{e}\left(i\partial_\mu A_\nu - i\partial_\nu A_\mu\right) = -\mathrm{e}i\hat{F}_{\mu\nu}, \tag{7.132}$$

where $\hat{F}_{\mu\nu}$ is the Maxwell tensor (7.89), we can finally derive

$$\left\{(i\partial_\mu - \mathrm{e}A_\mu)^2 - \frac{\mathrm{e}i}{4}\hat{F}_{\mu\nu}\left[\hat{\gamma}^\mu, \hat{\gamma}^\nu\right] - m^2\right\}|\,\psi\rangle = 0, \tag{7.133}$$

where I have used the properties (7.120). Let us now put

$$\hat{S}^{\mu\nu} = \frac{1}{2}\hat{\sigma}^{\mu\nu} = \frac{i}{4}\left[\hat{\gamma}^\mu, \hat{\gamma}^\nu\right], \tag{7.134}$$

so that, taking into account matrices (7.119), we have (in the so-called Weyl representation, after the name of H. Weyl)

$$\hat{S}_{ij} = \frac{1}{2}\epsilon_{ijk}\begin{bmatrix}\hat{\sigma}^k & \hat{\mathbf{0}} \\ \hat{\mathbf{0}} & \hat{\sigma}^k\end{bmatrix} \text{ and } \hat{S}_{0j} = -\frac{i}{2}\begin{bmatrix}\hat{\sigma}^j & \hat{\mathbf{0}} \\ \hat{\mathbf{0}} & -\hat{\sigma}^j\end{bmatrix}, \tag{7.135}$$

which allows us to rewrite Eq. (7.133) as

$$\left[\left(i\partial_\mu - \mathrm{e}A_\mu\right)^2 - \mathrm{e}\hat{F}_{\mu\nu}\hat{S}^{\mu\nu} - m^2\right]|\,\psi\rangle = 0. \tag{7.136}$$

Note that, the matrices $\hat{S}^{\mu\nu}$ have zero trace like the Pauli spin matrices (1.325) and the electromagnetic tensor (7.89). Since the elements of the Maxwell tensor are (7.90) and map to the two previous matrices, from Eq. (7.136), we finally derive the relativistic form of Eq. (7.99):

$$\left\{(i\partial_\mu - \mathrm{e}A_\mu)^2 - m^2 - \mathrm{e}\begin{bmatrix}(\hat{\mathbf{B}} + i\hat{\mathbf{E}})\cdot\hat{\boldsymbol{\sigma}} & \hat{\mathbf{0}} \\ \hat{\mathbf{0}} & (\hat{\mathbf{B}} - i\hat{\mathbf{E}})\cdot\hat{\boldsymbol{\sigma}}\end{bmatrix}\right\}|\psi\rangle = 0, \tag{7.137}$$

which corresponds to a magnetic (and not electric) dipole moment μ_B, whose size, with appropriate normalisation, is given by the (7.100).[31] In other words, while free spinors satisfy an equation of motion of the kind (7.130), when spinors are coupled to a photon we get an additional interaction term that corresponds to the magnetic dipole. The size of the electron's magnetic moment can be read off as the coefficients of this additional term.

Poincaré Group

Since, in a relativistic context, we like to recover Lorentz invariance, our problem is how to relate the unitary Pauli matrices with the Lorentz group, defined by transformations (2.95) and (2.99). The Lorentz group is obviously a set of transformations that preserves the Minkowski (flat) metric (2.88). We can identify particles with unitary representations of the Poincaré group or inhomogeneous Lorentz group, characterised by both mass m and spin s (such a group is the so-called semidirect product of the space-time translations and the Lorentz group). Thus, the transformation (2.82) is generalised to

$$x^\mu \mapsto x'^\mu = \Lambda^\mu_{\ \nu} x^\nu + a^\mu. \tag{7.138}$$

The translational motion is described by the energy–momentum 4 vector: the components $\hat{p}_0$, $\hat{p}_1$, $\hat{p}_2$, $\hat{p}_3$ of the energy–momentum 4 vector (2.85) give time translation (E) and the three spatial displacements ($\hat{\mathbf{p}}$), respectively. Thus, we have 10 elements of the group as a whole. Apart from the commutation relations (2.104), we have

$$\left[\hat{J}_l, \hat{p}_m\right] = i\epsilon_{lmn}\hat{p}_n, \left[\hat{J}_l, \hat{p}_0\right] = 0, \left[\hat{K}_l, \hat{p}_m\right] = i g_{lm}\hat{p}_0, \left[\hat{K}_l, \hat{p}_0\right] = -i\hat{p}_l. \tag{7.139}$$

A final comment is opportune here. A problem affecting quantum-field theory are the effects of self-interaction. For instance, an electron is surrounded by a cloud of virtual particles, especially photons, that will alter its mass and charge. An ensemble of techniques called *renormalisation* replaces the originally postulated ones with those observed. Renormalisation is especially helpful for dealing with infinities that arise in this way.

[31] For the electric dipole moment, I suggest the reader to study Schwartz (2014, Sect. 29.5.3).

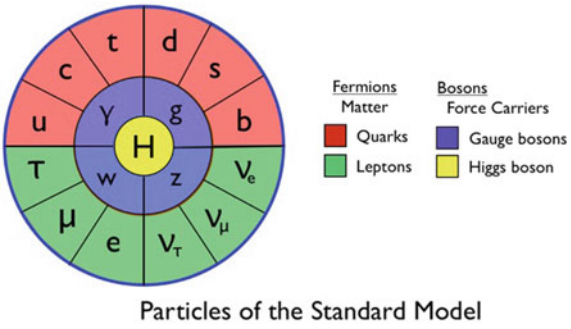


Fig. 7.3 Standard model. Recently, also the Higgs boson has been added to the model. Adapted from http://theoryandpractice.org/2013/08/a-fresh-look-for-the-standard-model/

7.1.4 Matters

Quarks and Gluons

The previous examination shows that quantum electrodynamics can be understood in terms of fields, gauge bosons as vehicles of the fields, and particles. Also, other particles can be treated with quantum-field theory. This has been subsequently integrated with both the so-called *chromodynamics* for the strong force and the model for weak interactions. The theory covering these three kinds of interaction is called the *standard model* (see Table 1.3, Sect. 1.3, and Fig. 7.3).

Let us have a short look to quarks and weak interaction. The goal is to make the reader more acquainted with the quantum-field treatment of basic problems in physics. In 1968, deep inelastic scattering experiments at the Stanford Linear Accelerator Centre (SLAC) showed that the proton contained much smaller, point-like objects and was therefore not an elementary particle as it was previously believed. This was the experimental evidence for the model of quarks developed by the American physicist Murray Gell–Mann. Quarks are the constituents of massive particles like protons and neutrons,[32] and come in three generations of doublets: (d, u), (s, c), (b, t), standing for the (quite bizarre) names of the so-called *flavours*: (down, up), (strange, charm), (bottom, top). Like other particles, quarks have of course their antiquarks.

Let us consider how quarks are combined. Note that, the strong interaction between the quarks comes to dominate their quantum-mechanical resistance to localisation. A quantity that needs to be conserved is called *colour*, a charge that has of course nothing to do with ordinary colours (see Fig. 7.4). Essentially, quarks can assume any of the 3 colours (represented by red, blue and green) and antiquarks any

[32] Gell-Mann (1964). The model was independently worked out also by the Russian–American physicist George Zweig. For a very accessible (but not formal) introduction to this subject see Watson (2004). For a formal treatment see Greiner et al. (1994).

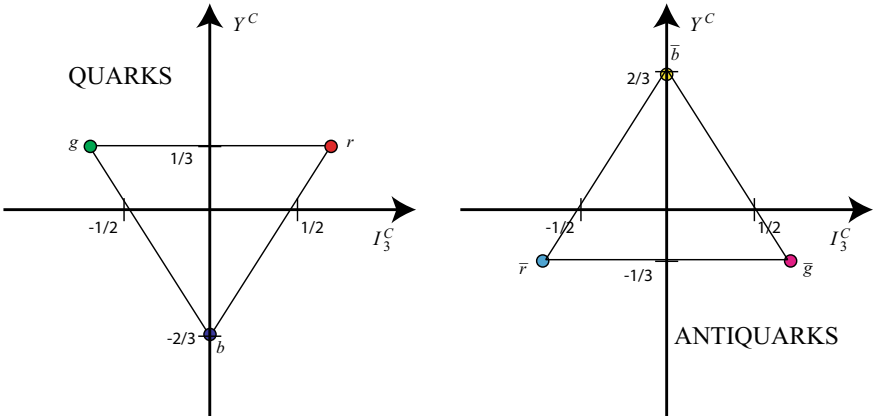


Fig. 7.4 Colour states (7.140) (as well as anticolour states) can be labelled by two quantum numbers: the colour hypercharge Y^C and the isospin in the z direction I_3^C. The isospin or isobaric spin is not an ordinary spin but is only related to the strong interaction (see Table 7.1). Particles that are affected equally by the strong force but have different electric charges (e.g. protons and neutrons) can be treated as being different states of the same particle with isospin values related to the number of charge states. Also, hypercharge is related to strong interaction and unifies isospin and flavour into a single charge. In fact, according to the Gell–Mann–Nishijima formula we have $Q = I_3 + 2^{-1}Y^C$, where Q is the electric charge, as well as $Y^C = B + S$, where B is the baryon number and S is strangeness (after the discovery of charm, top, and bottom quark flavours, the latter formula has been generalised)

of the 3 anticolours (antired, represented as cyan, antiblue, represented as yellow, and antigreen, represented as magenta). This means that we have 6 quark flavours times 3 colours, namely 18 quarks (and 18 antiquarks). The 3 colours can be, for example represented by 3 columnar vectors constituting a orthonormal basis

$$|r\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \; |b\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \; |g\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{7.140}$$

Hadrons (that comprehend baryons and mesons) need to have a neutral or white colour, and quarks need to be combined in appropriate way (Fig. 7.5). The fact that we have only 'neutral' particles, shows that colour is in fact unobservable. Note that for the above reason, the allowed combinations must give rise to matrices with 0 trace like the Pauli spin matrices (1.324).

The strong interaction connecting quarks is mediated by massless bosons called *gluons*, possessing themselves colour. Since gluons both respond to the presence and motion of colour charge and they carry themselves colour charge, then, differently from photons (that only mediates among electrons), they respond directly to one another (Fig. 7.6). This is a remarkable difference relative to the electromagnetic case and is mathematically rooted in the fact that, while the U(1) group of the electromagnetic theory is Abelian (commutative), the group of the strong force is

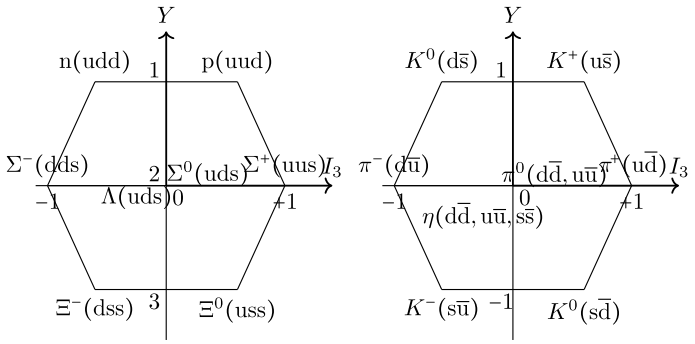


Fig. 7.5 On the left the baryon octet; on the right the light meson octet. Note that, Σ and Ξ baryons are short living. The baryon octet can be expanded to a decuplet by inserting Δ and Ω baryons, which are also short living. Recently (2017), the $\Xi^{++}(\mathrm{ccu})$ particle has been discovered. It is the sole baryon to have two massive (charm) quarks, giving a total mass of 3,621 MeV

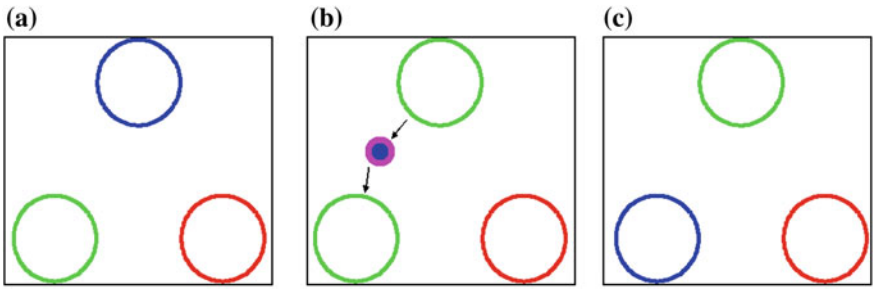


Fig. 7.6 The figure shows both the conservation of colour and how changes of colour happen. **a** An initial particle (e.g. a proton) with the three quarks, each with a different colour (the sum is white). **b** The top quark emits a blue-antigreen gluon (in fact, the gluon is in a superposition of blue-antigreen and green-antiblue, that leaves a green quark behind. **c** Interacting with the bottom quark on the left, it causes the latter to become blue. Adapted from http://en.wikipedia.org/wiki/Color_charge

not. Gluons have always two (unbalanced) colours (see Fig. 7.7), so that, although, abstractly speaking, $3 \times 3 = 9$ different colour gluons are possible, there is one of them (the singlet) that responds equally to all charges and therefore cannot be realised in our universe (as experiments show).[33]

The baryon number is a strictly conserved additive quantum number of a system (Table 7.1) and it is defined as

$$B = \frac{1}{3}\left(n_{\mathrm{q}} - n_{\bar{\mathrm{q}}}\right), \tag{7.141}$$

[33] Wilczek (2006, pp. 85–86).

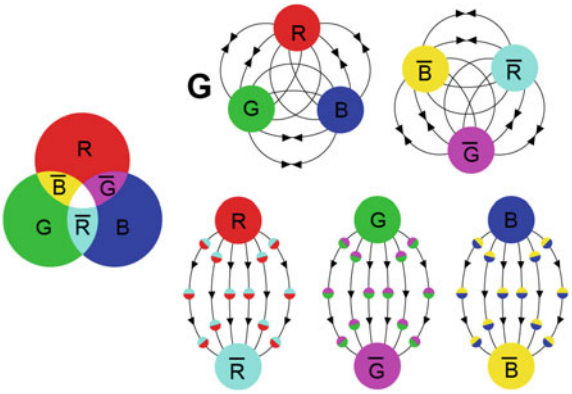


Fig. 7.7 Fields due to colour charges, as in quarks (**G** represents here the gluon field). These are neutral combinations. The overbar denotes anticolours (and antiparticles). Top: Colour charge has "ternary neutral states", as well as binary neutrality (analogous to electric charge) generating baryons like protons and neutrons. Bottom: Quark–antiquark combinations generating pions (π mesons). Adapted from http://en.wikipedia.org/wiki/Color_charge

Table 7.1 The six quark flavours (up, down, strange, charm, bottom, top) and their attributes. Note that, we cannot have massive (composed) particles with fractional charge or fractional baryon number. This can be considered another manifestation of quantisation and explains why isolated quarks cannot be found while stable massive particles always are composed of three quarks, and this in turn implies that fractional electric charge is also unobservable. Note that, antiquarks have the same electric charge and baryon number as quarks but with opposite sign

Quark attribute	Quark flavour					
	u	d	s	c	b	t
Electric charge	$+2/3$	$-1/3$	$-1/3$	$+2/3$	$-1/3$	$+2/3$
Mass [GeV]	0.004	0.007	0.135	1.3	4.2	174
Isospin I	$+1/2$	$+1/2$	–	–	–	–
Isospin I_3	$+1/2$	$-1/2$	–	–	–	–
Strangeness	–	–	-1	–	–	–
Charm	–	–	–	$+1$	–	–
Bottomness	–	–	–	–	-1	–
Topness	–	–	–	–	–	$+1$
Baryon number	$1/3$	$1/3$	$1/3$	$1/3$	$1/3$	$1/3$

where n_{q} is the number of quarks, and $n_{\bar{\mathrm{q}}}$ is the number of antiquarks. Baryons (composed of three quarks) have a baryon number of +1, while antibaryons (composed of three antiquarks) have a baryon number of -1; mesons, like gluons (composed of one quark and one antiquark) as well as particles not composed of quarks (like electrons), have a baryon number of 0.

Note that according to Table 7.1, quarks possess fractional electric charge (although in Nature it is not observable), what shows that the assumed indivisible quanta are not absolutely fundamental things.[34]

Weak Interaction and Higgs Field

I recall that parity refers to the symmetry between right and left (Sect. 1.3.1), while chirality refers to an intrinsic distinction between right and left. As mentioned, one of the fundamental processes of our world, the weak interaction (which is responsible of both radioactive decay and nuclear fusion, when two nuclei collide and merge to form a new nucleus), distinguishes between right and left at the opposite of many physical subatomic processes (it may be noted that also biochemicals show often chirality).[35] We could support a principle of maximal parity non-invariance according to which only left-handed (LH) particles and right-handed (RH) antiparticles participate in the weak interaction (and this would restore some symmetry). However, this would be inconsistent with relativity: what appears as an LH electron to a stationary observer appears as RH to an observer moving in the electron's direction. However, with photons, which are massless and moving at the maximal speed attainable in our physical universe, cannot be overtaken, this incongruence disappears. Therefore, the natural habitat of chirality is a massive world, and this is insightful since one of the main current physical investigations is about how masses have arisen at all in our universe during its evolution starting from a world populated from massless bosons, a process that gives rise to some asymmetries.

Thus, one of the main problems of the standard model of particles physics is the fact that W and Z bosons of the weak force are massive while they should be massless for preserving gauge symmetry. By the early 1960s, physicists had realised that a given symmetry law might not always be followed (or 'obeyed') under certain conditions. For solving the previous problem, three different teams of scientists proposed in the 1960s that there is a scalar field (which, I recall, does not change under Lorentz transformations (2.70)), known subsequently as Higgs field (after the name of the British theoretical physicist Peter Higgs), that permeates our universe and takes a non-zero constant value almost everywhere.[36] Since the Higgs field is a scalar field, the Higgs boson (the particle that mediates the field) has no spin (see previous subsection). The Higgs boson is also its own antiparticle, is CP-even, and has zero electric and colour charge. Thus, we can explain how W and Z bosons acquire subsequently mass when interacting with the Higgs field. Later on, this model was

[34] Laughlin (2005, pp. 76–78). Similar considerations are developed for the fractional quantum Hall effect (whose discovery conferred the Nobel Prize in physics to R. Laughlin).

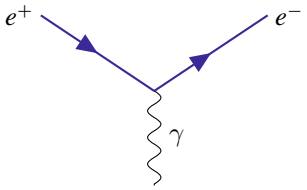
[35] Wilczek (2006, pp. 77–78).

[36] Higgs (1964). In fact, there were simultaneous contributions of two other teams, whose members were Robert Brout and François Englert, on the one hand, and Gerald Guralnik, C. Richard Hagen, and Tom Kibble, on the other.

thought to be also helpful for explaining how other particles (like electrons or quarks) acquire the specific mass that they have.

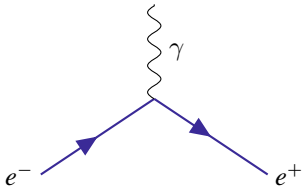
Feynman's Diagrams

R. Feynman introduced very helpful diagrams making us visualise possible interactions in which both particles and antiparticles are involved. Of course, such diagrams are known as *Feynman's diagrams*.[37] Since in Sect. 1.2.5 we have remarked that, according to the CPT theorem, there is conservation of product of charge conjugation, parity and time, this means that antiparticles (see Table 1.3) must not only have an opposite charge but also opposite motion direction and reversed time. In other words, antiparticles must evolve backwards in time and their spatial direction be the mirror image of that of the relative particle. This shows how important is symmetry in QM. For instance, the generation of the pair electron–positron can be represented as follows:



(7.142)

where it is understood that our arrow of the time goes from the bottom to the top. In other words, in certain conditions, a photon can produce a particle–antiparticle pair. This is an example of the conversion of energy into mass according to Einstein's formula (1.5). Analogously, the pair electron–positron annihilation can be represented as

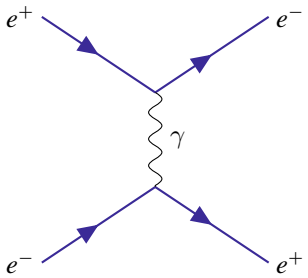


(7.143)

where the same convention for time has been followed. In this case, the pair particle–antiparticle has given rise to a photon. It is an example of conversion of mass into energy, always in agreement with Einstein's formula (1.5). These kinds of diagrams

[37] Feynman (1949). See also Bjorken and Drell (1964, Chaps. 6–7).

can obviously be combined. For instance, we can represent a pair annihilation followed by a pair creation as



(7.144)

It may be noted here that if we interpret both the creation and the annihilation as quantum events (what appears necessary), these diagrams confirm that quantum events are in principle independent of space-time, as recalled in Sects. 5.1.1, 5.1.5, and 6.2.5. As we shall see, they can be even considered the *source of two opposite space-time orders*.

The single-photon annihilation of an electron–positron pair cannot occur in free space because it is impossible to conserve energy and momentum together in this process. However, in the Coulomb field of a nucleus, the translational invariance is broken and single-photon annihilation may occur. The reverse reaction (always in free space and without an atomic nucleus) is also impossible for this reason. In quantum-field theory, such processes are allowed only as *intermediate* quantum states for times short enough that the violation of energy conservation can be accommodated by the uncertainty relation (1.292). This opens the way for *virtual* pair production or annihilation in which a one-particle quantum state may fluctuate into a two-particle state and back. These processes are important in the vacuum state and renormalisation of a quantum-field theory (see Eq. (7.53)).

Thus, in a particle universe (like our own), antiparticles can only be virtual (and vice versa for a world of antiparticles). In some cases, however, it is possible to boost the pair apart using external energy so that they avoid annihilation and can become *actual* particles. This may occur in one of two ways. In an accelerating frame of reference, the virtual particles may appear to be actual to the accelerating observer; this is the Unruh effect, on which I shall come back. In short, the vacuum of a stationary frame appears, to the accelerated observer, to be a warm gas of actual particles in thermodynamic equilibrium. An astronaut at rest may think he or she is in a vacuum, whereas an astronaut in an accelerating spaceship will feel immersed in a thermal bath of innumerable particles. This discrepancy between viewpoints also occurs at the perimeter of black holes and leads to paradoxical conclusions about the fate of in-falling matter.[38] Another example is pair production in very strong electric fields, sometimes called vacuum decay. For example, suppose a pair of atomic nuclei are merged to very briefly form a nucleus with an electric charge

[38] For a non-technical summary, see Susskind (1997).

greater than about 140, that is, larger than about the inverse of the (dimensionless) fine structure constant,[39] also known as Sommerfeld's constant, after the name of the German physicist Arnold Sommerfeld (1868–1951),

$$\alpha = \frac{1}{4\pi\epsilon_0}\frac{e^2}{\hbar c} = 0.0072973525664(17), \tag{7.145}$$

characterising the strength of the electromagnetic interaction between elementary charged particles, where I recall that ϵ_0 is the electric permitittivity. Then, the strength of the electric field will be such that it will be energetically favourable to create positron-electron pairs out of the vacuum, with the electron attracted to the nucleus to annihilate the positive charge. It can be assumed that our universe came out of vacuum according to processes of this kind and that particles were generated out of photons on these outlines.[40] Since, according to formula (3.79), the Boltzmann entropy of a photon gas is proportional to the number of photons, we can build a quantity (see Eq. (7.141))

$$\eta = \frac{n_{\mathfrak{b}} - n_{\bar{\mathfrak{b}}}}{n_\gamma}, \tag{7.146}$$

where $n_{\mathfrak{b}}$ and $n_{\bar{\mathfrak{b}}}$ are the numbers of baryons and antibaryons, respectively, and n_γ is that of photons. This quantity is inversely proportional to the entropy of baryons.

Let me remark, in conclusion, that Feynman diagrams can be considered as a pictorial representation of propagators (both in forward and backward time direction) on the outline of those introduced in Sect. 6.2.4.

7.1.5 A Basic Exploration of General Relativity

Dealing with fields and the relations between QM and relativity, a short look at general relativity (GR) is important. In fact, GR simply eliminates the notion of gravitational field and therefore could put in danger the integration between QM and relativity. This is clearly a critical point that needs some scrutiny. When we deal with GR in a quantum-mechanical context, it is important to consider two different aspects[41]:

- As for SR, according to GR, the space-time manifold represents a classically deterministic and continuous structure, while, as said, according to QM we should expect relevant fluctuations at the Planck scale (of the order of the squares of Planck time (3.42a) and Planck length (3.42b)): at Planck scale, we have no time and no length in the proper sense. In fact, it may be noted that from Eqs. (3.42), we get

[39] Tegmark et al. (2006).

[40] Serjeant (2010, Chap. 2).

[41] Misner et al. (1970, pp. 12–13).

$$G = \frac{c^5 t_P^2}{\hbar} \text{ or } G = \frac{c^3 l_P^2}{\hbar}, \tag{7.147}$$

which allows us to interpret the gravitational constant G as a limitation on the discernibility of space and time (or space-time).[42] In general, the idea of a continuum appears not fully compatible with quantum-mechanical discreteness (Sect. 1.1.2).

- A collection of particles like electrons, even with zero density of electric charge (when e^- and e^+ world lines are present in equal numbers), according to GR, will have a density of mass that will curve the manifold under study. On the other hand, "investigation in infinite details mean unlimited density, and unlimited disturbance of the geometry".

Introduction

GR may have started by Einstein's conviction that there cannot exist something (like it would be the case with absolute space) without being acted upon in some way.[43] The basic idea of GR can be said to be that every physical quantity must be describable by geometry, so that there are in fact no forces and no fields.[44]

At the core of GR is the relation of mass energy. According to Einstein's hypothesis, every form of energy has a mass equivalent in agreement with Eq. (1.5)[45]: (i) if all mass exerts and suffers gravity, we would expect even (the energy of) an electromagnetic field to exert a gravitational attraction, and conversely, light to bend under gravity. (ii) We shall expect a gravitational 'field' itself to gravitate. (iii) Stretched or compressed objects have (minutely) more mass by virtue of the stored elastic energy. (iv) The total mass of the separate components of a stable atomic nucleus always exceeds the mass of the nucleus itself, since energy (that is, mass) would have to be supplied in order to decompose the nucleus against the nuclear binding forces. This is the reason for the well-known 'mass defect' (on which I shall come back). One kind of energy that does not contribute to mass is potential energy of position.

In CM, a particle moving in an electromagnetic (or gravitational) field is often said to possess potential energy, so that the sum of its kinetic and potential energies remains constant (Eqs. (1.12) and (1.43)). This is a useful view, but energy conservation can also be satisfied by debiting the field with an energy loss equal to the kinetic energy gained by the particle.[46] In relativity there are good reasons for adopting the latter alternative: the 'real' location of any part of the energy is no longer a mere convention, since energy (as mass) gravitates; that is, it contributes measurably (at least in principle) to the curvature of space-time at its location. The stress–energy tensor (7.77) is the source of the gravitational 'field' in the Einstein field equations

[42] Cohen-Tannoudji (1991, pp. 112–117).
[43] See Rindler (2001, pp. 7 and 90).
[44] Misner et al. (1970, p. 48).
[45] Rindler (2001, p. 114).
[46] Rindler (2001, p. 113).

of general relativity, just as mass density is the source of such a field in Newtonian gravity.

However, we expect that physics is everywhere locally Lorentzian (i.e. with a flat space).[47] Moreover, for a weak gravitational field with slowly moving sources, Newton's and Einstein's equations of motion are essentially equivalent.[48] A second caveat is: there certainly are GR solutions where the local standard of non-acceleration does not accord with the matter distribution. Thus, while in GR all matter, including its motion, undoubtedly affects local inertial behaviour, it appears not entirely to cause it.[49]

A third caveat is the following[50]: in relativity, it is especially important to distinguish between the set of events that an observer *sees* (directly experiences) at one instant and the set of events that the observer considers (infers) to have occurred at that instant. What an observer actually sees or can photograph at one instant is called as world picture. In Sects. 5.1.1, 6.2.5, and 6.3.4, I have already insisted on this distinction, but it becomes even more crucial in such a context.

A final caveat is that the expression *test particle* is preferred to that of "point particle" as far as it is assumed that the former shows negligible mass and energy to avoid to treat it as a field source (see also Sect. 7.1.1).[51]

Equivalence Principle

It may be said that the equivalence principle is the vehicle between flat and curved space-time.[52] There are at least two quite distinct types of mass entering into Newton's theory of mechanics and gravitation. These are

- The *inertial* mass, which occurs as the ratio between force and acceleration in Newton's second law (1.37) and thus measures a particle's resistance to acceleration, and
- The *gravitational* mass, which may be regarded as the gravitational analog of electric charge.

Why inertial mass (whose significance as "resistance to acceleration" makes apparently sense even in a world without gravity) should serve as gravitational 'charge' when there is gravity, is totally unexplained in Newton's theory and seems purely fortuitous. Thus, the proportionality of inertial and gravitational mass is a profoundly

[47] Misner et al. (1970, p. 19).

[48] Rindler (2001, pp. 221–223).

[49] Rindler (2001, p. 8). Note that in the same textbook (pp. 22, 35) the complex relations between Mach's standpoint and relativity theories, especially concerning the notion of inertia, are studied.

[50] Rindler (2001, pp. 61–62).

[51] Malament (2012, p. 105).

[52] Misner et al. (1970, p. 207).

mysterious fact. Such an assumed proportionality is often called the *weak equivalence principle*.[53]

A fully equivalent formulation is that all free particles experience the same acceleration at a given point in a gravitational field. More precisely, the field times passive mass gives the force, and the force divided by inertial mass gives the acceleration, so the acceleration equals the field, $a = g$, independently of the particle and its mass. Thus, a uniform gravitational field is equivalent to an acceleration and a free-falling body (not subject to any other field) does not experience gravity because it is 'cancelled' by the acceleration. It follows that free motion in a gravitational field is fully determined by the field and an initial velocity. This is a peculiar characteristic of the gravitational field and of no other force acting on bodies.

We can now state Einstein's *strong equivalence principle* as follows (which appears as a full counterpart (and generalisation) of Galilei's relativity: see Sect. 2.3.2): all freely falling non-rotating cabins are equivalent for the performance of all physical experiments. Moreover, the physics in all these cabins is SR, and the cabins themselves are called local inertial frames (LIFs). Such a principle is not so obvious. In fact, the eminent Irish relativist John Lighton Synge (1897–1995) tells us that, since every 'real' gravitational field g (as opposed to the 'fictitious' field in an accelerating rocket) is non-uniform, there will always be tidal forces (on which I shall say more below) present in the cabin, causing relative accelerations dg between neighbouring free particles.[54] And with perfect instruments these could be detected. As pointed out by Synge, the gravitational field is described by the Riemann tensor (on which we shall say more below), and according to whether it does vanish or not we either do not have or have a gravitational field, which is therefore an absolute property that does not depend on the observer's world line.[55] So, we should speak of the strong equivalence principle at most as a limiting property.

Geodesics and Metric

Gravitation manifests itself not only in its 'pulling' towards the centre of the 'field' but also in the deformation of space-time (this is the already mentioned tidal effect of gravity) (Fig. 7.8). In fact, in SR all worldlines of free systems are straight, while in GR they are only locally straight.[56] Following Minkowski's procedure in SR (Sect. 2.3.2), we can regard the set of all (actual or potential) events in the world as a 4D continuum (space-time), which can be coordinated with four arbitrary (Gaussian) coordinates x_μ ($\mu = 0, 1, 2, 3$).[57] According to the strong equivalence principle, we can find at each event a small freely falling box in which SR holds. This box being a LIF provides us with a local inertial coordinate system which can be used to assign

[53] See Rindler (2001, pp. 16–19, 36).

[54] Synge (1960, pp. VIII–X). See also Rindler (2001, p. 23).

[55] I also recommend the analysis in Friedman (1983, Sect. 5.4).

[56] See Rindler (2001, pp. 20–21).

[57] See Rindler (2001, pp. 177–178).

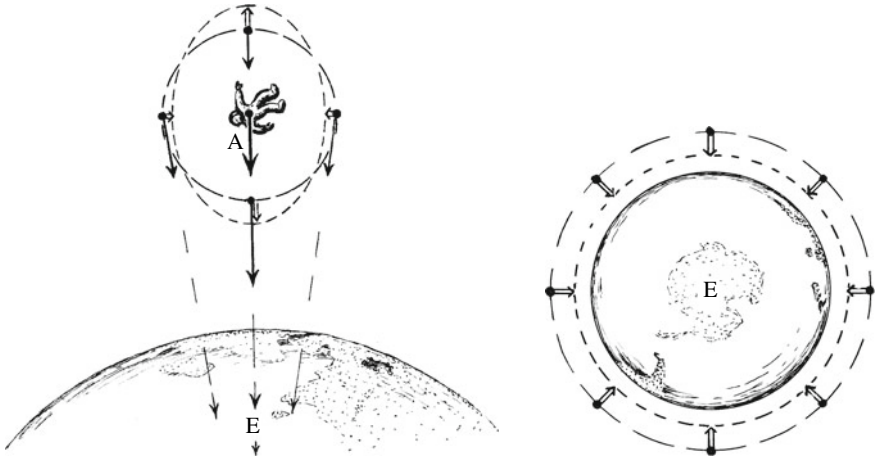


Fig. 7.8 Tidal (panel on the left) and contraction (panel on the right) effects of gravity. The tidal effect manifest itself in the stretching of a body into the direction pointing to the centre of the "field" and in squeezing it along the orthogonal direction. Adapted from Penrose (2004, p. 397)

a unique squared displacement from and to any neighbouring event according to the formula of the square distance (2.79). So there is a uniquely determinable metric structure on the global space-time. And since the formula for the metric is locally Minkowskian, as a generalisation of Riemannian (curved) space, it must be globally pseudo-Riemannian (the metric tensor, on which we shall come back, need not be positive definite). In GR, on the other hand, light in vacuum travels straight in the LIF, and therefore should travel geodesically in space-time, with the specification that, while free particles have timelike geodesic worldlines, the worldlines of photons need to be null. All these elements can be put in the form of axioms:

1. The space-time of events is pseudo-Riemannian with Minkowskian signature (it is often called Lorentzian).
2. Free test particles have timelike geodesic worldlines.
3. Timelike geodesics have the property of maximising the proper time of the test particle: the arc along any timelike worldline corresponds to c times the proper time of an ideal point clock that traces it out.
4. Light in vacuum follows null geodesics.
5. Einstein's field equations will relate the metric with the energy tensor of the sources.

For the theory to be useful, it should predict the orbits in the 'field' of a given mass distribution. Since the orbits (the geodesics) are now determined by the metric, we might expect the metric, in turn, to be determined by the sources—that is, by the gravitating matter. But there is an inherent logical obstacle to this: unless, we already know the metric, we do not know the space-time at the sources, and without that, we cannot precisely describe extended sources. So, it would seem that we are doomed

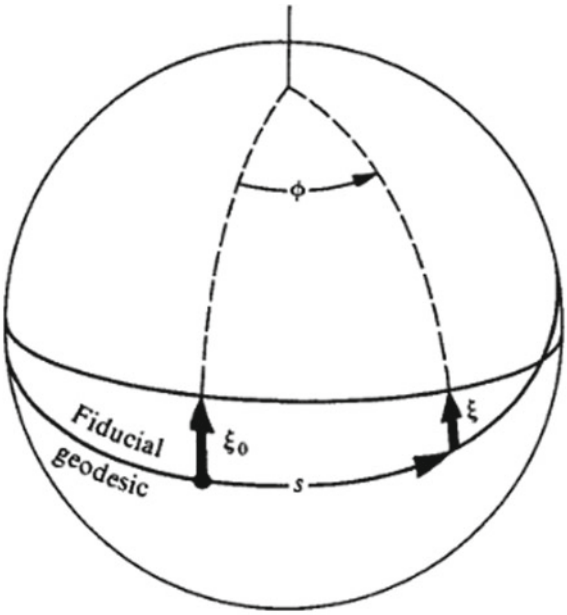


Fig. 7.9 Geodesic deviation in curved space-time. The geodesics, initially parallel and separated by a distance ξ_0, after a distance s are no longer parallel and their separation is now $\xi = \xi_0 \cos\phi = \xi_0 \cos(s/r)$, where r is the radius of the sphere and ϕ the angular separation between ξ_0 and ξ. Adapted from Misner et al. (1970, p. 31)

to use some sort of iterative mathematical process to get from the sources to the space-time.

On the basis of what said, in curved space-time, it is appropriate to think light cones as local structures in space-time, i.e. in the tangent space at the source event.[58] Since gravitation will affect both space and time, the presence of a mass will slow down clocks according to the time dilation

$$t_d(r) = e^{\frac{1}{c^2}\int_0^r g(r')dr'}, \tag{7.148}$$

where r is a ray (whose magnitude has been considered here) along the direction of the gravitational 'force' g. Moreover, the tidal effect determines geodesic deviations (Fig. 7.9; see also Fig. 5.6, Sect. 5.1). The equation for geodesic deviation is

$$\frac{d^2\xi}{ds^2} + R\xi = 0, \tag{7.149}$$

with R being the Gaussian curvature of the surface, which in the simplest case of a sphere is $R = 1/a^2$ (a being its ray), and where ξ is the separation from a fiducial geodesic.[59] What is crucial to understand is that, since their definition depends only on distance measurements in the surface, *geodesics are intrinsic* and remain geodesics

[58] Penrose (2004, p. 402).

[59] Misner et al. (1970, p. 30).

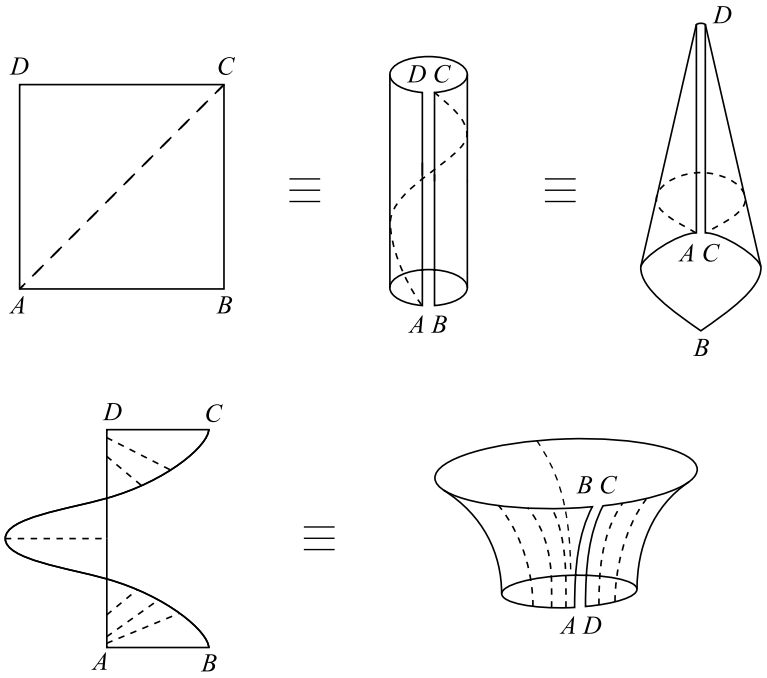


Fig. 7.10 Intrinsicity of geodesics: they are invariant under geometric deformation. Adapted from Rindler (2001, p. 166)

even when the surface is bent. This is so because also the Gaussian curvature is intrinsic.[60] This points to another characteristic of geodesics: they are the straightest possible lines on a curved surface (Fig. 7.10). For a sufficiently 'well-behaved' space, the two basic theorems on geodesics also hold: (i) there is a unique geodesic issuing from a given point P in a given direction and (ii) in a sufficiently small neighbourhood of point P each other point can be connected to P by a unique geodesic. Consider all the geodesics issuing from a point P in the directions of a linear 'pencil' $\lambda\mathbf{p} + \mu\mathbf{q}$ determined by two directions $\mathbf{p}$ and $\mathbf{q}$ at P. Such geodesics are said to generate a geodesic plane through P. Its curvature K at P is said to be the space curvature $K(\mathbf{p}, \mathbf{q})$ at P for the orientation $(\mathbf{p}, \mathbf{q})$. (In three dimensions K is completely known if it is known for 6 orientations, in four dimensions if it is known for 20.) A geodesic plane is the curved-space analog of a plane through a point, except that in general, it satisfies its defining property only with respect to that one point.[61] If K at P is independent of the orientation, P is an isotropic point.

[60] See Rindler (2001, pp. 166, 168).

[61] Rindler (2001, pp. 169–170).

Parallel Transport and Curved Space-time

In curved space-time having non-orthonormal but linearly independent basis vectors $\{\mathbf{e}_\alpha\}$ (such that any vector $\mathbf{u}$ can be written $\mathbf{u} = u^\alpha \mathbf{e}_\alpha$), the rules are identical to those of flat space-time except that we need to replace[62] (i) the covariant Lorentz component $g_{\alpha\beta}$ of the metric (Eq. (2.88)) by the GR components of the metric

$$g_{\alpha\beta} = \mathbf{g}(\mathbf{e}_\alpha, \mathbf{e}_\beta) := \mathbf{e}_\alpha \cdot \mathbf{e}_\beta, \tag{7.150}$$

with contravariant components $g^{\alpha\beta}$ replaced by $g^{\beta\gamma}$, so that

$$g_{\alpha\beta} g^{\beta\gamma} = \delta_{\alpha\gamma}, \tag{7.151}$$

and (ii) the Lorentz transformation matrix $\Lambda^{\alpha'}_{\beta}$ and its inverse $\Lambda^{\beta}_{\alpha'}$ (Eqs. (2.82)), where in such a context I have primed the indexes themselves, are replaced by an arbitrary nonsingular transformation matrix $\mathbb{L}^{\alpha'}_{\beta}$ and its inverse $\mathbb{L}^{\beta}_{\alpha'}$, such that

$$\mathbf{e}_\beta = \mathbf{e}_{\alpha'} \mathbb{L}^{\alpha'}_{\beta}, \quad \mathbf{e}^\beta = \mathbb{L}^{\beta}_{\alpha'} \mathbf{e}^{\alpha'}. \tag{7.152}$$

The Levi-Civita tensor like the metric tensor have elements that depend on the basis vectors. Let us consider a tensor field $\hat{\mathbf{T}}$ and introduce its gradient $\boldsymbol{\nabla}\hat{\mathbf{T}}$, with components (in the local Lorentz frame, these components are the directional derivatives of the components of $\hat{\mathbf{T}}$)

$$T^{\beta}_{\ \alpha,\gamma} = \frac{\partial T^{\beta}_{\ \alpha}}{\partial x^\gamma} = \boldsymbol{\nabla}_{\mathbf{e}_\gamma} T^{\beta}_{\ \alpha}. \tag{7.153}$$

We have that $\boldsymbol{\nabla}\hat{\mathbf{T}}$ gives $\boldsymbol{\nabla}_{\mathbf{u}}\hat{\mathbf{T}}$ when $\mathbf{u}$ is the last component of $\hat{\mathbf{T}}(\ldots, \mathbf{u})$. Then, no change in such Lorentz components would mean flat space and therefore parallel transport. Thus, we are asking here what would be the change in $\hat{\mathbf{T}}$ relative to the flat case.[63] To this purpose, let us define the covariant derivative $\boldsymbol{\nabla}_{\mathbf{u}}\hat{\mathbf{T}}$ of $\hat{\mathbf{T}}$ along a curve $f(\lambda)$, whose tangent vector is

$$\mathbf{u} = \frac{df}{d\lambda}, \tag{7.154}$$

namely (Fig. 7.11):

$$\left(\boldsymbol{\nabla}_{\mathbf{u}}\hat{\mathbf{T}}\right)_{\text{at } f(0)} := \lim_{\epsilon \to 0} \frac{\delta\hat{\mathbf{T}}}{\epsilon} = \lim_{\epsilon \to 0} \left\{ \frac{\hat{\mathbf{T}}[f(\epsilon)] - \hat{\mathbf{T}}[f(0)]}{\epsilon} \right\}, \tag{7.155}$$

where $\lambda = \epsilon$ is a point close to $\lambda = 0$ on the curve $f(\lambda)$. At that limit clearly, the field $\hat{\mathbf{T}}$ is parallel to the vector $\mathbf{u}$ and we have $\boldsymbol{\nabla}_{\mathbf{u}}\hat{\mathbf{T}} = 0$. We can then define a geodesic as

[62] Misner et al. (1970, pp. 201–202).

[63] Misner et al. (1970, pp. 208–211).

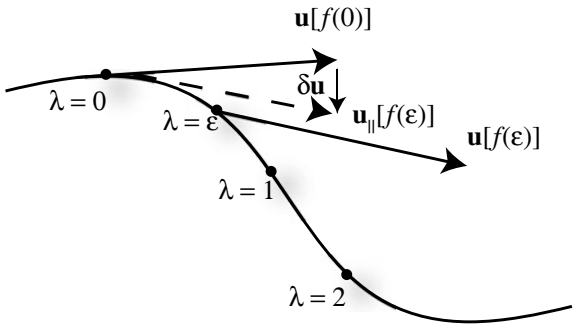


Fig. 7.11 Parallel transport in curved space. The curve λ represents a world line. From an operative point of view, we: (1) choose a point $f(0)$ on the curve λ at which we evaluate $\nabla_{\mathbf{u}}\mathbf{u}$; (2) choose a nearby point $f(\epsilon)$ on the curve; (3) parallel transport the vector $\mathbf{u}[f(\epsilon)]$ back to $f(0)$ getting the vector $\mathbf{u}_{\parallel}[f(\epsilon)]$; and finally, (4) evaluate the difference $\delta\mathbf{u} = \mathbf{u}_{\parallel}[f(\epsilon)] - \mathbf{u}[f(0)]$. In this way, we have evaluated the curvature. Adapted from Misner et al. (1970, p. 209)

a curve that parallel transports its tangent vector (7.154) along itself

$$\nabla_{\mathbf{u}}\mathbf{u} = 0. \tag{7.156}$$

Holonomy measures the extent to which the parallel transport along a closed loop fails to preserve the geometry.

In a local Lorentz frame we have components (7.153). It is not so for a general basis. Let $\{\mathbf{e}_\beta(f)\}$ vary arbitrarily but smoothly from point to point and $\{\mathbf{w}^\alpha(f)\}$ its dual basis, which is a 1 form (Fig. 7.12; see also Figs. 1.15–1.16, Sect. 1.2). Then,

$$\nabla\hat{\mathbf{T}} = \nabla\left(T^{\beta}_{\ \alpha}\mathbf{e}_\beta \wedge \mathbf{w}^\alpha\right) \tag{7.157}$$

will have contributions from both $\nabla\mathbf{e}_\beta$ and $\nabla\mathbf{w}^\alpha$ as well as from

$$\nabla T^{\beta}_{\ \alpha} = T^{\beta}_{\ \alpha,\gamma}\mathbf{w}^\gamma. \tag{7.158}$$

For quantifying these contributions (in order to account for twisting, turning, expansion, and contraction of both the basis vectors and the 1 form), we introduce the *connection coefficients*

$$\Gamma^{\alpha}_{\ \beta\gamma} := \langle \mathbf{w}^\alpha, \nabla_{\mathbf{e}_\gamma}\mathbf{e}_\beta\rangle, \text{ with } \langle \nabla_{\mathbf{e}_\gamma}\mathbf{w}^\alpha, \mathbf{e}_\beta\rangle = -\Gamma^{\alpha}_{\ \beta\gamma}, \tag{7.159}$$

where the above one represents the α component of change in $\mathbf{e}_\beta$, relative to parallel transport, along $\mathbf{e}_\gamma$. In terms of these coefficients and of the terms (7.153), the components of the gradient, denoted now by $T^{\beta}_{\ \alpha;\gamma}$, are

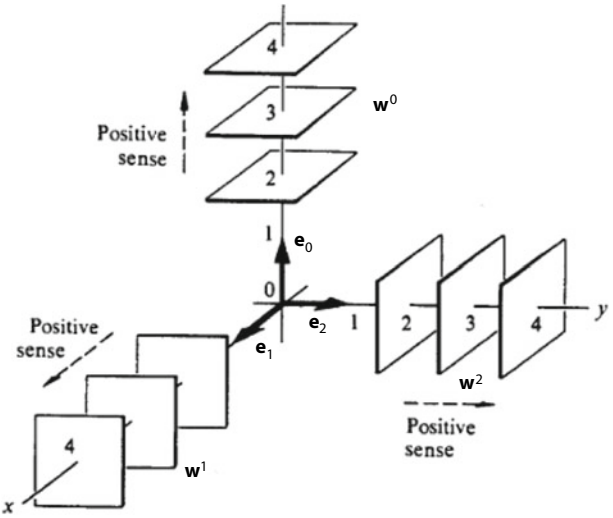


Fig. 7.12 Relation between basis vectors and dual 1 forms. Note that, according to the general properties of orthonormality, $\langle \mathbf{w}^\alpha, \mathbf{e}_\beta \rangle = \delta^\alpha_\beta$. Adapted from Misner et al. (1970, p. 60)

$$
\begin{aligned}
T^{\beta}_{\ \alpha;\gamma} &= T^{\beta}_{\ \alpha,\gamma} + \Gamma^{\beta}_{\ \mu\gamma} T^{\mu}_{\ \alpha} - \Gamma^{\mu}_{\ \alpha\gamma} T^{\beta}_{\ \mu} \\
&= \nabla_{\mathbf{e}_\gamma} T^{\beta}_{\ \alpha} + \langle \mathbf{w}^\beta, \nabla_{\mathbf{e}_\gamma} \mathbf{e}_\mu \rangle T^{\mu}_{\ \alpha} + \langle \nabla_{\mathbf{e}_\gamma} \mathbf{w}^\mu, \mathbf{e}_\alpha \rangle T^{\mu}_{\ \beta}.
\end{aligned} \tag{7.160}
$$

As said, if the basis at which the event $\nabla \hat{\mathbf{T}}$ is calculated were a local Lorentz frame, then the elements of $\nabla \hat{\mathbf{T}}$ would be just the $T^{\beta}_{\ \alpha,\gamma}$ (the first term on the RHS of Eq. (7.160)). When it is not so, such a Lorentz frame need to be corrected for twisting, turning, expansion and contraction of both the basis vectors and the 1 form.

Riemann Tensor

The geodesic deviation is a measure of the curvature of space-time. Let us consider a family of geodesics $f(\lambda, n)$ (Fig. 7.13). The smoothly varying parameter n discriminates among the different geodesics.[64] For fixed n, $f(\lambda, n)$ is a geodesic with tangent vector (7.154) fulfilling the condition (7.156). Then, the vector

$$
\mathbf{n} := \frac{\partial f}{\partial n} \tag{7.161}
$$

measures the separation between points with the same value of λ on neighbouring geodesics. Suppose that an observer falling freely along a fiducial geodesic $n = 0$ watches a test particle along the test geodesic $n = 1$. The velocity of the latter

[64] Misner et al. (1970, pp. 218–219).

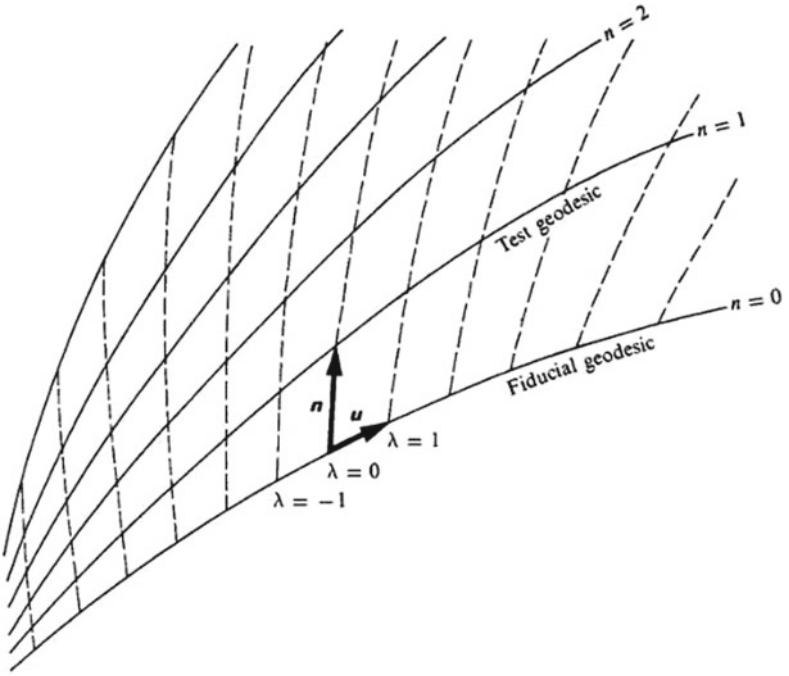


Fig. 7.13 A family of geodesics with selecting parameter n. Adapted from Misner et al. (1970, p. 219)

relative to the observer is given by $\nabla_{\mathbf{u}}\mathbf{n}$. Such a relative velocity as well as the separation vector $\mathbf{n}$ are arbitrary initial conditions but what is not arbitrary is the relative acceleration of the test particle with respect to the observer, given by $\nabla_{\mathbf{u}}\nabla_{\mathbf{u}}\mathbf{n}$. It would be zero in flat space-time but in curved space-time is given by

$$\nabla_{\mathbf{u}}\nabla_{\mathbf{u}}\mathbf{n} + \hat{\mathbf{R}}(\ldots, \mathbf{u}, \mathbf{n}, \mathbf{u}) = 0, \tag{7.162}$$

where $\hat{\mathbf{R}}$ is the Riemann tensor, after the name of B. Riemann. The relative acceleration in 3D and higher dimensions is a vector, and the following Eq. (7.149) one could be tempted to say that its components are $d^2\xi^\alpha/ds^2$ (Fig. 7.9), where $\boldsymbol{\xi}$ need to replace $\mathbf{n}$. However, this would not take into account the fact that change in length is only one factor of geodesic deviation, the other one being represented by rotations (Fig. 7.14).[65] Thus, it is suitable to introduce a new derivative symbol (D) that takes into account both contributions. By considering the proper time τ of the test particle, we can then write

$$\frac{D^2\boldsymbol{\xi}}{d\tau^2} + \hat{\mathbf{R}}(\mathbf{u}, \boldsymbol{\xi}, \mathbf{u}) = 0. \tag{7.163}$$

[65] Misner et al. (1970, pp. 31–37).

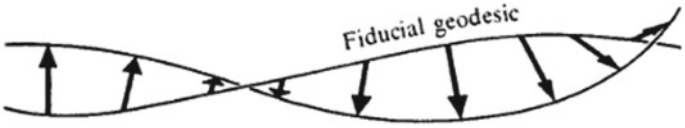


Fig. 7.14 Note that geodesic deviation manifests itself not only in the change of the length of vector ξ but also in possible rotations that it makes around the fiducial geodesic. Adapted from Misner et al. (1970, p. 31)

Considering that the Riemann tensor has trilinear components, we can rewrite the previous equation as

$$\frac{D^2\xi^\alpha}{d\tau^2} + R^\alpha{}_{\beta\gamma\delta}\frac{dx^\beta}{d\tau}\xi^\gamma\frac{dx^\delta}{d\tau} = 0, \tag{7.164}$$

which does for gravitation what Lorentz force law (which combines the effects of electric and magnetic forces on a point charge) makes for electromagnetism.[66] In fact, by rewriting the previous equation as

$$\frac{D^2\xi^\alpha}{d\tau^2} = -R^\alpha{}_{\beta\gamma\delta}u^\beta\xi^\gamma u^\delta, \tag{7.165}$$

we see that it is quite similar to the Lorentz force law

$$\frac{d^2x^\alpha}{d\tau^2} = \frac{\mathrm{e}}{m}\hat{F}^\alpha{}_\beta u^\beta. \tag{7.166}$$

Note that the Lorentz force law need to be rewritten in geometric terms by using the particle's own clock and its 4 velocity. In the previous equation, $\hat{F}^\alpha_\beta$ is the Maxwell tensor (7.89) or (7.91) but here is not expressed in covariant (or contravariant) components but rather as

$$\hat{F}^\alpha{}_\beta = \hat{F}^{\alpha\mu}g_{\mu\beta} = \begin{bmatrix} 0 & E_x & E_y & E_z \\ E_x & 0 & B_z & -B_y \\ E_y & -B_z & 0 & B_x \\ E_z & B_y & -B_x & 0 \end{bmatrix}, \tag{7.167}$$

with α being the rows and β being the columns.

As an example of Riemann tensor, consider two particles moving slowly at distance r from the Sun. In the standard and nearly inertial coordinates of celestial mechanics, all components can be neglected except $dx^0/d\tau = 1$. The space components of the equation for geodesic deviation are

$$\frac{d^2\xi^k}{d\tau^2} + R^k{}_{0j0}\xi^j = 0, \tag{7.168}$$

[66] Misner et al. (1970, pp. 72–73).

with

$$\begin{bmatrix} R^x_{0x0} & R^y_{0x0} & R^z_{0x0} \\ R^x_{0y0} & R^y_{0y0} & R^z_{0y0} \\ R^x_{0z0} & R^y_{0z0} & R^z_{0z0} \end{bmatrix} = \begin{bmatrix} m/r^3 & 0 & 0 \\ 0 & m/r^3 & 0 \\ 0 & 0 & -2m/r^3 \end{bmatrix}. \tag{7.169}$$

Einstein's theory can determine the value also of other terms but these nine ones are crucial for many applications.

From the Riemann tensor one can get several other curvature tensors by contraction.[67] One of the most used one is *Ricci curvature tensor*, after the name of the Italian mathematician Gregorio Ricci–Curbastro (1853–1925), whose components are

$$\mathbb{R}_{\mu\nu} = \Gamma^{\alpha}_{\mu\nu,\alpha} - \Gamma^{\alpha}_{\mu\alpha,\nu} + \Gamma^{\alpha}_{\beta\alpha}\Gamma^{\beta}_{\mu\nu} - \Gamma^{\alpha}_{\beta\nu}\Gamma^{\beta}_{\mu\alpha}. \tag{7.170}$$

Instead of, Einstein's tensor is a sort of average of the Riemann tensor over all directions.[68] In particular, in c units, it is equal to

$$\mathbf{G}_{\mu\nu} = \mathbb{R}_{\mu\nu} - \frac{1}{2} R \mathbf{g}_{\mu\nu}, \tag{7.171}$$

where $\hat{\mathbb{R}}$ is the Ricci tensor, $\mathbf{g}$ is the metric tensor and $R = \mathbf{g}^{\mu\nu}\mathbb{R}_{\mu\nu} = \mathbb{R}^{\nu}_{\nu}$ is the scalar curvature, i.e. the trace of the Ricci tensor with respect to the metric. Then, the *Einstein's field equations* can be written in natural units as

$$\mathbf{G}_{\mu\nu} + \Lambda \mathbf{g}_{\mu\nu} = 8\pi \mathbf{T}_{\mu\nu}, \tag{7.172}$$

where $T_{\mu\nu}$ is the stress–energy tensor (7.77) and Λ is the cosmological constant. The stress–energy of matter generates an average curvature in its neighbourhood. Simultaneously, the field equation is a propagation equation for the remaining, anisotropic, part of the curvature, governing not only the external space-time curvature of a static source but also the emission and propagation of gravitational waves. The latter represent a spectacular confirmation of GR. Gravitational waves, first proposed by Henri Poincaré, are ripples in the curvature of space-time that are generated by accelerated masses and propagate as waves outward from their source at the speed of light. In fact, they are produced when two black holes or two neutron stars collapse into each other. Their observation is relatively recent (from 2015 onwards) and requires giant interferometers of the kind shown in Fig. 7.15.

Einstein's Clock-Shutter Device

At the sixth Solvay Conference in 1930, Einstein proposed a device consisting of a box with a hole in one of its sides and a shutter moved by means of a clock inside

[67] Misner et al. (1970, pp. 220–222).

[68] Misner et al. (1970, pp. 39–43).

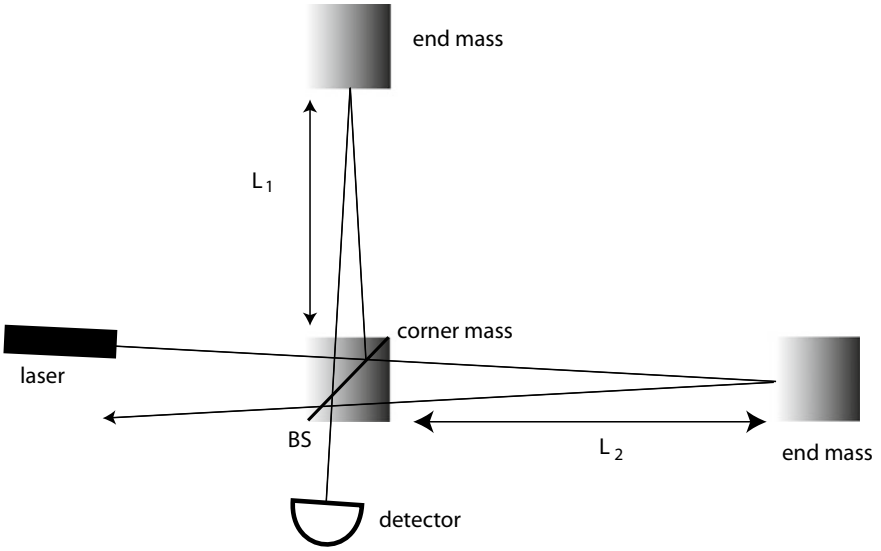


Fig. 7.15 A Michelson–Morley-type interferometer (see Fig. 2.11, Sect. 2.3) for detecting gravitational waves (top view). Three masses hang by wires from the overhead supports at the corner and ends of the interferometer. When the first crest of a gravitational wave enters the laboratory, its tidal forces should stretch the masses apart along the L_1 arm while squeezing them together along L_2. When the wave's first crest has passed and its first through arrives, the directions of stretch and squeeze will be changed. By monitoring the difference $L_1 - L_2$, one may look for gravitational waves. This is provided by a laser beam which shines onto a symmetric BS on the corner mass. The two outgoing beams go down the two arms and bounce off mirrors at the end of the arms and then return to the BS. The beams will be combined and split so that one part of each beam goes back to the laser and another part goes towards the photodetector. When no gravitational wave is present, the contributions from the two beams interfere in such a way that all the light goes back to the laser. Adapted from Auletta et al. (2009, p. 17)

the box.[69] If in its initial state the box contains a certain amount of radiation and the clock is set to open the shutter after a chosen short interval of time, it could be achieved that a single photon is released through the hole at a moment which is known as exactly as desired. Moreover, if we weight the box before and after this event, we could measure the energy of the photon as exactly as we want, and these two measurements would disprove the time–energy uncertainty relation (1.292). We see again that Einstein aimed at disproving the uncertainty relations (Sects. 2.3.1 and 3.3.1).

In his reply, Bohr first pointed out that, in order to successfully perform the experiment, the box needed to be suspended in a spring–balance support in the middle of a gravitational field, and to be furnished with a pointer to read its position on a graduate scale fixed to the balance support (Fig. 7.16). The weighing of the support may be thus performed with any given accuracy Δm. Bohr's reply was that any determina-

[69] Bohr (1949, pp. 224–228).

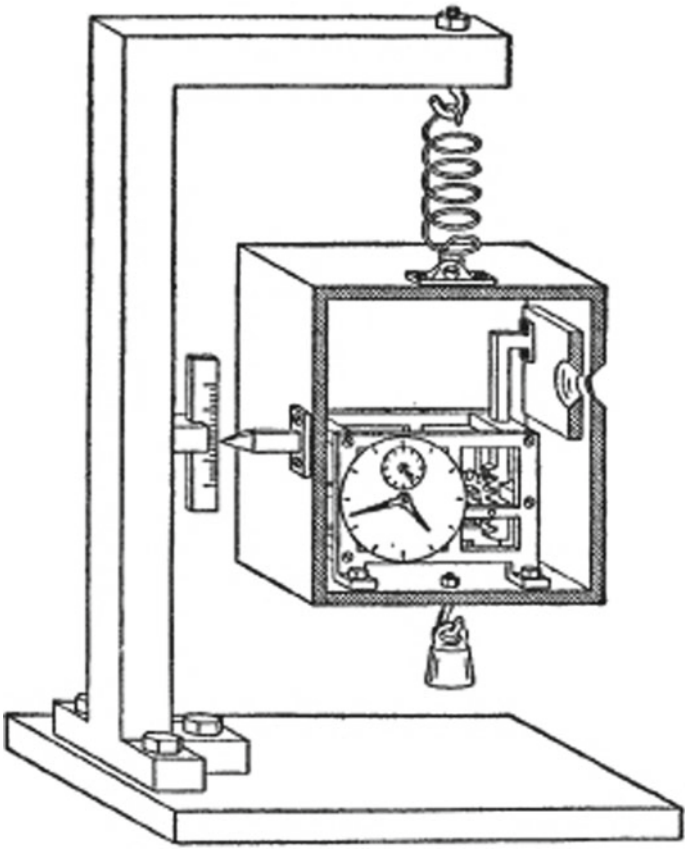


Fig. 7.16 A graphical representation of the apparatus proposed in the Einstein–Bohr debate to test Eq. (1.292). Adapted from Bohr (1949, p. 227)

tion of the position of the balance's pointer is given with an accuracy Δx, which will involve an uncertainty Δp_x in the control of box momentum according to Eq. (1.287). The target of the discussion was, in fact, the energy–time uncertainty relation and not the position–momentum uncertainty, which to many physicists appeared at that time less problematic than the former one. The uncertainty Δp_x must be smaller than the total momentum which, during the whole interval δt of the balancing procedure, can be imparted by the gravitational field to a body with mass Δm, i.e.

$$\Delta p_x < \delta t \cdot g \cdot \Delta m \ , \tag{7.173}$$

where g is the gravitational force. The greater the accuracy of the reading x of the pointer, the longer must the balancing interval δt be if a given accuracy Δm of the weight is to be obtained. But according to GR, the gravitational redshift will affect

the speed of the clock, yielding uncertainty Δt in the time δt required for the pointer to return to its original position. This uncertainty is given by

$$\Delta t = \frac{1}{c^2}\delta t g \Delta x \ . \tag{7.174}$$

By substituting the value of δt given by Eq. (7.174) into Eq. (7.173) we obtain

$$\Delta p_x < \frac{c^2 \Delta t \Delta m}{\Delta x} \ . \tag{7.175}$$

Finally, by applying Eq. (1.287) again, assuming that we have about the minimum uncertainty, we obtain

$$\Delta t > \frac{\hbar}{c^2 \Delta m} \ . \tag{7.176}$$

This, together with Einstein's formula (1.5), from which we get $\Delta E = \Delta m c^2$, gives Eq. (1.292) apart from a constant. It is worth mentioning that, after this discussion, Einstein became convinced that we cannot perform simultaneous measurements on the same quantum system, which can explain the reason why he introduced the EPR model with two particles later on.

Conclusions

Not all problems are solved. A difficulty is the explanation of why certain parameters or constants have exactly the value that they have.[70] In particular it is important here the dimensionless gravitational coupling constant (see Eq. (7.147))

$$\alpha_G = \frac{G m_p^2}{\hbar c} = \left(\frac{m_p}{m_{\mathrm{P}}}\right)^2 \approx 5.91 \times 10^{-39}, \tag{7.177}$$

characterising the gravitational attraction between a given pair of elementary particles, where G is Newton's gravitational constant, m_p is the proton's mass and m_{P} is the Planck mass (3.42c).

In general terms, the previous examination shows that, except gravitation, we can account for all basic physical processes and parameters in terms of fields with their relative gauge bosons as carriers. It is true that there is current research on quantum gravity and that sometimes a graviton is postulated to be the vehicle of the quantum gravitational field.[71] This research aims often at assimilating gravitation to a kind of quantum-field. In such a way, the noted discrepancies between QM and GR could be solved. Clearly, the price would be the renunciation to Einstein's construction of gravity in geometrical terms. Similar considerations are true for string theory, which

[70] Tegmark et al. (2006).

[71] Kiefer (2007), Rovelli (2004). See also Penrose (2004).

appears up to now to be far away from testability.[72] I also mention that the Pavia team is trying to infer from the general principles of quantum information (Sect. 6.4) the basic equations of quantum-field theory.[73] In the following, we shall explore ways to theoretically integrate QM and GR that appears to be economical and simultaneously faithful to Einstein's original insight about gravity.

In the following, we shall explore two different approaches to this kind of problems. The first one, loop quantum gravity, is centred on the Planck scale, while the second one, entropic gravity, deals more with the mesoscopic and macroscopic effects of gravity.

7.2 Space-time, Gravity, Signals

7.2.1 Quantum and Space-time

Physical Interactions

Let us recapitulate what we have learnt about the physical notion of *interaction*. The notion of interaction was well defined in CM: it was the local exchange of dynamical quantities (through collision, pulling, pushing, friction) between bodies being in direct contact. In particular, always at least a quantum of interaction is involved in such processes (Sect. 6.2.5). Such a notion acquires some ambiguities in QM since the previous examination shows that it covers three different aspects:

- *Correlations*. It has been noted that the premeasurement step is formally described by an interaction Hamiltonian (Sect. 3.1). On the other hand, the experiments of Mandel et al. and Scully et al. (Sect. 5.2.3) clearly show that this can happen in the absence of any exchange of momentum or energy, thus of any kind of physical interaction in the proper (or, at least, classical) sense of the word: it is sufficient the simple *presence* of an object or a device to influence the behaviour of the quantum system. We have even excluded any kind of mechanical process when information is involved (Sect. 6.3.4). Here, it is again manifest a fundamental aspect of quantum systems: their sensitivity to the environmental details and even to its geometry (for instance, the arrangement of an interferometer). As mentioned in Sect. 5.2.6, since in this way, the state of the system will be changed, this means that quantum systems map those environmental details in their own informational state (it does not matter here whether this change is reversible). Thus, it seems natural to infer that quantum systems integrate in such a state not only the correlations that they have with virtually all the other systems of the universe but also the correlations that they have with *local* environmental details.

[72] For an introduction to string theory see Zwiebach (2004).

[73] D'Ariano and Perinotti (2014).

- *Fields*. In the case of fields, we clearly have physical interactions in the proper sense (when gauge bosons are exchanged). On the other hand, correlations play also a crucial role contributing to the structure of the field. Those interactions, as far as the field and its dynamical evolution is concerned, do not matter in their 'collapse' quality in the same sense as ordinary detection events. Rather, these interactions are involved in the *propagation* of the field itself and are treated from this perspective. In fact, photons (and their interactions) can appear less fundamental than the electric and magnetic fields themselves (which are observables). This is true. Nevertheless, the occupation number (Sect. 7.1.2) is fundamental for describing the field, and so photons, although not individually distinguishable (Sect. 1.3.1). Here, the focus is no longer on correlations but on observables, i.e. the kind of parameters that describe the field.
- *Events*. Also here, we have an exchange of at a least a quantum of interaction, but the stress is on the kind of event and the signal it sends. For instance, in the experiment shown in Fig. 5.15, Sect. 5.2, a photon emitted during the passage through the microcavities is to us a sign of which–path information. Thus, according to what said in Sect. 5.2.6, events need to be understood as the cross of two dynamical paths (of the two interacting systems, in the simplest case), which, in principle (according to quantum laws), could be read also in reversed time order (Sect. 7.1.4). However, as said, quantum systems are sensitive to the different details of their local environment, and this determines a kind of irreversibility when their dynamical paths cross. What happens is that they superpose two different 'memories', which although transient (and reversible) before interaction are no longer so when the systems interact. Now, the perspective that I find correct to adopt is that it is the *discrepancy* in the informational maps of the *local environments* of two systems, as coded in their respective states, that generates the collapse (while their maps of the whole universe should plus or minus agree when they are close to each other), allowing us to understand the jump to a specific component of some initial state as a point of agreement between the two local maps of the involved systems, that is, as a result that is *congruent with both local maps*. If we speak in spatial terms for simplicity, the point of agreement (and of collapse) is exactly where the two systems cross. Clearly, these considerations are also true for fields as far as we consider local interactions.

Resuming, while in the first kind of interactions the focus is on the involved correlations, in the case of fields it is on the observables that describe the field. Finally, when we deal with events, the issue is which kind of events have happened and which kind of signal they have sent or can send.

Events Outside of the Light Past Cone Having Influence

In Sect. 7.1.2, we have seen that Lorentz and unitary transformations are not the same. The possible solution to this problem has already been anticipated in Sect. 6.2: as far as quantum systems like particles or photons do not interact they evolve unitarily,

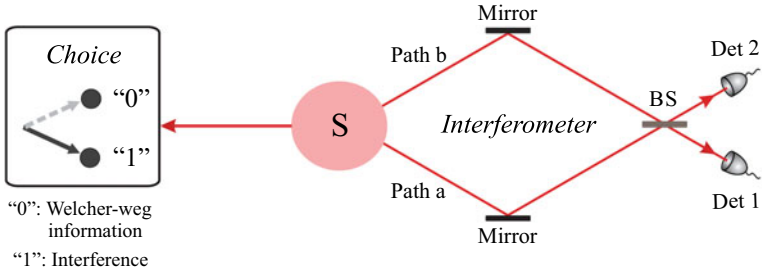


Fig. 7.17 The concept of Ma et al. experiment. The source S emits path polarisation entangled photon pairs (7.179). The system photon goes through the interferometer on the right while the environment photon undergoes polarisation measurement on the left. Adapted from Ma et al. (2013)

but when causal connections, even *potential*, with other quantum systems are considered, what is obviously the case with fields, these processes are ruled by Lorentz transformations. This can be shown on the basis of an already quoted experiment (Fig. 7.17).[74] This is a mix of delayed–choice and which–path experiments. First, we prepare a two-photon polarisation entangled state

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\left(|h\rangle_S\,|v\rangle_E + |v\rangle_S\,|h\rangle_E\right), \tag{7.178}$$

where as usual, $|h\rangle$, $|v\rangle$ stand for horizontal and vertical polarisation state, respectively, and the subscript S, E stand for the system and environment photon, respectively. Then, the orthogonal polarisation states of the system photon are coherently converted into two different interferometer path states a and b via a polarising beam splitter and two fibre polarisation controllers. In general, a polarisation beam splitter transmits photon in horizontal polarisation and reflects those in vertical polarisation (what gives two degree of freedom for each component). This approximately generates the hybrid entangled state

$$|\Psi_H\rangle = \frac{1}{\sqrt{2}}\left(|b\rangle_S\,|v\rangle_E + |a\rangle_S\,|h\rangle_E\right). \tag{7.179}$$

In such a configuration, the environment photon carries which–path (*Welcher–Weg*) information about the system photon. We can now perform two complementary polarisation projection measurements on the environment photon and acquire or erase which–path information of the system photon. This can be done by choosing between two options: either (i) we project the environment photon into the linear polarisation basis $\{|h\rangle_E, |v\rangle_E\}$, which reveals the path and displays no interference, or (ii) we project the environment photon into the $\{|r\rangle_E, |l\rangle_E\}$ basis, with

[74] Ma et al. (2013).

$$|r\rangle_E = \frac{1}{\sqrt{2}}\left(|h\rangle + i\,|v\rangle\right)_E \text{ and } |l\rangle_E = \frac{1}{\sqrt{2}}\left(|h\rangle - i\,|v\rangle\right)_E \qquad (7.180)$$

representing right and left circular polarisation states (Eq. (1.104)), respectively, in which case we erase which–path information. In the latter case, we describe the two systems as being in the state

$$|\Psi_H\rangle = \frac{1}{2}\left[\left(|a\rangle_S + i\,|b\rangle_S\right)|l\rangle_E + \left(|a\rangle_S - i\,|b\rangle_S\right)|r\rangle_E\right]. \qquad (7.181)$$

This choice is performed by a quantum random number generator (QRNG). To rule out any classical causal influence between (i) the choice of the polarisation measurement basis of the environment photon and its projection on one of the two basis and (ii) interferometer-related events of the system photon (from its entry into, up to its exit from, the apparatus), we set up the respective experimental apparatus in three distant laboratories (for polarisation beam splitting, polarisation projection, and QRNG) (Fig. 7.18). The conclusion of the experimenters is that no signal can have

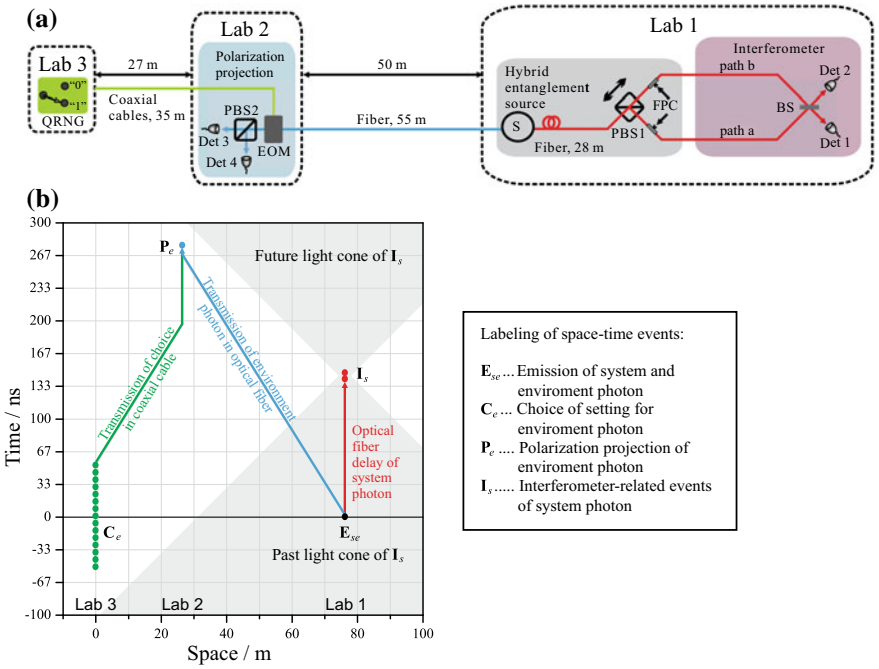


Fig. 7.18 Panel **a** the three labs with their fundamental operations are shown. The polarisation filter PBS1 allows the path distribution of the system photon's polarisation, while PBS2, controlled by the quantum random number generator (QRNG). The projection of the environment photon into a polarisation basis is effected through PBS2 jointly with an electro-optic modulator (EOM). Panel **b** The time–causal connections between events (emission of the photon pair, interferometer-related events, choice of polarisation and projection of the environment photon into one basis). Adapted from Ma et al. (2013)

been sent by one photon to the other. The experiment demonstrates and confirms that whether the correlations between two entangled photons reveal which–path information or an interference pattern of one (system) photon depends on the measurement choice on the other (environment) photon, even when all of the events on the two sides are space-like separated.

In other words, (i) a signal sent by the QRNG in lab 3 to the lab 2 (and here received), together with (ii) the propagation of the environment photon into the apparatus of the latter lab, and with (iii) the propagation of the system photon into the interferometer (lab 1), can determine the kind of detection event in lab 1, although the event determining the polarisation basis of the environment photon happens outside of the past light cone of that detection. As previously stated (Sect. 5.1.5), correlations are totally independent of space-time. This means that, as far as we consider the interdependencies *intrinsic* to a (even compound) quantum system (i.e. the quantum features), the system behaves in pure quantum-mechanical way and propagates according to the typical multipath dynamics (i.e. according to unitary evolution). When there is a source event, then it also gives rise to a signal propagating according to relativity. Here, the photon brings with itself the memory of the source event. Nevertheless, after the source event and before the final detection that signal can undergo further physical processes (Sect. 6.2.5). Thus, we can alter such multipath dynamics in ways that can be again pure quantum mechanical. It is the mechanical model of signal exchanging characterising CM that connects the dots and (at both small and large scale) sees a classical trajectory where there is none, opening in this way the door to a potential conflict between QM and SR (see also Sect. 5.1.1). Thus, we need to agree about two main conclusions: (i) the quantum dynamics between the initial and final event need to be ruled by quantum propagators (Sect. 6.2.4), but (ii) the initial and final events themselves need to be related by a Lorentz transformation.

Causality, Information, Signal

Therefore, quantum theory is essentially about how information is network-like processed being kept conserved (Sect. 6.1) while SR is a theory of how signals are spread in our universe. This means that SR is essentially an explanation of causal processes in our world and it is the merit of Einstein to have lead physics to such a causal structure. However, what is really extraordinary is that also quantum theory, as far as local events are considered, is a causal theory as well (Sect. 6.4.1). And this is necessary, otherwise the initial and final events in a communication protocol could not be related by a Lorentz transformation. In other words, SR is concerned with the propagation *rate* of the signal (and therefore with the causal connection between initial and final events), while QM, validating such a causal structure) is concerned with the (multipath) *modality* of such a propagation.

The only difference is that the notion of causality in quantum theory has been developed on the basis of controlled process while SR deals also (and mainly) with uncontrolled processes happening in our universe: in fact, together with GR, is widely used in cosmology. Of course, also QM is employed at this scale but, for the men-

tioned reason, its epistemological status is never fully clear. However, I shall make an effort here for overcoming this dichotomy, at least to a certain extent. I have already stressed that a relational view that is located halfway between quantum 'subjectivity' and classical 'objectivity' can be helpful (Sects. 3.2.3 and 5.2.6). Thus, the principle of causality can cover both fields of investigation. D'Ariano and co-workers say that the "no–signaling from the future" is the only requisite for an operational probabilistic theory "in order to satisfy the principle of no–signalling without interaction".[75] Let us reformulate this in general terms in order to disentangle the issue of causality from the operations that we perform in controlled situations:

> A signal spreading from a source event cannot be affected by any of its causal effects.

The words "causal effects" seem to display a certain circularity. However, they stand there only for making sure about the kind of "effects" that we are speaking here about, given a certain ambiguity of the term. We could also speak of "final" or "detection" event, but these formulations would be less general. Note that in such a formulation, the causality principle is neutral relative to the direction of the time arrow (and therefore relative to particle/antiparticle distinction). In fact, the occurring of creation and annihilation events of couples particles–antiparticles, although virtual in ordinary physical conditions, shows that events are primary to the spatial–temporal relations and determine that order locally. This is clearly displayed by Feynman diagrams (Sect. 7.1.4). Thus, we can conceive both processes of pair annihilation or creation as generated from a single quantum event but responding to the two different time directions.[76] For reasons that are not completely clear (perhaps due to initial and small fluctuations in the population of particles/antiparticles), such symmetry can be broken giving rise to universes (like our own) in which matter predominates and other universes in which antimatter predominates. Thus, what we assume is that whatever is the direction, signals spread in that way and consequently determine the causal structure of our universe. It has been in fact noted (Sect. 6.2.5) that no signal can go back to its source. At most, we can send another or the same, deflected, signal 'towards' the source event. But this is not a backpropagation.

---

[75] D'ariano et al. (2017, Chap. 6). Moreover, the authors stress the intimate relation of their principle with Einstein's notion of causality.

[76] Also the transactional interpretation of QM stresses the relevance of backpropagated (so-called *advanced*) waves and according to the view that signals cannot implode, correctly interprets such a wave as *expanding* in the other time direction: see Cramer (1986), Kastner (2013, Sect. 3.1). Such an interpretation is rooted in a proposal of Wheeler and Feynman to consider electromagnetic radiation as composed of a retarded wave emitted by a source and an advanced wave coming out from the target (absorber): see Wheeler and Feynman (1945, 1949). In the original proposal the advanced waves cancel out so that only retarded waves survive. Subsequently, it was shown that such a model does not account for a kind of self-interaction of the charged particle that is manifest in the Lamb shift (a difference in energy between two energy levels of the hydrogen atom), after the name of the American physicist Willis Lamb (1913–2008). Thus, the proponents of the transactional interpretation introduced the idea that a quantum event is produced when the retarded and advanced waves meet producing a kind of self-interaction (the transaction). The main difference with the standpoint supported in the present book seems to be in the fact that all advanced waves pertain to the domain of antiparticles and do not concern the way in which events are produced in our world.

Emerging Space-time

Thus, quantum events, giving rise to local irreversible processes (and to a global time arrow as the result of many interactions), being the source of signals, allow for the causal processes that we observe. The fundamental point is that there is no dynamics within space-time.[77] Now, if we cannot speak of space-time as far as pure quantum-mechanical aspects are considered but at the same time any quantum event is a source of signals, so that light spreading from quantum source events determines the general framework in which causal processes happen, then we are allowed to understand this as the *generation of the space-time manifold* itself through the networks established by those propagation processes as far as interactions are involved (Sect. 5.1.5). So, space-time was a consequence of quantum events spreading out light signals and the 'points' in space-time be relativistic events embedding quantum source (and reception) events. I mention that this should be partly in agreement with the point of view of the American science and religion scholar Ian Barbour (1923–2013),[78] if I am not wrong. It is the fact that events are sources of causal connections and not causes themselves (Sect. 6.2.5) that make of them generators of space-time and of its causal array. Note, again, that each quantum event replicates the characters of the presumable original event of our universe: the Big Bang as a generator of space-time.

This point of view finds some support in the work of the Israeli physicists Avshalom Elitzur and Shahar Dolev[79]:

> Suppose that there is […] a "now" front, on the one side of which there are past events, adding up as the "now" progresses, while on its other side there are no events, and hence, according to Mach, not even space-time. Space-time thus "grows" into the future as history unfolds. Time's asymmetry would therefore be naturally anchored in this alleged progress of the "now". […] The "now" does not move on some pre-existing dimension but rather creates that dimension. This is not "movement" in the ordinary sense, so no endless series of time parameters is entailed by it. Now let this Becoming be made quantum mechanical. What role does the wave function play in this creation of new events? The dynamically evolving space-time allows a radical possibility. Rather than conceiving of some empty space-time within which the wave function evolves, the reverse may be the case: The wave function evolves beyond the "now", i.e. outside of space-time, and its "collapse" due to the interaction with other wave functions creates not only the events, but also the space-time within which they are located in relation to one another. The famous peculiarities of the quantum interaction—nonlocality, the coexistence of mutually exclusive states, backward causation and the inconsistent histories […], thus become more natural.

In other words, the photon exchanges determine the abstract (flat) space-time network. Such a network could be taken to be simply ideal. In reality, it represents the pathways through which signals are spread and therefore causal connections are

[77] Geroch (1978, p. 20).

[78] Quoted in Brown (2005, p. 14).

[79] Elitzur and Dolev (2005, p. 346).

established. It even coincides with such signal spreading. The time arrow is thus not a matter of convention and neither it is the abstract topology of space-time.[80]

Then, such a network is necessarily local and not universal, since it depends on the local propagation of signals out of quantum events, notwithstanding the fact that, at least for the visible portion of our universe, all microphysical systems have had all the time since the Big Bang to exchange photons and so to become space-time connected. Thus, the fact that in general relativity we have a local flat Lorentzian space-time with Minkowskian geometry reflects the way in which physics works in reality.

The advantage of this proposal is that it integrates both the diameter about the Planck quantum-mechanical theory of fluctuations at the Planck scale. The hypersphere with a diameter about the Planck length (3.42b) surrounding a quantum event shows in fact no space-time and also no continuity. It is like a shielded singularity. It is only after the first effects of the signal on the surroundings that we have the emergence of a continuous space-time, in agreement with both the distinction between event and observed event (Sect. 5.1.1) and the information acquisition principle (Sect. 6.2.5). This could also explain why when we deal with fields and other continuous quantum phenomena we cannot have perfect localisation and only deal with probability densities (Sects. 1.2.3 and 7.1.3).

Stationary States and Events

We have seen (in Sect. 3.4.2) that the Dutch physicist G. 'T Hooft assumes that the whole universe as well as other quantum systems be in a ontological state.[81] Moreover, he considers the ontological states as stationary states and thus tells us that stationary states evolve into stationary states. Since a stationary state does not spontaneously evolve into an orthogonal energy eigenstate, likely the author thinks about the degeneracy of energy eigenvalues (Sect. 1.2.5): the more degrees of freedom a quantum system possesses (or the more it is compound), the more it has degeneracy of states. Already an electron confined in a hydrogen atom possesses momentum, orbital angular momentum and spin. On one extreme we have a 1D particle confined in a box or potential well which displays no degeneracy, on the other the whole universe which likely shows an infinity of such states. Thus, a quantum system that is sufficiently 'complex' can evolve from a stationary state to another that is degenerate relative to it.

Moreover, for elementary energy conservation reasons, the whole universe should be in a stationary state. Since it has evolved from an initial event (the Big Bang), such

[80] Malament's theorem, after the name of the American philosopher of science David Malament, tells us that continuous timelike curves determine the topology of space-time: see Malament (1977). This has several interesting consequences. One consequence is that if one knows of all points p and q whether it is possible that a particle travel from p to q, then one can recover the topology of space-time, and for this to be possible, it suffices the past and future distinguishability. Clearly, this has to do with the causal structure allowing that topology.

[81] 't Hooft (2016, p. 74–81).

an event must have produced an expanding quantum wave in a stationary state. I have already stressed the self-similarity of each quantum event; thus, although we cannot assume that all quantum systems are in a stationary state, all quantum events should generate or determine systems in stationary state (and they remain so if they do not interact with other systems and no dissipative process occurs). The paradigmatic case is the emission of a photon. This is evident for the emission spectrum: any chemical element releases a photon when jumping from a high to a lower energy state, and such a photon possesses a specific frequency (which allows us to determine the spectral composition of stars), but a specific frequency means a specific energy and thus an energy eigenstate. This confirms the strict connection between information and energy when signals are involved (Sect. 6.2.5). Moreover, as we shall see this seems to be well in agreement with the loop quantum gravity for which the only complete observables are those that commute with the Hamiltonian.

In other words, although we cannot tell which *particular* stationary state is that of a spontaneous emission if the signal is not received, we can say that it must be a stationary state. In other words, we have here a second-level equivalence class that considers all stationary states on the same foot. Since the signal brings with itself information that is in principle acquirable (otherwise no event could have occurred), this is in agreement with the fact that information is a second-level equivalence class (Sect. 6.3.4). However, we cannot represent or describe these unknown systems in the same way in which we describe controlled quantum systems. In other words, we cannot apply to them the notions of state (understood here as a first-level equivalence class), observable and property.

Let us now come back to the issue of the whole universe. The model of a universe evolving from one stationary state to another has been developed by the British physicist Jonathan J. Halliwell. The main idea of Halliwell is that the universe (at least at its global level) can be represented by a pure state of the general form $|\, p, E, n\rangle$ that (considering also relativistic aspects) is an eigenstate of momentum, energy and number operators.[82] Then, we could have an evolution of the kind

$$|\, p_1, E_1, n_1; t_1\rangle \longmapsto |\, p_2, E_2, n_2; t_2\rangle \longmapsto \ldots, \tag{7.182}$$

where the subscript depends on time. Note that, the number of systems (which, at least for relatively early stages of our universe, coincides with the number of photons), as Eq. (7.182) shows, is somehow fixed for our universe and clearly depends on the initial conditions (initial amount of energy) at the moment of Big Bang. This evolution would be perfectly unitary (keeping 0 entropy) but at the same time dependent on the whole state of correlations and interactions established and happening at various levels. This evolution should be a quantum trajectory (Sect. 6.2.4): the whole dynamics is a kind of itinerant self-decoherence: each time the universe is in one of the eigenstates (7.182) but opens again to innumerable possibilities out of which another particular degenerate eigenstate should be selected.

[82] Halliwell (2010).

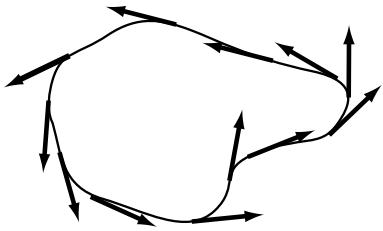


Fig. 7.19 A loop α and the distributional electric configuration represented by arrows. Adapted from Rovelli (2004, p. 11)

7.2.2 Loop Quantum Gravity and Entropic Gravity

Loop Quantum Gravity (LQG)

We shall now consider some proposals that stress the granular structure of our universe at a very small scale. As we have noted in the previous section, energy determines gravitational effects. Thus, we expect that also at the smallest scale (i.e. the Planck scale) there be some gravity: quantum gravity. This is the object of LQG. The basic ideas of LQG can be summarised as follows[83]: (i) it takes general relativity as it is (i.e. with its geometrisation of gravitation), (ii) it is a theory with background (space-time) independence, (iii) it renounces to any kind of unification of the four basic forces and (iv) it vindicates a 4D space-time without need of supersymmetries and strings. About Point (ii), LQG is grounded on a non-canonical algebra based on the holonomies of the gravitational connections (Sect. 7.1.5). Then, the natural consequence is to introduce loop states. This idea is quite general and the first hint is due to M. Faraday, who, having found Eq. (7.14), hypothesised that the electric field be zero everywhere except along a loop, and that at every point of the curve is tangent to the loop (Fig. 7.19), where it should be noted that such a formalism fulfils Coulomb law (7.12). Since the position of a loop state is only relative to other loops, no background space-time is necessary here (recall that we are at the Planck scale).

This leads to the natural consequence of the *granularity of space* (and time), what fits well with the quantisation of space-time discussed in Sect. 7.2.1. A quantum state of space $|s\rangle$ is formed by N grains of space, some of which are adjacent to each other. This can be represented by an abstract graph Γ with N nodes being the grains of space. The links of the graph link adjacent grains and represent the surfaces separating two adjacent grains. Imposing the labels n_i for the nodes and j_k for the links, such a graph is called a *spin network* (Fig. 7.20, Left panel).[84] Note that only abstract combinatorial relations defining the graph are relevant, nor its position, neither its shape. As a consequence, space itself (or space-time) has a combinatorial character, as expected from QM. Thus, the Hamiltonian, as the dynamical quantity ruling time translations, acting on the nodes, changes the combinatorial topology of the graph (Fig. 7.20, Right panel).

[83] Rovelli (2004, Chap. 1).

[84] The first idea is due to R. Penrose: see Penrose (1971).

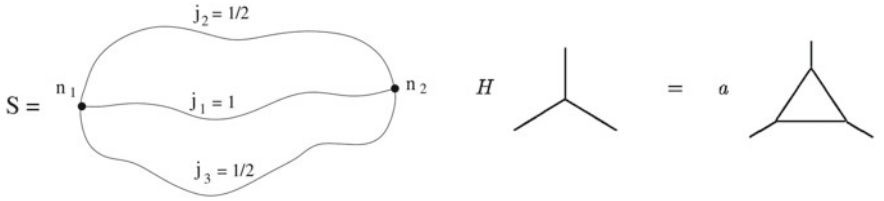


Fig. 7.20 Left panel: example of spin network. S is the surface with the links connecting the two nodes displayed. Right panel: example of the Hamiltonian's action on a node. Adapted from Rovelli (2004, pp. 13 and 17)

Since we are at Planck scale, the quantum dynamics of a particle can be described by probability amplitudes (of the kind of the Green's functions (6.138))

$$\mathrm{G}(x', t'; x, t) = \left\langle x' \left| e^{-\frac{i}{\hbar}\hat{H}_0(t'-t)} \right| x \right\rangle = \left\langle x', t' \mid x, t \right\rangle, \tag{7.183}$$

where $|x, t\rangle$ is an eigenstate with value x of the Heisenberg position observable $\hat{x}(t)$ and $|x\rangle = |x, 0\rangle$ for $t_0 = 0$. Applying this idea to fields and following Feynman's path integral formalism (Sect. 6.2.4), these amplitudes can be understood as sums over different field configurations that take value φ on surfaces Σ. Now, these fields include gravitation and this provides for the space-time geometry and therefore determines the relations among systems. These propagators take the form

$$\mathrm{G}(s'; s) = \langle s' | \hat{P} | s \rangle = \left\langle s' \mid s \right\rangle_{\mathcal{H}}, \tag{7.184}$$

where $\hat{P}$ is an improper projection operator to be determined, defining the scalar product between states of space on the relative Hilbert space $\mathcal{H}$. Here, the transition amplitudes display diffeomorphism invariance, what implies background independence. The term *diffeomorphism* denotes a mathematical notion, and in particular, it is an invertible function that maps one differential manifold to another (Sect. 2.3), and both the function and its inverse are smooth (having derivatives of all orders). An active diffeomorphism only drags a field and it is therefore gauge, i.e. physically irrelevant. What is preserved under such transformations are the coincidences between the values the gravitational field take at such and such a 'place' and the values the matter fields take there. If $|\emptyset\rangle$ represents the state with no spin network, we have

$$\mathrm{G}(s) = \left\langle \emptyset \left| \hat{P} \right| s \right\rangle = \langle 0 \mid s \rangle \text{ or } \hat{P} \, |\emptyset\rangle = |0\rangle \, , \tag{7.185}$$

where in such a context $|0\rangle$ is the covariant vacuum (which is not the state of minimal energy).

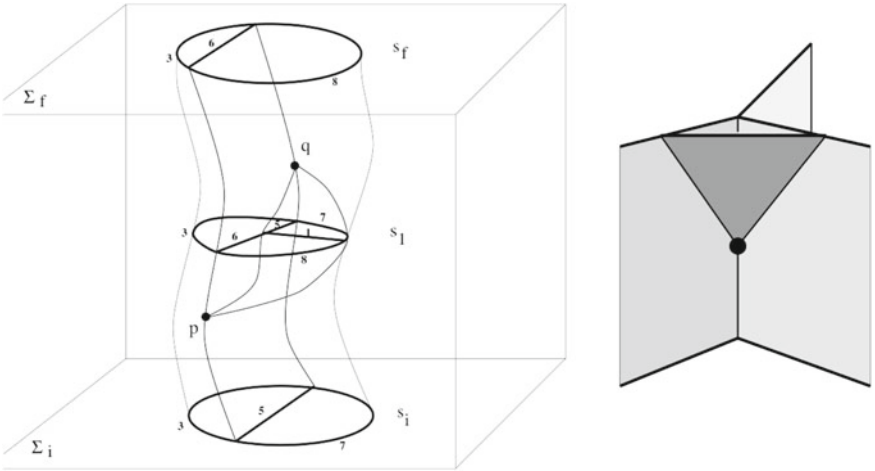


Fig. 7.21 Left panel: example of spin foam. Right panel: the vertex of a spin foam. Adapted from Rovelli (2004, pp. 18 and 19)

As said, probability amplitudes sum over paths. A path is the world history of a graph with interactions (events) happening at the nodes, while the nodes themselves are quanta of volume (loops). Thus, this world history is a collection of faces, namely the world histories of the links, and the faces join at edges that are the world histories of nodes; finally, edges join at vertices. Thus, while a Feynman graph is composed of edges and vertices (see Sect. 7.1.4), the two-complex world histories are composed of edges, vertices and faces. This is called *spin foam* (Fig. 7.21, Left panel). An example of vertex displaying the action of the Hamiltonian is shown in Fig. 7.21, Right panel. A spin foam is given by a measure term $\mu(\sigma)$ times the product over the vertices v of a vertex amplitude $\vartheta_v(\sigma)$. The physical probability amplitudes $\mathrm{G}(s'; s)$ are then obtained by summing over the spin foams bounded by the spin networks s and s':

$$\mathrm{G}(s'; s) \approx \sum_{\sigma, \partial\sigma = s \cap s'} \mu(\sigma) \prod_v \vartheta_v(\sigma). \tag{7.186}$$

This probability amplitude can be interpreted as a sum over space-times. It can be seen as a precise realisation of quantum gravity formalism as a sum over four geometries.

In this perspective, what remains are relations among dynamical entities[85]: two particles' worldlines intersect, a field has a certain value where another field has another value, or we deal with partial observables that can be measured together. In sum, physical objects can only be localised with respect to one another, in agreement with the relational point of view supported here (Sects. 5.2.6 and 7.2.1). Thus, physics is a correlation among partial observables without any assumption of a space-time background, as announced.

[85] Rovelli (2004, Chaps. 2–3).

The metric structure inherited by σ depends on the gravitational electric field E.[86] Considering in particular a 2D surface S embedded in a 3D surface σ, its area is given by

$$A(S) = \int_S d^2\sigma |E| = \int_S |E|, \tag{7.187}$$

where in the last step I have interpreted the expression as the surface integral of the norm of the 2 form $E_j = E^a_j \epsilon_{abc} dx^b \wedge dx^c$. Thus, it could be said that in gravity the "length of the electric field is the area" or the area of a surface is the flux (the norm) of the gravitational electric field across the surface. An analogous expression can be found for the volume.

Then, a state in GR is an equivalence class of 4D field configurations solving Einstein's equations under gauge transformations. The full dynamics of quantum GR is contained in the propagator $\mathrm{G}(\mathbf{A})$ of the field $\mathbf{A}$, which to the first relevant order in $\hbar$, is related to $e^{-\frac{i}{\hbar}S(\mathbf{A})}$ (see e.g. Eq. (6.143)). Note that predictions are here local, that is, concern particular regions of space-time. Moreover, as mentioned, it does not matter where the surface σ is located. Only relative space-time distances do count. Since the geometry is determined by gravitation, we only need to know the value of the dynamical fields on σ.

Let us denote by $\hat{\alpha}(t)$ the Heisenberg (3D) position operator in such a context and write its eigenvalue equation[87]

$$\hat{\alpha}(t)\,|\,\alpha; t\rangle = \alpha\,|\,\alpha; t\rangle\ . \tag{7.188}$$

Putting $|\,\alpha; 0\rangle = |\,\alpha\rangle$ for $t_0 = 0$, we clearly have

$$|\,\alpha; t\rangle = e^{-\frac{i}{\hbar}\hat{H}_\alpha t}\,|\,\alpha\rangle \tag{7.189}$$

where the Hamiltonian

$$\hat{H}_\alpha = \frac{\hbar^2}{2m}\frac{\partial^2}{\partial\alpha^2} + \frac{m\omega^2}{2}\alpha^2 \tag{7.190}$$

is built on the outline of the harmonic oscillator Hamiltonian (6.70). The corresponding propagator is

$$\begin{aligned} \mathrm{G}(\alpha', t'; \alpha, t) &= \langle \alpha'; t' \mid \alpha; t\rangle = \left\langle \alpha' \left| e^{-i(t'-t)\hat{H}_\alpha} \right| \alpha \right\rangle \\ &= \sum_n H_n(\alpha') e^{-i(t'-t)E_n} H_n^\star(\alpha), \end{aligned} \tag{7.191}$$

where $H_n(\alpha) = \langle \alpha; t \mid E_n\rangle$ is the eigenfunction of $\hat{H}_\alpha$ with eigenvalue E_n. Analogously to Eq. (7.184), we also have

[86] Rovelli (2004, Chap. 4).

[87] Rovelli (2004, Chaps. 5 and 7).

$$\mathrm{G}(\alpha', t'; \alpha, t) = \left\langle \alpha' \left| \hat{U} \right| \alpha \right\rangle. \tag{7.192}$$

Denoting any compact support complex function $f(\alpha, t)$, the state

$$\begin{aligned} |\, f\rangle &= \int d\alpha dt\ f(\alpha, t)\, |\, \alpha; t\rangle \\ &= \int d\alpha dt\ \langle \alpha; t \mid f\rangle\, |\, \alpha; t\rangle \end{aligned} \tag{7.193}$$

is the space-time state of the function $f(\alpha, t)$. This notion generalises the wave packet for which we have $f(\alpha, t) = f(\alpha)\delta(t)$ since they can be considered as associated with instantaneous position measurements. The corresponding wave function is

$$\begin{aligned} \psi_f(\alpha', t') &= \langle \alpha'; t' \mid f\rangle \\ &= \langle \alpha'; t' \,| \int d\alpha dt\ f(\alpha, t)\, |\, \alpha; t\rangle \\ &= \int d\alpha dt\ f(\alpha, t) \langle \alpha'; t' \mid \alpha; t\rangle \\ &= \int d\alpha dt\, \mathrm{G}(\alpha', t'; \alpha, t) f(\alpha, t), \end{aligned} \tag{7.194}$$

while the scalar product of two space-times is given by

$$\langle f' \mid f\rangle = \int d\alpha dt d\alpha' dt' f^{\star}(\alpha', t') \mathrm{G}(\alpha', t'; \alpha, t) f(\alpha, t), \tag{7.195}$$

with

$$|\, f'\rangle = \int d\alpha' dt' f(\alpha', t')\, |\, \alpha'; t'\rangle\,. \tag{7.196}$$

We can associate the state $|\, R\rangle$ to any space-time region R and write

$$|\, R\rangle = N_R \int_R d\alpha dt\, |\, \alpha; t\rangle\ , \tag{7.197}$$

where N_R is a normalisation factor. The quantum-gravity dynamics is defined by the so-called Wheeler–DeWitt equation, after the names of J. A. Wheeler and B. DeWitt,[88]

$$\hat{H}\psi(\alpha, t) = 0, \tag{7.198}$$

with the full 4D Hamiltonian

[88] DeWitt (1967).

$$\begin{aligned}\hat{H} &= \hat{H}\left(\alpha, t, -i\hbar\frac{\partial}{\partial\alpha}, -i\hbar\frac{\partial}{\partial t}\right)\\ &= -i\hbar\frac{\partial}{\partial t} + \hat{H}_\alpha,\end{aligned} \tag{7.199}$$

where $\hat{H}_\alpha$ is given by Eq. (7.190) and $\psi(\alpha, t) = \langle\alpha; t \mid R\rangle$. From this we can derive the explicit form for the improper projectors

$$\hat{P} = \int d\tau e^{-i\hat{H}\tau}, \tag{7.200}$$

where τ is a time interval. Now, let us expand $|\, f\rangle$ in terms of the energy eigenstates $|\, E_n\rangle$ as

$$|\, f\rangle = \sum_n |\, E_n\rangle\, \langle E_n \mid f\rangle\,, \tag{7.201}$$

where $\langle E_n \mid f\rangle = f(E_n)$ are the coefficient of such expansion. Then, an arbitrary function $f(\alpha, \tau) = \langle\alpha; \tau \mid f\rangle$, on the outline of Eqs. (1.198)–(1.199) and (7.191), can be written as

$$f(\alpha, \tau) = \sum_n f(E_n) H_n(\alpha) e^{iE_n\tau}. \tag{7.202}$$

Mathematically, if the time interval τ in the expression (7.200) extend to infinity, we know that it gives rise to a Dirac's delta. Suppose, moreover, that we accept uncertainty of time in a finite interval τ. Then, thanks to the energy–time uncertainty relation (1.292), we have $\hat{P} \simeq \delta(E_m)$, where E_m is a particular eigenvalue of the Hamiltonian. In this case, the improper projector acts on the function $f(\alpha, \tau)$ as

$$\begin{aligned}\hat{P} f(\alpha, t) &\simeq e^{-iE_m\tau} \sum_n f(E_n) H_n(\alpha) e^{iE_n\tau}\\ &= \sum_n \delta(E_n - E_m)\, \langle E_n \mid f\rangle\, \langle\alpha; \tau \mid E_n\rangle\\ &= \langle E_m \mid f\rangle\, \langle\alpha; \tau \mid E_m\rangle\\ &= H_m(\alpha) f(E_m),\end{aligned} \tag{7.203}$$

where $f(E_m) = \langle E_m \mid f\rangle$ is an energy eigenfunction and therefore represent a general solutions of Eq. (7.198). In other words, if we accept a certain indeterminacy in time, $\hat{P}$ sends (projects) arbitrary functions in(to) solutions of the Wheeler–DeWitt equation. Note that α and t (τ) are partial observables,, i.e. observables that we can measure but whose value we cannot predict.[89] At the opposite, energy (and any observable that commute with the Hamiltonian) can be considered complete and their value can be predicted. Complete observables are determined through correlations

[89] Rovelli (2002).

among partial observables. This is especially relevant for GR since here observables evolve with respect to each other and we cannot solve for one as a function of the other.

Assume now that (α, t) and (α', t') are two events: what is the probability that we observe the latter having observed the former one? Since we need to renounce the idea of a perfect resolution, we deal in fact with intervals $\Delta\alpha'$ and $\Delta t'$ or also with regions

$$R = (\alpha \pm \Delta\alpha, t \pm \Delta t) \text{ and } R' = (\alpha' \pm \Delta\alpha', t' \pm \Delta t'), \tag{7.204}$$

where

$$\left| R' \right\rangle = N_{R'} \int_{R'} d\alpha' dt' \left| \alpha'; t' \right\rangle . \tag{7.205}$$

If R and R' are much smaller than any other physical quantity involved, included the space-time separation of the two regions, it can be shown that the probability that the two events are observed is

$$\wp_{R,R'} = \gamma^2 \left| \left\langle R' \mid R \right\rangle \right|^2 , \tag{7.206}$$

where γ^2 is a dimensionless constant related to the efficiency of detectors and thanks to Eqs. (7.197) and (7.205), we have

$$\begin{aligned} \left\langle R' \mid R \right\rangle &= \int_{R'} d\alpha' t' \int_R d\alpha dt \mathrm{G}(\alpha', t'; \alpha, t) \\ &= \int_{R'} \int_R d\alpha' t' d\alpha dt \ \left\langle \alpha'; t' \mid \alpha; t \right\rangle , \end{aligned} \tag{7.207}$$

where for simplicity I have assumed $N_R N_{R'} = 1$. In particular, the probability amplitude is given by

$$\vartheta_{R,R'} = \gamma \frac{\left\langle R' \mid R \right\rangle}{\sqrt{\langle R' \mid R' \rangle} \sqrt{\langle R \mid R \rangle}} . \tag{7.208}$$

We easily recover the traditional quantum formulation for probability density. In fact, if R' is sufficiently small, the wave function

$$\psi(\alpha', t') = \left\langle \alpha'; t' \mid R \right\rangle \tag{7.209}$$

can be assumed to be constant in R', and then the probability to be detected in R' is

$$\wp_{R'} = \gamma^2 \left(V_{R'} N_{R'} \right)^2 \left| \psi(\alpha', t') \right|^2 , \tag{7.210}$$

where $V_{R'}$ is the volume of the region R'. If this region has sides $\Delta\alpha'$ and $\Delta t'$, and we assume that

$$\Delta t' \ll \frac{m\Delta(\alpha')^2}{\hbar}, \tag{7.211}$$

then the term $(V_{R'}N_{R'})^2$ can be shown to be proportional to $\Delta\alpha'$ and we can obtain

$$\wp_{R'} \propto \Delta\alpha' \left|\psi(\alpha', t')\right|^2, \tag{7.212}$$

as anticipated. At the opposite, when

$$\Delta t' \gg \frac{m\Delta(\alpha')^2}{\hbar}, \tag{7.213}$$

we have

$$\wp_{R'} \propto (\Delta t')^{-\frac{1}{2}} \left|\psi(\alpha', t')\right|^2, \tag{7.214}$$

and we cannot associate a probability density in t' to this detector, because the detection probability does not scale linearly with $\Delta t'$. In agreement with the amplitude (7.184), we have

$$\mathrm{G}(\alpha', t'; \alpha, t) = \left\langle \alpha'; t' \left| \hat{O} \right| \alpha; t \right\rangle, \tag{7.215}$$

where I recall that α and t (as well as the primed counterparts) are partial observables. We can see the involved (both input and output) states as kinematic states in the sense that they correspond to single events, while their scalar product expresses the relation between the two events.

Entropic Quantum Gravity

I have already mentioned (in Sect. 5.1.5) that the Dutch physicist Erik Verlinde has contributed to the idea of emerging space-time.[90] His proposal has also a fundamental relevance for the issue QM/GR. He makes use of the holographic principle proposed by the Dutch physicist Gerard 't Hooft,[91] according to which the description of a volume of space can be thought of as N bits of binary information encoded on a boundary of that region, namely a closed surface of area A (in the case of space-time, we need to consider a 3D hyper surface). Thus, it is a sort of projection to a lower level dimensional space (or space-time). This is a consequence of the work of the Israeli physicist Jacob Bekenstein (1947–2015) in computing the entropy of a black hole that turns out to be[92]

$$S_{\mathrm{BH}} = k_{\mathrm{B}} \frac{A}{4\ell_{\mathrm{P}}^2} = \frac{A}{4} \frac{k_{\mathrm{B}} c^3}{\hbar G}, \tag{7.216}$$

[90] Verlinde (2011).

[91] 't Hooft (1993). See also Susskind (1995).

[92] Bekenstein (1973). On black holes see Serjeant (2010, Chap. 6).

where A is the area of the event horizon (the boundary enclosing a black hole beyond which events cannot affect an outside observer) and ℓ_{P}^2 is the so-called Planck area, ℓ_{P} being the Planck length (3.45b): practically, we are dividing the area A in small cells of area $4\ell_{\mathrm{P}}^2$ representing units of entropy. In general terms, the information is evenly distributed on the surface and the number of bits is obviously given by (Fig. 7.22; see also Eq. (7.147))

$$N = \frac{A}{4\ell_{\mathrm{P}}^2} = \frac{Ac^3}{4\hbar G}. \tag{7.217}$$

Note the connection among the four constants, characterising QM, SR, GR and statistical mechanics (Sect. 6.2.5) in Eq. (7.216). This is also known as *Bekenstein–Hawking entropy*, after the name of the British physicist Steven Hawking (1942-2018). The reason for the holographic principle is that the entropy of a black hole saturates the *Bekenstein bound* on a given space volume:

$$S \leq \frac{2\pi k_{\mathrm{B}} r E}{\hbar c}, \tag{7.218}$$

where E is the energy contained in that region and r the ray of the sphere enclosing that volume, or, in terms of information,

$$I \leq \frac{2\pi r E}{\hbar c \ln 2}. \tag{7.219}$$

This means that a black hole (its surface) represents the maximal concentration of entropy in a given volume that is possible in our universe.

The effective temperature experienced due to a uniform acceleration a in a vacuum field according to the *Unruh effect* (mentioned in Sect. 7.1.4), after the name of the Canadian physicist William Unruh, is

$$T = \frac{\hbar a}{2\pi c k_{\mathrm{B}}}, \tag{7.220}$$

which allows, in conjunction with the previous equations, to derive Newton's gravitational force through the second law of dynamics (1.37). Thus, gravity would be also an emergent physical dimension out of information entropy. Following Jacob Bekenstein,[93] Verlinde assumes that we can write

$$\Delta S_{\mathrm{B}} = 2\pi k_{\mathrm{B}} \text{ when } \Delta x = \frac{\hbar}{mc} \tag{7.221}$$

for the change of Boltzmann entropy at the boundary, where the expression for Δx has been derived from Eq. (6.180) for the Compton length. We further assume that the change in entropy near the surface is linear in Δx:

[93] Bekenstein (1973).

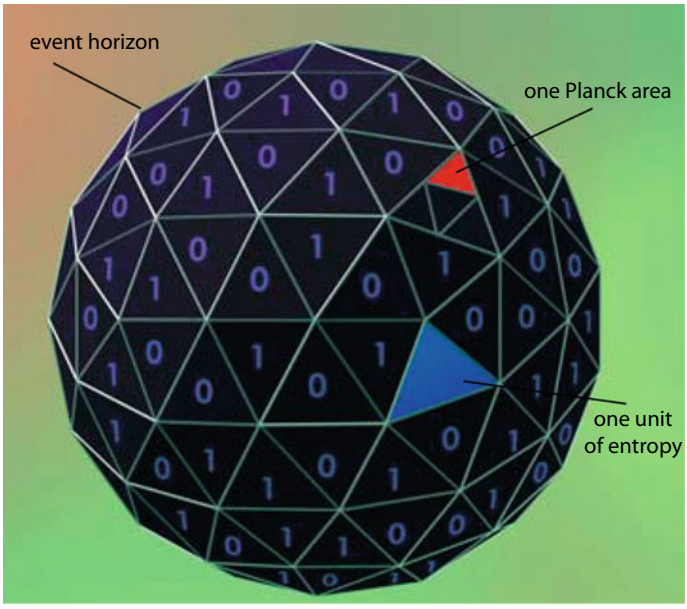


Fig. 7.22 Black hole's entropy: each unit of entropy is four times the Planck area. Adapted from Bekenstein (2003)

$$\Delta S_{\mathrm{B}} = 2\pi k_{\mathrm{B}} \frac{mc}{\hbar} \Delta x. \tag{7.222}$$

Since, in agreement with Eq. (6.182) as well as in agreement with QM, up to a constant each bit carries energy (i.e. $S_{\mathrm{B}} = k_{\mathrm{B}} \ln W = k_{\mathrm{B}}/2$)

$$E = \frac{1}{2} k_{\mathrm{B}} T, \tag{7.223}$$

we can write

$$mc^2 = \frac{1}{2} N k_{\mathrm{B}} T \tag{7.224}$$

for N bits of information. Note that $E = k_{\mathrm{B}} T$ is used as a scale factor for energy in statistical mechanics. When putting this equation into Eq. (7.222) and also use the acceleration a derived from Eq. (7.220), we obtain

$$\frac{\Delta S_{\mathrm{B}}}{N} = k_{\mathrm{B}} a \frac{\Delta x}{2c^2}. \tag{7.225}$$

We conclude that acceleration is related to an entropy gradient. Then, inertia is a consequence of the fact that a particle in rest will stay in rest because there are no entropy gradients. It is then natural to say that acceleration is related to the gradient

of the gravitational (Newton) potential Φ:

$$a = -\nabla\Phi, \tag{7.226}$$

allowing us to rewrite the previous equation as

$$\frac{\Delta S_{\mathrm{B}}}{N} = -k_{\mathrm{B}}\frac{\nabla\Phi}{2c^2}. \tag{7.227}$$

This establishes a connection between entropy gradient and gravity gradient. Thus, we can reformulate the temperature (7.220) as

$$T = -\frac{\hbar\nabla\Phi}{2\pi c k_{\mathrm{B}}}, \tag{7.228}$$

and the energy (7.223) enclosed in a surface S covered by N bits as

$$E = \frac{1}{2}k_{\mathrm{B}}\int_S T\,dN. \tag{7.229}$$

Inserting Eq. (7.228) into (7.229) and making us of Eqs. (1.5) and (7.147), we can derive the classical Gauss law for gravity, after the name of Carl Friedrich Gauss, for the total mass M enclosed by a surface S:

$$M = \frac{1}{4\pi G}\int_S \nabla\Phi\cdot dA, \tag{7.230}$$

where dA is the area of an infinitesimal piece of the surface S. By considering a number of text particles subjected to position shifts of $\delta\mathbf{r}_j$ due to gravitation, we can express the entropic forces on the particles as

$$\sum_j \mathbf{F}_j\cdot\delta\mathbf{r}_j = \frac{1}{4\pi G}\int_S (\delta\Phi\nabla\Phi - \Phi\nabla\delta\Phi)\,dA, \tag{7.231}$$

from which it is possible to derive the generalisation of Gauss law (7.230) to GR:

$$M = \frac{1}{4\pi G}\int_S e^{\phi}\nabla\phi\cdot A, \tag{7.232}$$

where

$$\phi = \frac{1}{2}\log(-\xi^{\mu}\xi_{\mu}) \tag{7.233}$$

is the generalisation of Newton potential to GR, where the ξ^{μ} are the so-called Killing vectors (after the name of the German mathematician Wilhelm Killing (1847–1923)) and e^{ϕ} represents the redshift factor that relates the local time coordinate to that at a

reference point with $\phi = 0$, which will be taken here to be at infinity. Reformulating the Lie bracket or commutator (1.210) in terms of the so-called *Lie derivative*, which acts as a derivative on vector field $\boldsymbol{\eta}$ at a point p depending on vectors $\boldsymbol{\xi}$[94]:

$$\hat{\mathcal{L}}_{\boldsymbol{\xi}}\boldsymbol{\eta}(p) := [\boldsymbol{\xi}, \boldsymbol{\eta}]\,(p) = \nabla_{\boldsymbol{\xi}}\boldsymbol{\eta}(p) - \nabla_{\boldsymbol{\eta}}\boldsymbol{\xi}(p), \tag{7.234}$$

we can define a Killing vector as a vector field $\boldsymbol{\kappa}$ such that annihilates a field g:

$$\hat{\mathcal{L}}_{\boldsymbol{\kappa}} g = 0. \tag{7.235}$$

The advantage of Verlinde's proposal is that it allows for predictions that agree quite well with measured values of gravitation without introducing the ad hoc hypothesis of dark matter. Recently, the distribution of gravity around more than 33,000 galaxies has been measured and the agreement is quite good.[95]

7.2.3 A Dynamical View

As we have seen at the end of Sect. 3.3.4 and in Sect. 5.3, Einstein supported a form of dynamical realism. Being also the father of GR, it seems opportune to ask whether his standpoint can help us here. It is one of the fundamental ideas of relativity theory that space-time structure is not a fixed backdrop against which the processes of physics unfold, but instead participates in that unfolding (Sect. 7.1.5)[96]: it posits a dynamical interaction between the space-time metric in any region and the matter fields there. The interaction is governed by Einstein's field equations (7.172). Let us see how we can apply these concepts to the problem of the foundations of a quantum theory of gravity.

The two formalisms presented above (entropic gravity and LQG) need to agree at the Planck scale, that is, at the level of Planck hypersphere. As a matter of fact, LQG well describes the entropy of a black hole, in agreement with Bekenstein–Hawking formula (7.216), that was the departure point of entropic quantum gravity. Since, according to LQG, the nodes of a spin foam represent interactions (events), it is reasonable to assume that, for elementary events, we have exactly one bit of classical information that is acquirable. When a single quantum event happens representing a bit of information (like the absorption of a photon by the photodetector wall as displayed in Fig. 6.3, Sect. 6.2, where I assume that a photon is thereafter emitted representing a signal of the happened detection), in a very small region of the Planck dimension we expect to have precisely a classical bit of information on the surface enclosing that region and an energy (7.223) within. In other words, this region, in the case of a 3D space, and assuming that its area is $4\ell_{\mathrm{P}}^2$ in agreement with the minimal

[94] Penrose (2004, pp. 314 and 320). See also Malament (2012, p. 74).

[95] Brouwer et al. (2017).

[96] Malament (2012, p. 139).

area of a black hole representing a bit, must have a ray given by

$$r = \frac{\ell_{\mathrm{P}}}{\sqrt{\pi}}. \tag{7.236}$$

Thus Eq. (7.236) could represent the joint between LQG and entropic gravity. This might be insightful because it would show that the projection postulated by the holographic principle describes in fact a physical phenomenon: the information instantiated in a small region (of the Planck 'size') of space (or space-time) is in fact the information that is *available* just at the surface of this region, it is information that is 'pushed' towards the outside of the Planck sphere (or hypersphere). This could explain where is the threshold between an infinite amount of (potential) quantum information and the information (1 bit) classically available out of a single quantum event (Sect. 3.2.3) and the principle of information acquisition (Sect. 6.2.5): in each Planck hypersphere we have that amount, but it is unreachable as such, only its (classical) projection at the surface it is. That projection is a signal telling us "I have happened" (it solves the Bernoulli trial: Sect. 2.1.2), and it is only, for example the set up of the apparatus that tells us that this is a signal for the presence of a particle in a given path.

If we consider the language of classical philosophy, such Planck hypersphere is a kind of Leibnizian monad (Sects. 2.3.3 and 6.2.5). Each event, from the Big Bang up to the emission of a single photon possesses this character. In the words of the Belgian physicist and Catholic priest Georges Lemaître (1894–1966) is a primeval atom.[97]

This seems to be also true for black holes. One of the most important discoveries about these objects is that they emit a radiation called *Hawking radiation*, which shows striking similarities with the emission of a black body. We can assume that a particle–antiparticle radiation is emitted from just beyond the event horizon and that the black hole gravitation could boost virtual particles into real ones so that a particle can escape the event horizon. The temperature of the radiation is

$$T = \frac{\hbar c^3}{8\pi G M k_{\mathrm{B}}}, \tag{7.237}$$

where M is the mass of the black hole. This is again is evidence for a signal emitted by the surface, while it is reasonable to assume that beyond the event horizon there is quantum information. The reason is that no matter structure or whatever physical entity can survive the gravitational squeezing of a black hole, so that 'inside' there is nothing but quantum information, while the units of entropy–information covering the surface are units of *classical* entropy–information. Thus, the repartition of the event horizon in units of entropy is another example of quantisation, each unit of entropy representing possible events giving rise to signals that can be sent and received.

[97] Lemaître (1946).

The use of propagators by LQG confirms that, although already a bit of classical information is available at the surface of the Planck hypersphere embedding a single quantum event, and thus, it can in turn generate one or more events, the propagation of signal outwards can proceed through a multipath (unitary) dynamics, at least in the interval between two events. During such dynamics, the information is propagated in quantum-mechanical form, while bits of classical information is generated through further events, where the first occurring ones need to be at least at the surface of the Planck hypersphere embedding the source event.

Here, gravity adds to this scenario. We need to distinguish between two different issues: one is the origin of mass and this could be explained through the entrenchment of loops according to LQG, although still progresses need to be made in reproducing Einstein's theory of general relativity as a classical limit of LQG; another is the macroscopic effect of mass *once that is there*, what is better described by entropic gravity: here we have pure geometrical issues according to GR. Thus, if light propagation determines the space-time allowing for causal processes (Sect. 7.2.1), then the role of masses and their motion would be to *generate geometric deformations* to that Minkowskian manifold.[98] In other words, gravitation gives the distortions of the signal and not its propagation. Clearly, light, following null geodesics, remains as the paradigmatic signal (pure kinetic energy) that connects the nodes of a network.

In fact, macroscopically gravity manifests itself through three different phenomena:

- The first one is contraction. This can be considered the fundamental one, since gravity becomes sensitive only when considerable mass are at play, and this demands gravitational contraction. Let us consider a paradigmatic situation: a cosmic cloud of hydrogen atoms. As far as we consider small portions of this cloud and the distances among atoms are not too narrow, the effects of gravity are irrelevant. This works as a local quantum-relativistic information network in which the atoms are the nodes and exchange signals. It has clearly a flat geometry. Now, independently of how the first steps of gas condensation can happen (through random collisions among the atoms or due to local maxima of atoms' concentration), overcoming a certain threshold the effects of gravitation become sensitive. This results in growing condensation of the gas and in the formation of molecular hydrogen which contributes in turn to condensation. In other words, the pressure of gravity shrinks the original quantum-relativistic network. Since the 'wires' (representing causal-signalling pathways) connecting the nodes cannot be reduced in size (since in such a case we would reversibly send the signal back to its source, with the consequence of violating Huygens principle: Sects. 6.2.4, 6.2.5, and 7.2.1), they can only bend. Thus, the curvature of space-time was an immediate effect of matter densification and of the consequent contraction. Note that, space-time is not a substance or a substrate but rather a network of relations. As for QM, we see that relations have

[98] The philosopher of science Harry Brown has tried to interpret in dynamical way GR as a theory of the dynamical behaviour of rods and clocks in motion constraining the kinematic relations Brown (2005, p. 82), a standpoint that could be brought in harmony with the point of view supported here.

a fundamental role in the constitution of reality. Of course, these relations are conveyed by the spreading of physical signals.

- Gravity manifested itself as an entropy gradient, according to Eqs. (7.227) and (7.231). Thus, every body preserves its state and is affected by gravity only if there is a gradient of entropy. In such a case, the curvature of space-time generates paths of free fall (timelike geodesics) in which the acceleration compensates the effects of gravity (see also Eq. (7.226)). This allowed us to propose a slight reformulation of the strong equivalence principle taking into account Synge's objection (Sect. 7.1.5): *mass* would be the resistance to any change of state (inertia) apart from the 'natural' effect of gravity. Although mass is a single entity, according to our proposal, it behaves differently when it is subject to gravitation (free fall) and when is subject to other forces. Thus, inertial and gravitational mass could be two aspects or sides of mass, and there was no full equivalence.
  As we have remarked in Sect. 2.3.2, 99% of the energy of matter is represented by mass. In other words, we can understand mass as a giant energy (entropy) store. The generation (and subsequent concentration) of mass can indeed be seen as a transfer from kinetic to potential energy (Sect. 7.1.4). Reciprocally, as remarked again in Sect. 7.1.5, a body falling in a gravitational 'field' transforms back gravitational potential energy into its kinetic energy. This makes a difference relative to light that remains pure signal (and pure kinetic energy).
- Finally, there is also a further effect of gravity on bodies that appears less natural: the tidal effect. This is generated as a compression on (and stretching of) bodies as extended physical systems. If the mass of a body were concentrated in a point, it would only experience free fall (that is therefore the natural behaviour of mass, as said), but falling along a geodesics naturally involves deformation effects on an extended body (stretching along the gedesics and compression along the orthogonal direction): this is due to the difference in the gravitational 'force' (or slope) between the front and rear part of the body relative to the geodesic line. Ultimately, this is due to the fact that the macroscopic conformation of a body is not simply mass but is a geometric configuration that involves chemical bonds. This is an issue that will be discussed in the next section.

In conclusion, many of the attempts at unifying QM with GR either try to assimilate the latter to the former or vice versa. At the opposite, in agreement with LQG, I propose that these three theories, QM, SR and GR, need to be taken as they are, and the kind of integration that we need could be accomplished if we assume that each of them rules a particular and fundamental aspect of our physical world. In other words, at the smallest scale (the Planck scale), we have no signal, no space-time and at most, we may have quantum events as cores of the space-time 'wave' that arises overcoming the Planck diameter around the event. Thus, at small scales, we have quantum nodes and pure quantum information processing. This, as shown by the delayed choice experiment (Sects. 2.4.3), remains totally undisturbed by the propagation of signals (ruled by SR) that concerns relations that are locally seen from a cosmological point of view. At a bigger scale, distortions due to masses and their motion are ruled by GR. Thus, QM, ruled by unitary evolution, SR ruled by the Poincaré group of

transformations, and GR with its Lorentz group of transformations describe three different kinds of physical phenomena at very different scales. Of course, the big challenge in the future is to find more and more appropriate theoretical bridges among these different scales. But the results already accomplished by LQG and entropic gravity make us hope for the best.

7.3 Emerging World

The fact that the laws of QM are probabilistic (Sects. 2.4.2, 3.2.4, and 5.1.1) allows for the establishment of new kinds of reality in our world when appropriate physical contexts and events are produced, according to the propagation of signals, the establishment of local classical correlations (structures), and the role of gravity. We shall synthetically explore such processes. In such a case, we speak of *emergence*.[99]

Emergence is a widespread notion[100] but it is likely that a modern treatment of the concept is due to the British philosopher John Stuart Mill (1806–1873), who in 1843 affirmed that addition of causes does not necessarily imply the proportional addition of effects, but may result in a new configuration (what is especially true when non-linear effects are present)[101]; it is also likely that the first use of the term as such is due to the British philosopher and literature critic George Lewes (1817–1878)) for expressing a similar idea.[102] Today, this concept is very much nuanced and many different formulations are available. However, in general terms, it can be agreed that we have emergence of a new structure that somehow represents or determines a new fundamental property or behaviour whenever such a property or behaviour is robust to fluctuations or variations of variables or conditions out of which it is emerged.[103] Thus, we can say that something is an emergent reality when it cannot be deduced from a theory describing its components or parts.[104]

We have already met emerging phenomena and processes. I have pointed out that the structure of space-time emerges out of quantum information and the propagation of signals, while the mass (for example, quarks) likely emerges out of a quantised field like the Higgs field. Thus, also for elementary systems that QM describes quite well, there are parameters that play an important role but that we are not able to justify quantum mechanically at the moment. This is evident in the case of the most simple atom of the universe: the hydrogen atom. This is very instructive for appreciating the power of the quantum formalism but also its limitations. Moreover, we shall see that, when we pass already to a simple three-body atom or molecule, the physical

[99] Chibbaro et al. (2014, p. 66). For an authoritative physicist's standpoint see Laughlin (2005).

[100] See e.g. Clayton and Davies (2006), Bedau and Humphreys (2008).

[101] Mill (1843, III, Chap. 6).

[102] Lewes (1875). See also Morgan (1923).

[103] Batterman (2002), Mitchell (2009).

[104] Hempel and Oppenheim (1965).

situation can no longer be described by the deterministic probabilistic equations of QM alone.

7.3.1 The Hydrogen Atom

The General Problem

The hydrogen atom represents the simplest chemical element present in our universe and we could expect that quantum equations explain its nature and behaviour quite easily. This is also true to a certain extent. But the number of constraints that are at work here is already amazing and gives us a feeling of how complex is to get a mesoscopic–macroscopic world out of basic quantum laws. Let us consider here some of the basic constraints (we shall simply touch upon some fundamental issues). Abstractly speaking, to put together two particles—a proton and an electron—seems to be an easy job. The libraries are filled with books that in the last centuries supported the idea (that ultimately goes back to the Greek philosopher Democritus (460–370 BC) and partly to the Latin literate Epicurus (341–270 BC)) according to which the pure mechanical collisions of some "atoms" (where obviously only mechanical causation is at play) will finally produce every macroscopic body that we observe. In fact, things are not so simple.

The most reasonable constraint is that the involved particles must be stable, otherwise, we would not have a stable (hydrogen) atom either (whose lifespan at the moment is indefinite). As said, the proton must be a particle composed of three quarks. Now, most of the quarks' combinations are short living or unstable, and this reduces the number of possible candidates (Sect. 7.1.4). In fact, only the quarks up (u) and down (d) can give rise to proton (and neutrons).[105] Moreover, the fact that the proton is composed of quarks, and is therefore heavier than the electron (which is not composed of quarks and thus is an elementary particle, as far as we know at present time) is relevant, since the proton (about 1836 times more massive than the electron) occupies the nucleus of the atom and therefore represents the centre of stability of the whole atomic system, which displays a central potential—an issue that becomes even more relevant when there are many protons and electrons, as is the case for heavier elements. The mass of protons and neutrons (that constitute the overwhelming part of the mass of matter) results in large part out of the residual 'local' (colour) field determined by exchanges of massless gluons among the quarks constituting those particles.[106]

A remarkable aspect of the problem is that the proton and the electron posses an opposite electric charge. The two charges need to balance (which excludes many

[105] Watson (2004).

[106] Wilczek (2006, pp. 74–75). Note that the contribution of quarks to the mass of protons and neutrons is modest, given that the latter have a mass of 938.3 and 939.6 MeV, respectively, while the masses of u and d are 1.7–3.1 and 4.1–5.7 MeV, respectively.

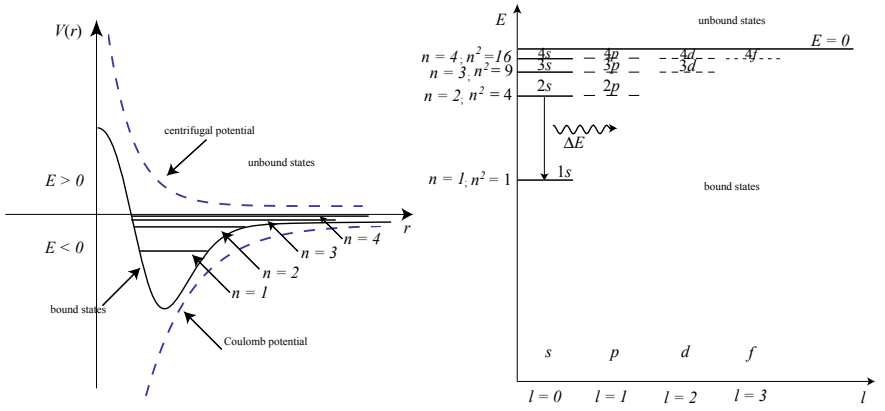


Fig. 7.23 On the left, the electromagnetic (Coulomb) potential as a function of the distance r (on the abscissa) of the electrons from the nucleus. With energy below zero we have trapped (bound) states for the electrons, i.e. possible orbitals at specific energy levels (see also Figs. 1.5, Sect. 1.1, and 6.3, Sect. 1.1). On the right is a schematic representation of the same populated energy levels (only the first 4 shown) as a function of the two quantum numbers n (indicating the orbital shell) and l (which is related to the orbital momentum of the electrons). The quantity n^2 indicates the number of suborbitals for each shell. The arrow shows the transition of an electron from the $n = 2, l = 0$ state to the $n = 1, l = 0$ state, with emission of a photon of energy $\Delta E = E(n = 2) - E(n = 1) = (-E_0/8) - (-E_0/2) = 3E_0/8 = 10.2\text{eV}$, where $-E_0/2$, in agreement with Eq. (7.223), represents the energy of the ground state (the lowest orbital shell attainable in a quantised atom). Adapted from Auletta (2011a, p. 162)

abstract combinations of quarks that would not give the appropriate charge value). For this reason, the only possible combination for protons is uud, since, according to Table 7.1, u has a charge of $+2/3$ while d a charge of $-1/3$, so that uud gives the right value $+1$. At the opposite, antiprotons show the combination $\bar{\text{u}}\bar{\text{u}}\bar{\text{d}}$ with charge -1. Note that the neutron (which does not have natural charge) is produced by the combination udd, what gives in fact 0 charge. In atoms with a nucleus encompassing several particles (both protons and neutrons) there is a residual strong force holding these particles together (especially for overcoming the repulsive electromagnetic force among protons) that is mediated by the massive particles composed of a quark and antiquark (always d and u and their antiparticles) called pions (or π mesons). Since both electric charge and baryon number are conserved quantities, this contributes to the stability of nuclei and to a phenomenal conservation of mass.[107] Obviously, radioactive decay is an exception to nuclear integrity and in fact, massive pions evolve here into massless photons.

Moreover, a dipolar force like the electromagnetic one is much more suitable than a mono-polar one (like the 'gravitational' one) when a combinatorics for producing the different chemical elements need to be at work. Indeed, in the case of a dipolar force, we must have the same number of differently charged particles and the electrons needs

[107] Wilczek (2006, p. 45).

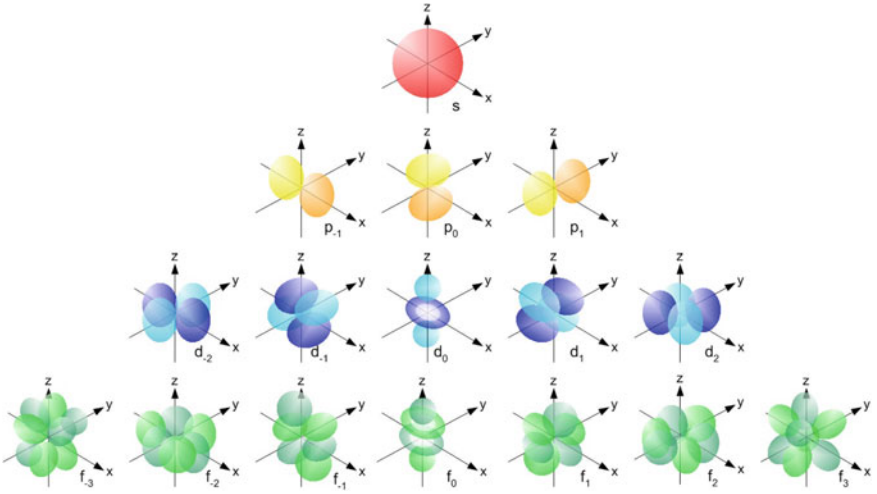


Fig. 7.24 Suborbitals for the first four states. Note that each orbital "position" can be occupied by two electrons with opposite spin, according to Pauli exclusion principle (Sect. 1.3.1). These suborbitals, according to Eqs. (1.270)–(1.272), are 1 for any so-called s state ($l = 0$), 3 for any p state ($l = 1$), 1 for each of its Cartesian components, 5 for any d state ($l = 2$: see also Fig. 1.23, Sect. 1.2), 7 for any f state ($l = 3$). Therefore, level $n = 1$ has energy $-E_0/2$ (representing the ground energy), while the number of orbitals is given by $n^2 = 1$. The level $n = 2$ has energy $-E_0/8$ and number of orbitals $n^2 = 4 = 1 + 3$. The level $n = 3$ has energy $-E_0/18$ and number of orbitals $n^2 = 9 = 1 + 3 + 5$. The level $n = 4$ has energy $-E_0/32$ and number of orbitals $n^2 = 16 = 1 + 3 + 5 + 7$. Adapted from https://ka-perseus-images.s3.amazonaws.com/05cb54e6ff5c2289b76027bb3d74ae8db658f41f.jpg

to fill the orbital levels at a specific position (dictated by the electron's energy) and in specific quantities: this more easily permits a combinatorics (note, however, that when we do not have such a balance as for ions, this contributes to the establishment of chemical connections).

The Schrödinger Equation

The building of atoms is ruled by some numbers.[108] Only electrons, due to the electromagnetic field (i.e. the attractive force exerted by the central proton), can be confined in an atom. When (for example, by absorbing photons) they approach the limit $E = 0$, the orbital shells can become more and more dense (as in the heaviest atoms) up to the jump from discontinuous to continuous spectrum, and then overcoming the centripetal force, they can escape the atom (becoming free particles). So, bound states are 'located' between the electromagnetic (Coulomb) potential and the centrifugal potential (Fig. 7.23, Left panel). In other words, the electric attraction

[108]For all what follows the reader is recommended to study (Auletta et al. 2009, Chap. 11).

mediated by photons contributes negatively to the whole energy of the system (as it happens for any other force or effect decreasing with the distance: see Sect. 7.1.5), and in such a way, any atom is a stable structure precisely because its components cohere into a whole with a total energy that is less than that if they would be separated.

For each energetic level (starting from the level with the lowest energy, that is, the nearest to the proton allowed by the quantisation of the field), there are (Fig. 7.23, Right panel)

- Specific values of the orbital angular momenta l, due to the rotation of electrons (and therefore representing the centrifugal component for that level): Sect. 1.2.4,
- Specific values of the principal quantum number n, indicating the orbital shell (Sect. 6.2.2), and
- Of the number n^2, indicating the number of suborbitals for each shell (Fig. 7.24).

To these numbers we need to add the electric charge e. Also the spin number s plays an important role but I shall not consider this aspect here for simplicity. When the number l grows, we have an increasing number of components for each l as well as for the suborbitals of each orbital shell. I recall that the components are called s, p, d, f states (the terms stand for sharp, principal, diffusive and fundamental). The s states ($n = 1, 2, 3, \ldots, l = 0, m_l = 0$) show a spherical symmetry, while p states ($n = 2, 3, 4, \ldots, l = 1, m_l = -1, 0, 1$) show a double bulk about the $x-$, $y-$ and z-axes, one bulk with a positive sign, the other one with a negative sign, corresponding to the positive or negative phase of the wave function (Fig. 7.24).

I recall that the numbers l, m, n are represented by integers and are *additive* quantum numbers, so that, in the case of composite systems, the resulting number is simply the addition of the numbers of the subsystems (in Sect. 1.2.4 we have considered the case of the addition of angular momenta).

Let us have a look at Fig. 7.25 and consider Eq. (1.265b) as the eigenvalue equation of the angular part of the wave function (1.263). In order to write down the Hamiltonian for the electron, we need to express the nabla operator in spherical coordinates, what is is a little cumbersome. Making use of the components (1.258) and of the chain rule for partial derivatives (1.262), the three components of the nabla operator in spherical coordinates are:

$$\frac{\partial}{\partial x} = \sin\theta\cos\phi\frac{\partial}{\partial r} + \frac{\cos\theta\cos\phi}{r}\frac{\partial}{\partial\theta} - \frac{\sin\phi}{r\sin\theta}\frac{\partial}{\partial\phi}, \tag{7.238a}$$

$$\frac{\partial}{\partial y} = \sin\theta\sin\phi\frac{\partial}{\partial r} + \frac{\cos\theta\sin\phi}{r}\frac{\partial}{\partial\theta} + \frac{\cos\phi}{r\sin\theta}\frac{\partial}{\partial\phi}, \tag{7.238b}$$

$$\frac{\partial}{\partial z} = \cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial\theta}. \tag{7.238c}$$

Then, the square of the nabla operator (or Laplacian) in spherical coordinates is

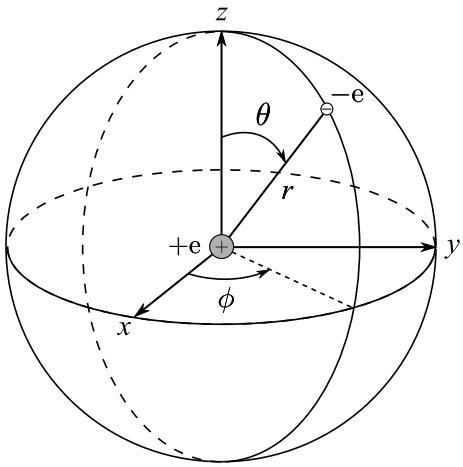


Fig. 7.25 Pictorial representation of an electron (white circle) with negative charge −e orbiting a proton (grey circle) with a positive charge +e in a hydrogen atom (note that the electron in fact does not occupy a specific position on its orbital). The electron wave function is expressed in spherical coordinates as $\psi(r, \theta, \phi)$, where r represents the distance of the electron from the proton, whose position is chosen as the origin, θ is the polar angle and ϕ is the azimuthal angle. Adapted from Auletta and Wang (2014, p. 224)

$$\begin{aligned}\nabla^2 = \boldsymbol{\Delta} &= \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \\ &= \frac{1}{r^2}\left[\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right]. \end{aligned} \tag{7.239}$$

It follows from Eq. (1.261b) that up to a factor of $\hbar^2$ the angular part of the Laplacian (the second and third terms on the second row) is the minus of the magnitude squared of the (orbital) angular momentum operator. As a result, in three dimensions the Hamiltonian for a central potential can be written as (see Eq. (1.158))

$$\begin{aligned}\hat{H} &= -\frac{\hbar^2}{2m}\nabla^2 + V(r) \\ &= -\frac{1}{2m}\left[\frac{\hbar^2}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) - \frac{1}{r^2}\hat{\mathbf{L}}^2\right] + V(r) \\ &= -\frac{\hbar^2}{2mr^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{\hat{\mathbf{L}}^2}{2mr^2} + V(r). \end{aligned} \tag{7.240}$$

In Sect. 1.2.4, when dealing with angular momentum, we could drop any reference to the radial part of the wave function because the radial distance r is conserved under rotations. However, we have here also to deal with the effects of the electromagnetic attraction between proton and electron, and the radial part plays a crucial role. Then,

the radial part of the Schrödinger equation takes the form

$$\left[-\frac{\hbar^2}{2mr^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right)+\frac{l(l+1)\hbar^2}{2mr^2}+V(r)\right]f(r)=Ef(r), \tag{7.241}$$

where $f(r)$ is the radial component of the wave function (1.263) and the three terms on the LHS correspond to the three parts in the last line of Eq. (7.240). It may be noted that the square of the angular momentum contribution (the second term) takes the form of its eigenvalues (1.253). Obviously, since the above expression represents the eigenvalue equation of the Hamiltonian, we have the eigenvalue E on the RHS. By performing now a change of variables

$$f(r)=\frac{\xi(r)}{r}\quad (r>0), \tag{7.242}$$

we can obtain a further simplification of the first term on the LHS of the above equation:

$$\begin{aligned}\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right)\frac{\xi(r)}{r}&=\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\frac{\xi(r)}{r}\right)\\&=\frac{1}{r^2}\frac{\partial}{\partial r}\left(r\xi'(r)-\xi(r)\right)\\&=\frac{1}{r}\xi''(r)+\frac{1}{r^2}\xi'(r)-\frac{1}{r^2}\xi'(r)\\&=\frac{1}{r}\xi''(r),\end{aligned} \tag{7.243}$$

where

$$\xi'(r)=\frac{\partial\xi(r)}{\partial r},\quad \xi''(r)=\frac{\partial^2\xi(r)}{\partial r^2}, \tag{7.244}$$

and use has been made of the product rule (1.206). This result allows us to simplify Eq. (7.241) to

$$-\frac{\hbar^2}{2m}\xi''(r)+\left[\frac{\hbar^2 l(l+1)}{2mr^2}+V(r)\right]\xi(r)=E\xi(r). \tag{7.245}$$

I remark that, expressed in terms of $\xi(r)$, the radial equation (7.245) is formally identical to the 1D Schrödinger equation (1.203) for a particle moving in an effective potential given by

$$V_{\text{eff}}(r)=V(r)+\frac{l(l+1)\hbar^2}{2mr^2}. \tag{7.246}$$

As mentioned, the additional term in the effective potential originates from the orbital angular momentum of the electron. It is called the *centrifugal potential barrier* and being repulsive, prevents the electron from reaching the centre of the potential (Fig. 7.23, Left panel).

Let us now apply the previous formalism to the hydrogen atom in particular. In this case, the parameter m is represented by the electron mass $m_\mathrm{e} = 9.109 \times 10^{-31}$ kg and the potential energy $V(r)$ is the celebrated Coulomb (electric) potential, after the name of Charles-Augustin de Coulomb,

$$V(r) = -\frac{1}{4\pi\epsilon_0}\frac{\mathrm{e}^2}{r}, \tag{7.247}$$

where ϵ_0 is the vacuum permittivity (Sect. 7.1.1). This potential represents the electrostatic interaction between a proton with charge e and an electron with charge $-\mathrm{e}$ at a distance r from the proton. Note that the mass of a 'free proton' and a 'free neutron' is $1.67262158 \times 10^{-27}$ kg and $1.67492716 \times 10^{-27}$ kg (which are largely lower than the Planck's mass (3.42c)), respectively. The sum amounts to $3.34754874 \times 10^{-27}$ kg. However, the mass of a 'free' deuteron is only $3.34358309 \times 10^{-27}$ kg. The difference is $-0.00396565 \times 10^{-27}$ kg. This problem is called "mass defect". The reason is that the binding of the proton and neutron to form the deuteron releases a certain amount of energy which is equivalent to the mass defect. Note that here we take the mass of the electron (as well as of quarks and the proton) as bare facts extrinsic to the quantum-mechanical description. The values of electric charge and the vacuum permittivity in SI units are given by

$$\mathrm{e} = 1.602 \times 10^{-19}\,\mathrm{C} \quad \text{and} \quad \epsilon_0 = 8.854 \times 10^{-12}\,\mathrm{F \cdot m^{-1}}, \tag{7.248}$$

respectively, where C is the symbol for the electric charge unit coulomb and $\mathrm{F} = \mathrm{C/V}$ (where V is the volt) is the symbol for the capacitance unit farad. Using the Coulomb potential and the electron mass, the radial equation (7.245), after rearrangement of the terms, can be rewritten as[109]

$$\xi''(r) + \frac{2m_\mathrm{e}}{\hbar^2}\left[\frac{1}{4\pi\epsilon_0}\frac{\mathrm{e}^2}{r} - \frac{l(l+1)\hbar^2}{2m_\mathrm{e}r^2} + E\right]\xi(r) = 0. \tag{7.249}$$

In physics, it is always advisable to make equations dimensionless by introducing natural units characteristic to the problem. In the present problem, let us first introduce two natural units that can be built from the physical constants m_e, $\hbar$, e and ϵ_0: the

[109] Our analysis here assumes that the proton is infinitely massive with respect to the electron. For actual proton mass $m_\mathrm{p} \approx 1836\, m_\mathrm{e}$, the electron mass m_e that appears in the Hamiltonian (7.240) should be replaced by the so-called reduced mass

$$m = \frac{m_\mathrm{e} m_\mathrm{p}}{m_\mathrm{e} + m_\mathrm{p}} \approx 0.995\, m_\mathrm{e},$$

which to an excellent approximation can be identified with the electron mass.

natural unit of length is the *Bohr radius* a_0, and the natural unit of energy E_0. They are, respectively, defined as

$$a_0 := \frac{4\pi\epsilon_0\hbar^2}{m_e e^2} = 5.29 \times 10^{-11}\,\text{m} \quad \text{and} \quad E_0 := \frac{m_e e^4}{\hbar^2} = 27.2\ \text{eV}, \tag{7.250}$$

where eV stands for electron volt and is a unit of energy equal to approximately $1.60217653 \times 10^{-19}$ J. The Bohr radius is equal to the most probable distance between the proton and the electron in a hydrogen atom in its ground state, while $E_0/2$ represents the ground state energy of the hydrogen atom.

The previous definitions allow us now to define the dimensionless variables $\tilde{r}$ and $\tilde{E}$:

$$\tilde{r} := \frac{r}{a_0}, \quad \tilde{E} := \frac{E}{E_0}, \tag{7.251}$$

with

$$r_0^2 \frac{d^2}{dr^2}\xi = \frac{d^2}{d\tilde{r}^2}\xi\,, \tag{7.252}$$

from which we can derive the dimensionless squared principal quantum number

$$n^2 = \frac{1}{2\tilde{E}}. \tag{7.253}$$

With these choices, it is possible to rewrite Eq. (7.249) as

$$\left[\frac{d^2}{d\tilde{r}^2} + \frac{2}{\tilde{r}} - \frac{l(l+1)}{\tilde{r}^2} + 2\tilde{E}\right]\xi(\tilde{r}) = 0\,. \tag{7.254}$$

The asymptotic behaviour of the solution of Eq. (7.254) for $\tilde{r} \to \infty$ and for $\tilde{r} \to 0$ suggests the assumption that $\xi(\tilde{r})$ has the form

$$\xi(\tilde{r}) = \tilde{r}^{l+1} e^{-\frac{\tilde{r}}{n}} W(\tilde{r}), \tag{7.255}$$

for some function $W(\tilde{r})$. Note that the first and second derivatives of $\xi(\tilde{r})$ are

$$\frac{\partial}{\partial\tilde{r}}\xi(\tilde{r}) = \left(\frac{l+1}{\tilde{r}} - \frac{1}{n} + \frac{W'}{W}\right)\tilde{r}^{l+1} e^{-\frac{\tilde{r}}{n}} W, \tag{7.256a}$$

$$\frac{\partial^2}{\partial\tilde{r}^2}\xi(\tilde{r}) = \left[-\frac{l+1}{\tilde{r}^2} + \frac{W''}{W} - \frac{(W')^2}{W^2} + \left(\frac{l+1}{\tilde{r}} - \frac{1}{n} + \frac{W'}{W}\right)^2\right]\tilde{r}^{l+1} e^{-\frac{\tilde{r}}{n}} W. \tag{7.256b}$$

Substituting Eqs. (7.256b) and (7.255) into Eq. (7.254), and using Eq. (7.253), we obtain

$$0 = \left[\left(-\frac{l+1}{\tilde{r}^2} + \frac{W''}{W} - \frac{(W')^2}{W^2} + \frac{(l+1)^2}{\tilde{r}^2} + \frac{1}{n^2} + \frac{(W')^2}{W^2} - \frac{2(l+1)}{n\tilde{r}} + \frac{2(l+1)W'}{\tilde{r}W} - \frac{2W'}{nW}\right)\right.$$
$$\left. - \left(\frac{1}{n^2} - \frac{2}{\tilde{r}} + \frac{l(l+1)}{\tilde{r}^2}\right)\right]\xi$$
$$= \left[\frac{W''}{W} + \left(\frac{2(l+1)}{\tilde{r}} - \frac{2}{n}\right)\frac{W'}{W} + \frac{2}{\tilde{r}}\left(1 - \frac{l+1}{n}\right)\right]\xi\,, \tag{7.257}$$

and by multiplying by $W\tilde{r}$, we have

$$\left[\tilde{r}W'' + 2\left(l+1-\frac{\tilde{r}}{n}\right)W' + 2\frac{n-l-1}{n}W\right]\xi(\tilde{r}) = 0\,, \tag{7.258}$$

which allows us to split the equation into two parts and to consider only the composed term between square brackets. By setting

$$\eta := \frac{2\tilde{r}}{n}, \tag{7.259}$$

with $n > 0$, we can obtain the final Schrödinger equation for the radial part

$$\eta W''(\eta) + n\left[(l+1) - \frac{\eta}{2}\right]W'(\eta) + (n-l-1)W(\eta) = 0. \tag{7.260}$$

A similar result can be derived for $n < 0$. This is due to the fact that

$$W(\eta) = F\left(-|n| + l + 1; 2l + 2; \eta\right)\,, \tag{7.261}$$

where F is called a confluent hypergeometric function.

Eigenvalues and Eigenfunctions

From the previous equation, we infer that the general form of the involved functions must be:

$$W(\eta) = \sum_{j=0}^{\infty} c_j \eta^j, \tag{7.262}$$

where the coefficients c_j need to be determined. The first two derivatives of $W(\eta)$ with respect to η are given by

$$W'(\eta) = \sum_{j=1}^{\infty} j c_j \eta^{j-1}, \quad W''(\eta) = \sum_{j=2}^{\infty} j(j-1) c_j \eta^{j-2}. \tag{7.263}$$

Substituting the above expressions into Eq. (7.260), we can obtain the following recurrence relation between successive coefficients:

$$c_{j+1} = \frac{j+l+1-n}{(j+1)[j+2(l+1)]} c_j \quad (j = 0, 1, 2, \dots). \tag{7.264}$$

For instance, the first three terms are given by

$$c_0, \quad c_1 = \frac{l+1-n}{2(l+1)} c_0, \quad c_2 = \frac{l+2-n}{2(2l+3)} c_1. \tag{7.265}$$

If we arbitrarily fix $c_0 = 1$, then all c_j can be obtained. Since $W(\eta)$ cannot grow as e^{η} when $\eta \to \infty$, we need to require that the series (7.262) be truncated at a certain value $j = k$: the radial wave function need to contain only a finite number of terms. In order for this to happen, it is necessary that we have

$$n = l + 1 + k \ , \tag{7.266}$$

where k must be a non-negative integer. Since l is an integer number, also n must be an integer number. Thus, for each integer value of n, using the definition (7.253), the corresponding value (eigenvalue) of energy is given by

$$E_n = E_0 \tilde{E} = -\frac{1}{2n^2} E_0 \ , \tag{7.267}$$

where $n = 1, 2, \dots$. As anticipated, the energy levels of the bound states become more and more dense as n goes to infinity, i.e. as the energy approaches the limiting value $E = 0$.

This analysis allows us to compute the radial wave functions $f_{nl}(r)$ for the different orbital levels. For $f_{10}(r)$, we have $c_0 = 1$ and thanks to the recursive relation (7.264), we have that $\forall j \geq 1$, $c_j = 0$. Taking advantage of the definition (7.242), we have that $f(r) = \xi(\tilde{r})/\tilde{r}$. Then, $W(\eta) = c_0 \eta^0 = 1$, from which it follows that (see Eq. (7.255)) $\xi(\tilde{r}) = \tilde{r} e^{-\tilde{r}}$. Now, the eigenfunctions f_{nl}, being the spectrum discrete, need to satisfy the normalisation condition

$$\int_0^{+\infty} dr \, |f_{nl}(r)|^2 \, r^2 = 1 \ , \tag{7.268}$$

so that we have

$$f_{10} = \mathcal{N}_{10} e^{-\tilde{r}}, \tag{7.269}$$

where $\mathcal{N}_{10}$ is the normalisation constant. This gives

$$\mathcal{N}_{10}^2 \int_0^{+\infty} dr e^{-\frac{2r}{a_0}} r^2 = 1 \ . \tag{7.270}$$

Integrating by parts (Eq. (1.30)), we have $\forall x$

$$\int_0^{+\infty} dx e^{-x} x^2 = 2 \,. \tag{7.271}$$

This implies

$$2\left(\frac{a_0}{2}\right)^3 \mathcal{N}_{10}^2 = 1 \text{ or } \mathcal{N}_{10} = 2a_0^{-\frac{3}{2}} \,. \tag{7.272}$$

For f_{20}, we have

$$c_1 = \frac{1-2}{2} c_0 = -\frac{1}{2} c_0, \quad c_2 = \frac{2-2}{6} c_1 = 0, \tag{7.273}$$

with all other $c_j = 0$ for $j > 2$. This means that we have

$$W(\eta) = c_0 - c_0 \frac{1}{2} \eta = 1 - \frac{\widetilde{r}}{2}. \tag{7.274}$$

By taking into account Eqs. (7.255) and (7.259), we derive

$$\xi(\tilde{r}) = c_0 \widetilde{r}\, e^{-\frac{\widetilde{r}}{2}} \left(1 - \frac{\widetilde{r}}{2}\right), \tag{7.275}$$

which finally gives

$$f_{20} = \mathcal{N}_{20} \left(1 - \frac{r}{2a_0}\right) e^{-\frac{r}{2a_0}}, \tag{7.276}$$

where $\mathcal{N}_{20}$ is the normalisation constant.

For $n = 2$ and $l = 1$, the only non-zero coefficient is c_0. This means $W(\eta) = 1$ and

$$\xi(\tilde{r}) = \widetilde{r}^2\, e^{-\frac{\widetilde{r}}{2}}. \tag{7.277}$$

Hence we have

$$f_{21} = \mathcal{N}_{21} \left(\frac{r}{a_0}\right) e^{-\frac{r}{2a_0}}, \tag{7.278}$$

where $\mathcal{N}_{21}$ is the normalisation constant.

Summarising and inserting the explicit forms of the normalisation constants, the first few radial eigenfunctions f_{nl} are (Fig. 7.26)

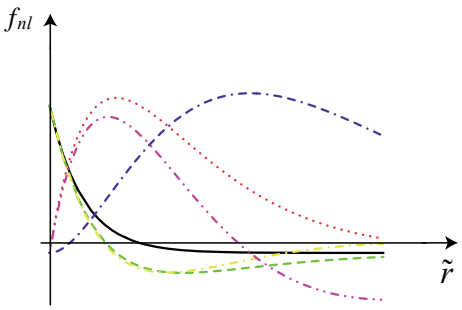


Fig. 7.26 Plot of the first few radial eigenfunctions of the hydrogenoid atom: f_{10} (solid line), f_{20} (dashed line), f_{21} (dotted line), f_{30} (dot-dashed line), f_{31} (dot–dot–dashed line) and f_{32} (dot–dash–dashed line). Adapted from Auletta et al. (2009, p. 412)

$$f_{10}(r) = 2\left(\frac{1}{a_0}\right)^{\frac{3}{2}} e^{-\frac{r}{a_0}}, \tag{7.279a}$$

$$f_{20}(r) = \frac{1}{\sqrt{2}}\left(\frac{1}{a_0}\right)^{\frac{3}{2}}\left(1-\frac{r}{2a_0}\right) e^{-\frac{r}{2a_0}}, \tag{7.279b}$$

$$f_{21}(r) = \frac{1}{2\sqrt{6}}\left(\frac{1}{a_0}\right)^{\frac{3}{2}}\frac{r}{a_0} e^{-\frac{r}{2a_0}}. \tag{7.279c}$$

Further calculations show that we must have

$$f_{30}(\tilde{r}) \propto e^{-\frac{\tilde{r}}{3}}\left(1-\eta+\frac{1}{6}\eta^2\right) = e^{-\frac{\tilde{r}}{3}}\left(1-\frac{2}{3}\tilde{r}+\frac{2}{27}\tilde{r}^2\right), \tag{7.279d}$$

$$f_{31}(\tilde{r}) \propto \eta\left(1-\frac{1}{4}\eta\right)e^{-\frac{\eta}{2}} = \frac{2}{3}\tilde{r}\left(1-\frac{1}{6}\tilde{r}\right)e^{-\frac{\tilde{r}}{3}}, \tag{7.279e}$$

$$f_{32}(\tilde{r}) \propto \eta^2 e^{-\frac{\eta}{2}} = \frac{4}{9}\tilde{r}^2 e^{-\frac{\tilde{r}}{3}}, \tag{7.279f}$$

where I have not considered the normalisation constants.

I also note that the radial probability density for the electron, giving the maximal probability to find it at a distance r from the nucleus (Fig. 7.27) is

$$\wp_{nl}(\tilde{r}) = |f_{nl}(r)|^2 r^2 . \tag{7.280}$$

By rewriting Eq. (1.263) in the explicit form

$$\psi_{nlm}(r,\theta,\phi) = f_{nl}(r)Y_{lm}(\theta,\phi) , \tag{7.281}$$

we can write down the zero-angular momentum total eigenfunctions $\psi_{nlm}(r,\theta,\phi)$ for $n = 1, 2, 3$:

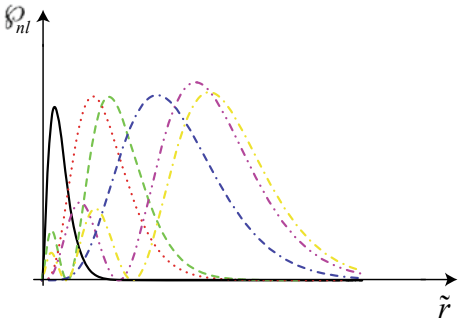


Fig. 7.27 Plot of the radial probability densities (7.280) for the first few radial eigenfunctions of the hydrogenoid atom (same convention as in Fig. 7.26). Adapted from Auletta et al. (2009, p. 413)

$$\psi_{100}(\mathbf{r}) = \frac{1}{\sqrt{\pi}} \left(\frac{1}{a_0}\right)^{\frac{3}{2}} e^{-\frac{r}{a_0}} , \tag{7.282a}$$

$$\psi_{200}(\mathbf{r}) = \frac{1}{4\sqrt{2\pi}} \left(\frac{1}{a_0}\right)^{\frac{3}{2}} \left(2 - \frac{r}{a_0}\right) e^{-\frac{r}{2a_0}} , \tag{7.282b}$$

$$\psi_{300}(\mathbf{r}) = \frac{1}{81\sqrt{3\pi}} \left(\frac{1}{a_0}\right)^{\frac{3}{2}} \left(27 - 18\frac{r}{a_0} + \frac{2r^2}{a_0^2}\right) e^{-\frac{r}{3a_0}} . \tag{7.282c}$$

As mentioned, I have made several simplifications. In the previous forms of the total eigenfunctions, I have not considered the angular momentum. Moreover, in the Schrödinger equation I have not considered the spin component (see Eqs. (7.99) and (7.137)). Note also that there are additional constraints for getting a hydrogen atom than those considered here, for instance, due to magnetic (like the Paschen-Bach or Zeeman) effects (related to the spin and angular momentum) or to relativistic (like the Darwin) correction.[110] As said, the simplest element of the universe is already much more than a system described by a basic Schrödinger equation.

A Look at the Chemical Elements

Generalising what previously said, I shall try to summarise now the building of the chemical elements.[111] In fact, they can be built in a 'mechanical way', although the details are not described by deterministic equations. The spherical symmetry of the s states is a consequence of QM and especially of the spatial symmetry of the wave function. Also the other symmetries of molecules are a consequence of quantum-mechanical distributions of atoms and electrons. Taking into account the limitation imposed by the exclusion principle, we can have two electrons in the first shell (denoted by $1s$), 8 electrons in the second shell (2 electrons for $2s$ and two times the three axes = 6 for $2p$), 18 electrons for the third shell (2 for $3s$, 6 for $3p$ and 10

[110] For these issues, I suggest the reader to study Auletta et al. (2009, Sects. 11.3–11.4).

[111] For this and the next subsection see Atkins and De Paula (2006), Atkins and Friedman (2005).

for $3d$), 32 for the fourth shell (2 for $4s$, 6 for $4p$, 10 for $4d$ and 14 for $4f$) and so on, where recall (Fig. 7.24) that the s states can have at most two electrons, the p states at most six electrons, the d states at most ten, the f states at most forteen. The order of occupation is

$$1s \;\; 2s \;\; 2p \;\; 3s \;\; 3p \;\; 4s \;\; 3d \;\; 4p \;\; 5s \;\; 4d \;\; 5p \;\; 6s \;\; 4f \;\; 5d \;\; 6p \;\; 7s \;\; 5f \;\; 6d \;\; 7p$$

In such a way, we obtain the 118 elements, from hydrogen (with 1 electron and 1 proton) up to organessum (with 118 protons). All chemical elements are organised in 4 major blocks (Fig. 7.28):

- The s block (labelled by a red bullet), with 2 columns, each column ranging from $1s$ to $7s$ (groups 1–2 plus helium),
- The p block (labelled with a yellow bullet), with 6 columns, each column ranging from $2p$ to $7p$ (groups 13–18),
- The d block (labelled by a blue bullet), with 10 columns, each column ranging from $3d$ to $6d$ (groups 3–12),
- The f block (labelled by a green bullet), with 14 columns, each column ranging from $4f$ to $5f$ (the only exception represented by thorium).

The term *period* refers to the occupation number identifying those elements with the same electron shells (s and p). For instance, elements with period 2 always involve the $2s$ and $2p$ orbitals, elements of period 3 the $3s$ and $3p$ orbitals, and so on (the only exception represented by palladium of period 5, which presents $4s^2$ instead of $5s$). Note that exceptions depend on the order of occupation. The term *group* identifies common characters of the outmost electron shells and thus similar chemical and physical (emergent) properties. In particular, by increasing the group number, we have less metallic elements, an increase both in ionisation energy and electronegativity, and a progressive relative reduction of the atomic radius.

As said, any orbital (energy level) of an atom can be occupied at most by two electrons with opposite spin. Let us consider some elements. Hydrogen (H) has the atomic number 1 (1 electron, 1 proton, no neutrons), followed by helium (He), with atomic number 2 (2 electrons at the same energy level, 2 protons, 2 neutrons). Lithium (Li) has the atomic number 3 (3 electrons disposed within two orbitals, 3 protons and 4 neutrons). Berillium (Be) has the atomic number 4 (4 electrons in two orbitals, 4 protons, 5 neutrons), the boron (B) which has the atomic number 5 (5 electrons in 3 levels, 5, 6). The carbon (C) has the atomic number 6 (6, 6, 6), nitrogen (N) a. n. 7 (7, 7, 7), oxygen (O) a.n. 8 (8, 8, 8), fluorine (F) a. n. 9 (9, 9, 10), neon (Ne) a. n. 10 (10, 10, 10). The second layer of elements (from sodium (Na) to argon (Ar)) have atomic numbers from 11 to 18. The third layer goes from potassium (K) to krypton (Kr), with atomic numbers from 19 to 36. The fourth layer from rubidium (Rb) to xenon (Xe), with atomic numbers from 37 to 54.

As anticipated, it is evident that we have here a kind of combinatorics that allows the building of more and more heavy elements by simply adding protons in the nucleus and relative electrons in the appropriate orbital shells. This combinatorics

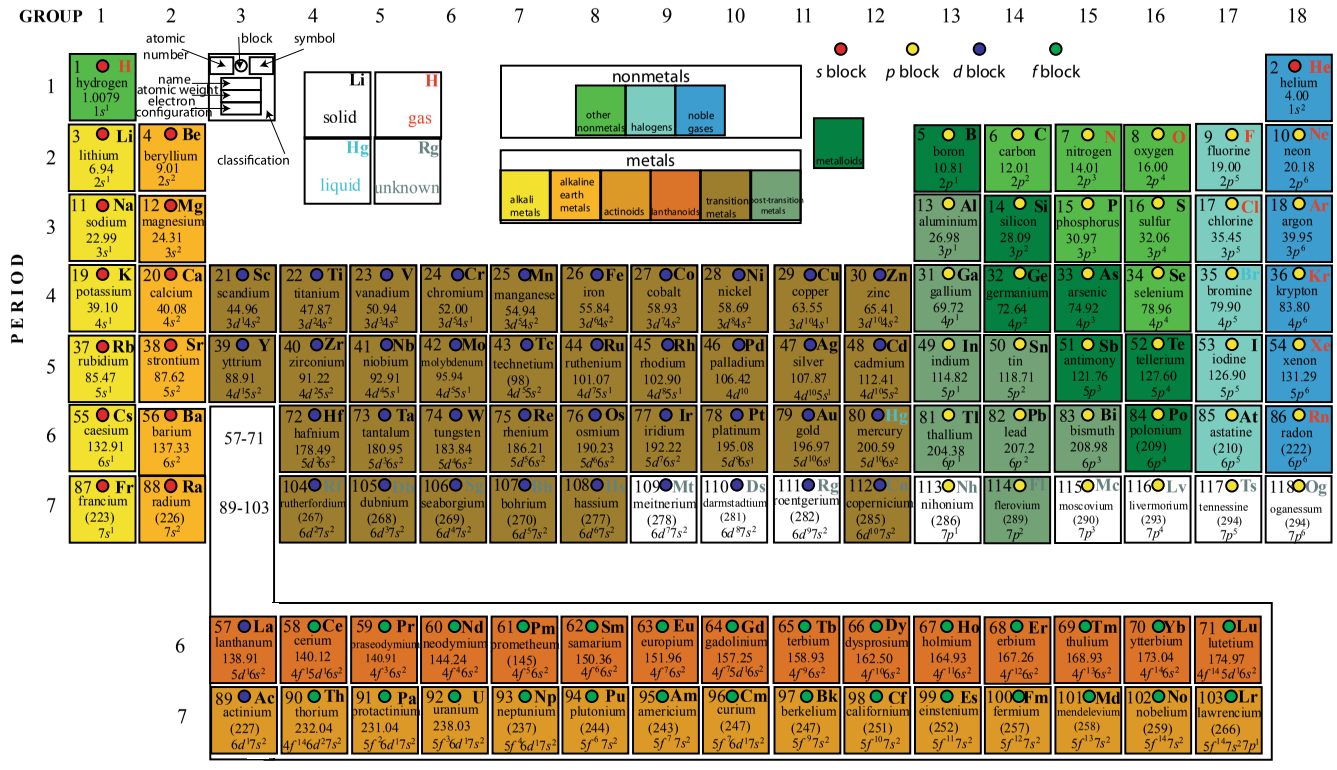


Fig. 7.28 Periodic table of elements. The atomic number of the element (number of protons), symbol, name, atomic weight, element category and relevant orbitals are shown. Note that, all lanthanoids and actinoids have period 6 and 7, respectively; lanthanum (a.n. 57) and actinium (a.n. 89) pertain to group 3 while the remnant 28 ones do not pertain to numbered groups. Adapted from Auletta (2011a, p. 165)

is also evident in the building of molecules, as we shall see later. However, we shall discover that there are also characteristic differences relative to the combinatorics of all possibilities for the state of an isolated quantum system.

7.3.2 Approximation Methods

In general, it is not possible to find the exact solution (i.e. the exact energy eigenvalues and eigenfunctions) for the Schrödinger equation of any molecule and even any atom except hydrogen. Already when just three particles interact (as it is the case, for example for the helium atom or the hydrogen molecular ion, with two protons and one electron: see Fig. 7.30), no deterministic description (it does not matter whether quantum or classical) is possible.[112] This problem is already known in classical physics and is called the problem of the three interacting bodies,[113] because there is no closed-form solution of the set of differential equations describing this physical situation.[114] For instance, in order to solve such a problem for the planetary system, the astronomer Johannes Kepler treated the bodies as point masses and reduced the problem to a set of equations describing one-to-one interactions (either sun–planet or planet–satellite). In such a case, we need to use approximations and statistical methods.

Here, I shall summarise the most popular methods used for the estimate of the electronic states in the diatomic molecules: the *Linear Combination of Atomic Orbitals* (LCAO) and the *Valence Bond Method* (VBM). Their starting point is slightly different, however, both these methods have the goal to find the best estimate of the exact solution. What is very helpful in both methods is that we can omit any consideration of the angular momentum since in most cases the involved particles occupy stable relative 'positions' or can be treated so.

LCAO

The LCAO method assumes that the wave function of a set of atoms in a molecule can be appropriately dealt with as a linear combination of the original atomic wave functions with certain weights that have to be determined. In other words, according to this method the original wave functions are only slightly modified due to the formation of the molecule. In order to find the best estimate, one has then to apply a variational method and obtain the value of the parameters (the weights of the atomic components) that minimise the energy.

Let us consider two atoms, labelled a and b, whose electronic states are described by the wave functions ψ_a and ψ_b, respectively. We then make the assumption that

[112] Auletta et al. (2009, Chap. 12).

[113] Poincaré (1890).

[114] Barrow-Green (1997).

the electronic state of the a-b molecule may be expressed as the superposition

$$\Psi = c_a\psi_a + c_b\psi_b \ , \tag{7.283a}$$

or

$$\Psi = \psi_a + b\psi_b \ , \tag{7.283b}$$

where $b = c_b/c_a$ (with $c_a \neq 0$). In this case, if ψ_a and ψ_b are normalised, the resulting wave function Ψ is not normalised. Denoting by $\hat{H}$ the total Hamiltonian of the molecule, we have the Schrödinger equation

$$\hat{H}\Psi = E\Psi \ , \tag{7.284a}$$

or

$$\hat{H}\psi_a + b\hat{H}\psi_b = E\psi_a + bE\psi_b \ . \tag{7.284b}$$

Multiplying both sides first by ψ_a^* and then by ψ_b^* in two separate steps and integrating, we obtain the set of two equations

$$H_{aa} - E + b(H_{ab} - ES_{ab}) = 0 \ , \tag{7.285a}$$
$$H_{ab} - ES_{ab} + b(H_{bb} - E) = 0 \ , \tag{7.285b}$$

where

$$H_{ab} = \int d\mathrm{V}\psi_a^*\hat{H}\psi_b = \beta \tag{7.286}$$

is the so-called *resonance integral*, and

$$S_{ab} = \int d\mathrm{V}\psi_a^*\psi_b = S \ , \tag{7.287}$$

with $0 \leq S \leq 1$, is the so-called *overlapping integral*, $d\mathrm{V}$ being the electron's volume element. Similarly,

$$H_{aa} = \int d\mathrm{V}\psi_a^*\hat{H}\psi_a \text{ and } H_{bb} = \int d\mathrm{V}\psi_b^*\hat{H}\psi_b. \tag{7.288}$$

From the product of Eqs. (7.285) we can obtain

$$(E_a - E)(E_b - E) - (\beta - ES)^2 = f(E) = 0 \ , \tag{7.289}$$

where I have used the simplified notation $E_a = H_{aa}$, $E_b = H_{bb}$, which should not be confused with the energy eigenvalues of the atoms a and b, respectively. Eq. (7.289) is an implicit quadratic equation for E, given E_a, E_b, β and S. The function $f(E)$ is

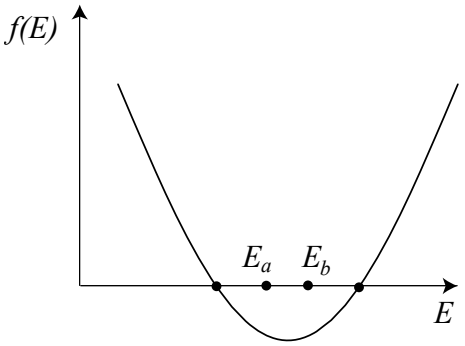


Fig. 7.29 Schematic diagram of the function $f(E)$ (see Eq. (7.289)). The zeros of $f(E)$ represent the LCAO solutions for the energy of the molecule a-b. Adapted from Auletta et al. (2009, p. 450)

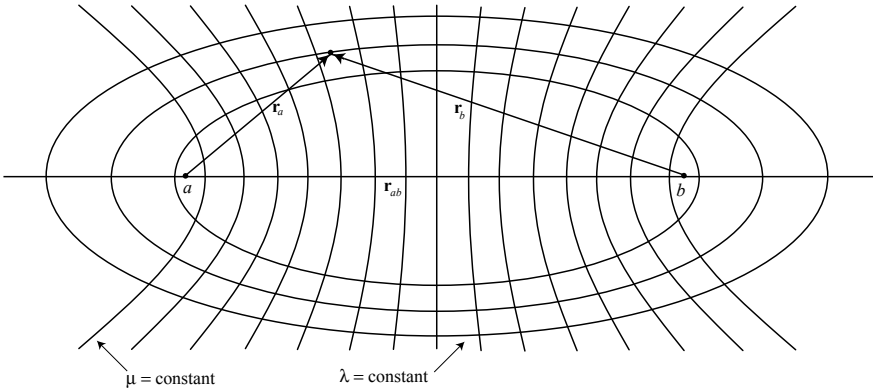


Fig. 7.30 Spheroidal coordinates for the H_2^+ ion. The two hydrogen nuclei (i.e. the two protons) a, b are located at distance $\mathbf{r}_{ab}$ at the foci of the ellipses (λ = constant) and hyperbolas (μ = constant). The electron is located at distances $\mathbf{r}_a$ and $\mathbf{r}_b$ from the nuclei a and b, respectively. Besides λ and μ, the third coordinate is given by the rotation angle ϕ about the nuclear axis ab. Adapted from Auletta et al. (2009, p. 440)

schematically represented in Fig. 7.29, and its zeros represent the LCAO approximate solutions to the molecular energy.

In the case of the hydrogen molecular ion (H_2^+, composed of two protons and one electron), by effecting the separation of variables within the framework of the so-called *spheroidal* (or elliptical) coordinates (see Fig. 7.30), where

$$\lambda = \frac{\mathbf{r}_a + \mathbf{r}_b}{\mathbf{r}_{ab}} \text{ and } \mu = \frac{\mathbf{r}_a - \mathbf{r}_b}{\mathbf{r}_{ab}} , \tag{7.290}$$

with $1 \leq \lambda \leq +\infty$ and $-1 \leq \mu \leq +1$, we can write the Hamiltonian as (see the first row of Eqs. (7.240) and (7.247))

$$\hat{H} = -\frac{\hat{\mathbf{p}}^2}{2m} + V(\mathbf{r}) = -\frac{\hbar^2}{2m}\boldsymbol{\nabla}^2 - \frac{\mathrm{e}^2}{\mathbf{r}_a} - \frac{\mathrm{e}^2}{\mathbf{r}_b} + \frac{\mathrm{e}^2}{\mathbf{r}_{ab}} \,. \tag{7.291}$$

Since the molecule is homonuclear (i.e. composed of the same atomic elements), we have $|b|^2 = 1$, or $b = \pm 1$, which implies a symmetric and an antisymmetric wave function

$$\Psi_+ = \psi_a + \psi_b \,, \tag{7.292a}$$
$$\Psi_- = \psi_a - \psi_b \,, \tag{7.292b}$$

with

$$E_a = E_b = E_0 \,. \tag{7.293}$$

Thus, in the case of Ψ_+, Eqs. (7.285) become

$$(E_0 - E_+) + \beta - E_+ S = 0 \,, \tag{7.294a}$$
$$\beta - E_+ S + (E_0 - E_+) = 0 \,, \tag{7.294b}$$

that lead to

$$E_+ = \frac{E_0 + \beta}{1 + S} \,. \tag{7.295a}$$

Analogously, we can derive

$$E_- = \frac{E_0 - \beta}{1 - S} \,. \tag{7.295b}$$

The corresponding wave functions are given by

$$\psi_a(r) = C_a e^{-\frac{r}{a_0}} \,, \tag{7.296a}$$
$$\psi_b(r) = C_b e^{-\frac{(r - r_{ab})}{a_0}} \,, \tag{7.296b}$$

where C_a and C_b are constants and a_0 is the Bohr's radius of the hydrogen atom (7.250). In the LCAO approximation, therefore, two energy levels for the molecule are possible, and are separated by an energy gap $E_- - E_+$ (see Fig. 7.31).

In Fig. 7.32, the graph of Ψ_+ and Ψ_- as a function of $\mathbf{r}_{ab}$ are shown. As expected, only Ψ_+ has a minimum for $\mathbf{r}_{ab} \approx 2a_0$, with $E \approx -0.56\mathrm{e}^2/a_0$. The excitation from the state Ψ_+ to the state Ψ_- is already sufficient to break the molecule. Similar considerations can be developed for the case of other homonuclear diatomic molecules.

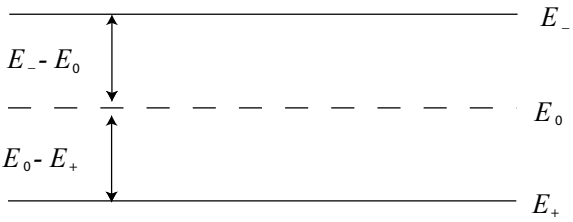


Fig. 7.31 LCAO energy solutions for the H_2^+ molecular ion. Only he symmetric Ψ_+ state gives rise to a stable molecule with energy eigenvalue E_+. On the contrary, the antisymmetric state Ψ_- with energy eigenvalue E_- is unstable. Adapted from Auletta et al. (2009, p. 451)

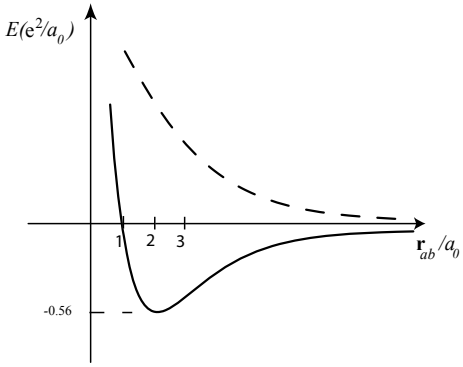


Fig. 7.32 Graphic representation of the symmetric (solid line) and antisymmetric (dashed line) states of the ground level of the H_2^+ molecule. Adapted from Auletta et al. (2009, p. 452)

VBM

The valence bond method is for certain respects similar to the LCAO approximation. Nevertheless, it attempts to obtain the molecular wave function as a product—rather than a linear superposition—of the original atomic wave function. While in the LCAO method we first build the molecular states that are successively filled with the existing electrons, the VBM makes use of the bonds between decoupled electrons, forgetting about the internal electronic levels (complete shells).

Differently from previous cases, we need here to consider the case of the hydrogen molecule H_2, since, as mentioned, the VBM works with bonds, i.e. with molecules presenting at least two electrons. We start with the product wave function

$$\Psi(1,2) = \psi_a(1)\psi_b(2) \,, \tag{7.297}$$

where 1 and 2 denote the two electrons. Since the wave function (7.297) is not symmetrised, a better assumption is

$$\Psi^{\pm}_{\text{cov}}(1,2) = \psi_a(1)\psi_b(2) \pm \psi_a(2)\psi_b(1) \,, \tag{7.298}$$

sometimes called the *covalent wave function*: Ψ^{+}_{cov} yields a dissociation energy of 3.15 eV (relative to the experimental value of about 4.8 eV). The approximation can be further improved by accepting a small probability of a *ionic* wave function

$$\Psi^{\pm}_{\text{ion}}(1,2) = \psi_a(1)\psi_a(2) \pm \psi_b(1)\psi_b(2) \ . \tag{7.299}$$

A covalent bond involves the sharing of electron pairs between atoms, while a ionic bond involves the electrostatic attraction between oppositely charged ions. Then, we can write the total wave function as

$$\Psi(1,2) = \Psi^{\pm}_{\text{cov}}(1,2) + \alpha\Psi^{\pm}_{\text{ion}}(1,2) \ , \tag{7.300}$$

where α should be calculated with the variational method. For the case of the hydrogen molecule, it turns out that $\alpha = 0.25$, i.e. the weight of the ionic bond is about 25%. With this improvement, we arrive at a value for the dissociation energy of about 4.1 eV. Other small improvements may be obtained if one accepts a certain polarisation of the atomic orbitals due to the electrons. I also note that, with this simple instruments, it is also possible to foresee the existence (or the non-existence) of some diatomic molecules, such as He_2 or Li_2.

As for the LCAO method, in the case of heteronuclear molecules, besides what we have seen above, one should also consider whether to put a particular weight in front of the product wave function terms. An interesting example is given by the fluoride acid HF, where $\alpha = 0.5$, due to the fact that fluorine is very electronegative.

7.3.3 The Emergence of the Molecular World

The Notion of Emergence

I have already introduced the notion of emergence. This becomes particularly relevant when, starting by QM, we pass to higher levels of complexity of our physical world for which the use of approximation methods can become insufficient. In fact, as a conclusion of the previous examination, we can say that both the macroscopic level does not disturb quantum processes at their level and the details of the microscopic interactions (for instance, the atomic scale) become largely irrelevant for many questions of physical interest (for instance the behaviour of fluids), in particular, when we approach critical points. This happens when the number of relations and parameters involved in the description of the systems requires completely different methods relative to those applied to more basic physical situations. In the following, I shall introduce in a sketchy form of some of these aspects.

The notion of emergence has a robust logical categorical basis. As we shall see in the next chapter, from a category theory point of view, emergence deals with both restrictions in the number of the components of a system and increasing of relations among them, what has the consequence that these relations are equivalence classes of

lower level relations. In particular, operations at a higher level represent equivalence classes of operations or processes at a lower level. For instance, biological operations are equivalence classes of physical–chemical processes. Evidence for that is splicing (separation of exons (DNA codifying sequences) from introns (DNA non-codifying sequences)), which can be done at least in three different ways.[115]

In particular, I have proposed to deal with the problem of emergence on the basis of four basic ideas[116]:

(1) The different levels of emergence are characterised by distinct conservation laws: at any level, a particular quantity is conserved or optimised (in a normative sense). Although conservation laws always need to be satisfied, they can be considered as going from more general to more specific ones when passing from the less complex to the more complex, so that new conservation laws are established while the previous ones can be 'locally' violated at the emerging level, and this is precisely what defines the new class of operations for this level.
(2) The emergence process starts from some initial instability.
(3) The driving force of emergence is given by selection processes allowing canalisation along specific (emerging) paths thanks to constraints.[117] This fits well with the idea of a universal Darwinism.[118]
(4) New forms of stability are determined by new kinds of operations that are allowed by those constraints. So constraints are built both in due course and as a final result of the process.

In few words, this view could be called *emergent monism*.[119] Complexity is not easy to define. In general terms, it can be taken as resulting from a compromise between order and disorder in which the components of a system share among them some information but locally display a certain degree of autonomy and variance relative to each other.[120] In general, a complex system is hierarchically organised working as a combination of modularisation (and therefore relative local autonomy) and centralisation for some crucial functions or processes. This confers to the system a significant degree of plasticity (and adaptation). A paradigmatic example of the transition from the less complex to the more complex is represented by the evolution from prokaryotes (bacteria) to eukaryotes. In eukaryotes, the biological operations are much more controlled and the organism is simultaneously more modularised in relation to prokaryotes.

A general behaviour in Nature seems to be the following: since every break of symmetry or other kinds of disorder need to be compensated once a local selection event has already occurred (in order to keep the same amount of global entropy:

[115] Auletta et al. (2008).

[116] Auletta (2015b).

[117] See also Auletta (2011a, Sect. 2.4.2).

[118] As supported in Zurek (2004, 2009). This seems to have been also the point of view of Deutsch (1997), D'Espagnat (2011).

[119] See Auletta (2011c, Sect. 3.3.5) and literature quoted there.

[120] Auletta (2011a, Sect. 6.3) (Auletta 2011b).

Sects. 3.2.2–3.2.3) yet, on the other hand, it is impossible to restore exactly the same initial state (due to the presence of both quantum and classical correlations this would imply to ultimately restoring the same state of the universe), this confers an *itinerant dynamics* to the evolution of our universe. Even quantum systems, if perturbed, will rarely come back to the same state after a cycle (Sects. 5.1.3 and 7.1.5). This has in turn a relevant consequence for systems that are more complex: only *higher integration* can rebuild order when discontinuous critical events have already happened that endanger their 'survival', what represents a spontaneous drive to higher complexity. In fact, if a system, in its search for equilibrium, cannot rebuild its ordered initial state, under the pressure of selective mechanisms it is obliged to somehow integrate the perturbation or at least its consequences in a new form of stability, otherwise it will be still exposed to the danger of a breakdown. However, since it was previously unable to do so (otherwise, it would not have been perturbed), this implies a higher level of complexity in which order and disorder can be integrated in a stable (and new) way. It is true that abstractly speaking less complex systems could be more robust to some fluctuations than more complex ones. However, the argument here is not about the degree of robustness but about the necessity to add further constraints or protecting mechanisms for managing a specific perturbation once it has already occurred. Complexity is a *consequence* and not a condition of (cosmic) evolution. Summarising, this itinerant dynamics implies recurrent situations of instability at different levels that can be solved by reaching levels of further complexity thanks to additional constraints.[121]

Of course, many kinds and forms of emergence can be distinguished. However, there are what could be called the major emergences of our physical world[122]:

- The transition from QM to the macroscopic world,
- The transition from the physical macroscopic world to life,
- The transition from life to mind.

I shall deal only with the first point below and come back in sketchy form to the other two in the next chapter, due to their relevance to the epistemological issues raised by quantum theory.

Laws, Operations and Entropy

One should avoid a possible confusion: one issue is the *global conservation laws* and another is the issue of the *typical behaviour of a system* at a certain level (with its distinctive operations). If QM is right, the global amount of entropy is still conserved, and mesoscopic or macroscopic systems cannot modify this situation. Nevertheless, the typical behaviour of quantum and classical systems are totally different. Isolated or relatively isolated quantum systems are reversible and display a superposition of all possibilities, and therefore they conserve entropy. In other words, a significant

[121] Auletta (2011a, Sect. 8.2.7).

[122] I was inspired by Maynard Smith and Szathmáry (1995).

large quantum system reproduces with some approximation the same behaviour of the whole quantum universe. Quite opposite, an isolated and relatively large classical system does not. In other words, we have a new kind of allowed operations at this classical level: building and destroying (for example, chemical) structures. This is a pure physical dimension and therefore the function or meaning of these structures is of no relevance here: any structure can be built that satisfies the general requirement of energy conservation (and minimisation) in a portion of the environment being able to make simultaneous use of free energy that is locally liberated through the spontaneously occurring quantum selection events.

A possible worry is that the generation of a higher level of order contradicts the second law of thermodynamics. In fact, it is not so. When a system integrates more parts or factors or also more systems merge (as it is the case for the emergence of a classical world), the maximal entropy attainable becomes higher than the simple sum of the entropy of the parts or subsystems taken separately before the merging.[123] In the previous subsection, we have already considered this when passing from two interacting bodies to three. This allows for the spontaneous constitutions of new relations among those parts, and these relations imply an increase in order. In fact, every time we put things together some relations are spontaneously generated[124] filling part of the window that opens between the global entropy of the independent systems and the maximal entropy of the compound system.

Singular Limits

On this basis, we are trying here to describe as far as possible the way in which quantum systems give rise to a reality that is less based on the combination of possibilities and much more on determination of patterns and structures. In other words, what is commonly known as reality appears to be a construction but out of quantum systems.

Physically speaking, we can rephrase the difference between quantum and mesoscopic–macroscopic worlds as a problem of limits (Sect. 1.3.2). In fact, for many transitions from one domain to another singular limits are involved, that is, situations in which the transitions from the equations ruling one domain to another become abrupt and anomalous.[125] In other words, we call a limit *singular* if the behaviour of some constant or parameter when equal to zero differs from the behaviour of the same constant or parameter tending to zero. For instance, in the case in which we consider the transition from the quantum wave-like description of light and the classical geometrical (Newtonian) ray theory, we have a singular limit at $\lambda = 0$. Near that limit, a

[123] Landsberg (1984a, b). See also Auletta (2011a, Sect. 9.9).

[124] Peirce (1878, p. 310). Already Leibniz said: "Et si quelcun traçoit tout d'une suite une ligne qui seroit tantost droite, tantost cercle, tantost d'une autre nature, il est possible de trouver une notion ou regle ou equation commune à tous les points de cette ligne en vertu de la quelle ces même changemens doivent arriver" Leibniz (1686, p. 431).

[125] Berry (1994, 2001, 2002). See also Chibbaro et al. (2014, p. 33).

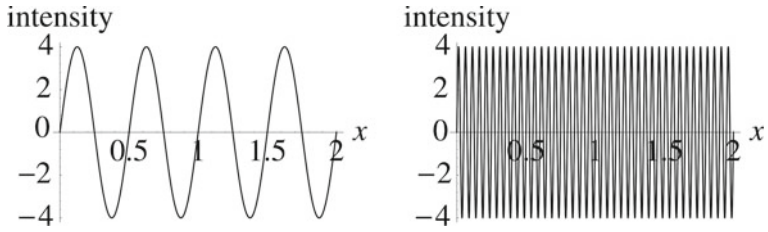


Fig. 7.33 We are assuming a photon moving along the x axis. Left panel: the quantum interference function for a wavelength $\lambda = 0.5$. Right panel: when $\lambda = 0.05$, that is for λ that tends to 0, we have the intensity of light attaining all values between 0 and 4. Adapted from Berry (2001)

faster and faster oscillation of the intensity of light (given by the square modulus of the wave function describing the photon) occurs, and in the limit, it attains all values between 0 and 4 in arbitrarily small ranges of x (Fig. 7.33).

Singular limits can be found also in other contexts and other levels of complexity. Indeed, further evidence for emergence 'out' of singular limits is phase transition (from a liquid to a solid state, for instance). Although phase transition is well defined in general terms (the phase that minimises the free energy is selected), the specific parameters describing a phase transition of a concrete chemical (like water) cannot be guessed a priori since experimental data are necessary, and therefore needs to be simulated.[126] We speak of first-order phase transitions when the equilibrium state of a given macroscopic system is not uniquely specified by the usual thermodynamic quantities, such as the temperature and the pressure.[127] In such a case, two different phases characterised by different properties, e.g. liquid and solid or solid and gas, may equally exist at a phase transition and typically, they coexist. Furthermore, these phase transitions are characterised by discontinuous changes of certain material properties, as in the case of ice turning liquid, whose density changes discontinuously.

In the so-called second-order phase transitions, such discontinuities do not occur, nevertheless, the overall behaviour of the object of interest changes drastically. These are related to the presence of critical points in the phase diagram for equilibrium states, for example the PT (pressure–temperature) plane. Note that, the difference between liquid and gas density (the order parameter here) vanishes when we approach a critical temperature, around which the liquid phase turns continuously into gas and the distinction between the phases fades away. Resuming with the words of Robert Laughlin, "One unambiguous signature of an organisational phenomenon is …sharp phase transition. The transition itself, however, is only a symptom. The important thing is not the transition but the emergent exactness that necessitates it".[128]

[126] Atkins and De Paula (2006, p. 177).

[127] Chibbaro et al. (2014, pp. 62–63).

[128] Laughlin (2005, p. 40).

Mononuclear and Heteronuclear Molecules

Let us first discuss the formation and geometry of single molecules and then what happens when larger populations are considered. What is extraordinary with the molecular world is that both emergent phenomena and general quantum-mechanical principles cooperate in surprising ways. The LCAO approximation helps us to understand that electrons accumulate in regions where atomic orbitals interfere constructively.[129] In the case of mononuclear molecules, recalling Eq. (7.292), we have bonding, for which we can write down the probability $\psi^2 = \psi_a^2 + \psi_b^2 + 2\psi_a\psi_b$ for two atoms a and b (where for simplicity I have used only real terms). These positively interfering regions (called internuclear) allow for orbital shrinkage that improves electron–proton interaction more than it is decreased by the migration to this region (Fig. 7.34). They are regions of lower energy. We have antibonding when $\psi^2 = \psi_a^2 + \psi_b^2 - 2\psi_a\psi_b$, an orbital that, when occupied, contributes to a reduction of the cohesion of the two atoms because the electrons are found everywhere but between the two nuclei, exposing the latter in this way and determining a repulsive reaction between the atoms.

When applied to the nitrogen molecule N_2 and taking z as the internuclear axis (connecting the two nuclei), we imagine each $2p_z$ orbital pointing towards a $2p_z$ orbital of the other atom. Here, a cylindrical symmetry is established along the internuclear axis (the so-called σ bond). The other $2p$ orbitals ($2p_x$ and $2p_y$), instead, establish a side-by-side bond through the nodal plane (a so-called π bond) (Fig. 7.35). Let us also consider oxygen, whose electron configuration is $1s^2 2s^2 2p_x^2 2p_y^1 2p_z^1$ (Fig. 7.28). When giving rise to water, the unpaired electrons in the O$2p$ orbitals can each be paired with an electron in H$1s$ orbital. Since the $2p_y$ and $2p_z$ orbitals theoretically lie at 90° to each other, the two constituted σ bonds also theoretically lie at 90° to each other (Fig. 7.36).

The case of heteronuclear molecules is more complicated. All heteronuclear diatomic molecules are polar, i.e. with a permanent electric dipole moment: this is the measure of the polarity of a system of electric charges (like electron and proton) with a displacement vector pointing from the negative charge to the positive charge. Molecular symmetry is of the greatest importance for the issue of molecular polarity, even more than the problem of whether or not the atoms constituting the molecules belong to the same element. Indeed, ozone is homonuclear but polar, since the central O atom is different from the other two (having two bonds instead of one) and the dipole moments are not cancelled[130] (Fig. 7.37). Interaction between molecules, especially considering electric dipoles and ions give rise to a large variety of multipoles.

The position of each atom in a molecule is determined by the nature of the chemical bonds by which it is connected to its neighbouring atoms.[131] The molecular geometry can be described by the positions of these atoms in space, evoking bond lengths of

[129] Clayden et al. (2001, pp. 95–105).

[130] Atkins and De Paula (2006, pp. 621–635).

[131] I recommend the excellent article https://en.wikipedia.org/wiki/Molecular_geometry.

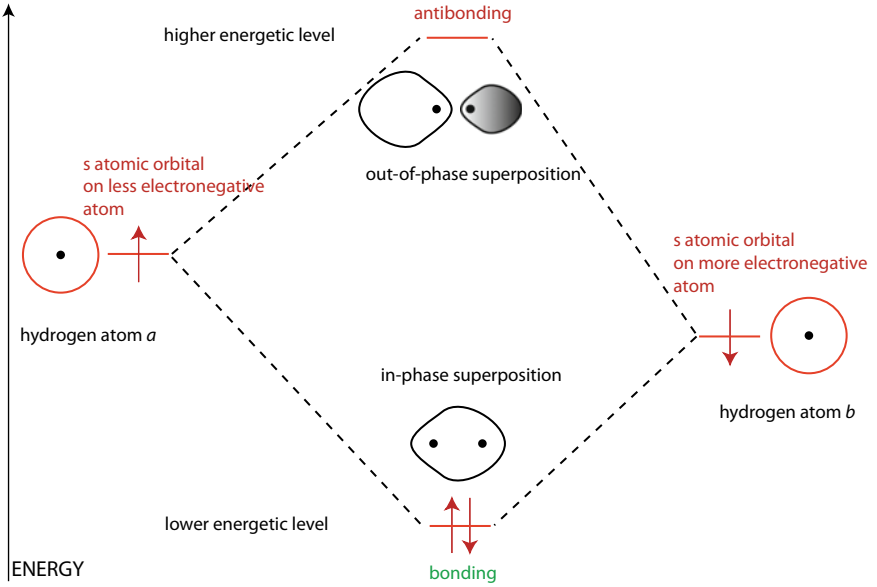


Fig. 7.34 The bonding and antibonding between two hydrogen atoms, *a* (on the left) and *b* (on the right). Shadowed and unshadowed regions mean position probabilities for electrons being in opposite phases. The elements like O or F are the more electronegative atoms which contribute more to the bonding orbital (the two opposite spins of the electrons are indicated with double vertical arrows below), while the atoms like C are the less electronegative ones, that contribute more to the antibonding orbital. The case in which the electrons are shared equally by the two atoms occupying the same intermediate energetic level, portrays a pure *covalent bond*. When the difference between the two electronegativities is too large, we find that a filled orbital on the anion (negatively charged ion) has the same energy level as the atomic orbital on one of the atoms and the empty orbital on the cation (positively charged ion) has the same energy level as the atomic orbital on the other atom. In this case an *ionic bond* is established (like between metals and non-metals). Here, an intermediate case is shown. Adapted from Auletta (2011a, p. 167)

two joined atoms, bond angles of three connected atoms, and torsion angles (dihedral angles) of three consecutive bonds. The bond length is defined to be the average distance between the nuclei of two atoms bonded together in any given molecule. A bond angle is the angle formed between three atoms across at least two bonds.

Some common shapes of simple molecules are the following ones. Linear: in a linear model, atoms are connected in a straight line. The bond angles are set at 180°. For example, carbon dioxide and nitric oxide have a linear molecular shape. Trigonal planar: molecules with the trigonal planar shape are somewhat triangular and in one plane (flat). Consequently, the bond angles are set at 120°. For example, boron trifluoride or ozone. Bent: bent or angular molecules have a non-linear shape. For example, water (H^2O), which has an angle of about 105°, what explains the particularity remarked in the caption of Fig. 7.36. Tetrahedral: this shape is found when there are four bonds all on one central atom, with no extra unshared electron

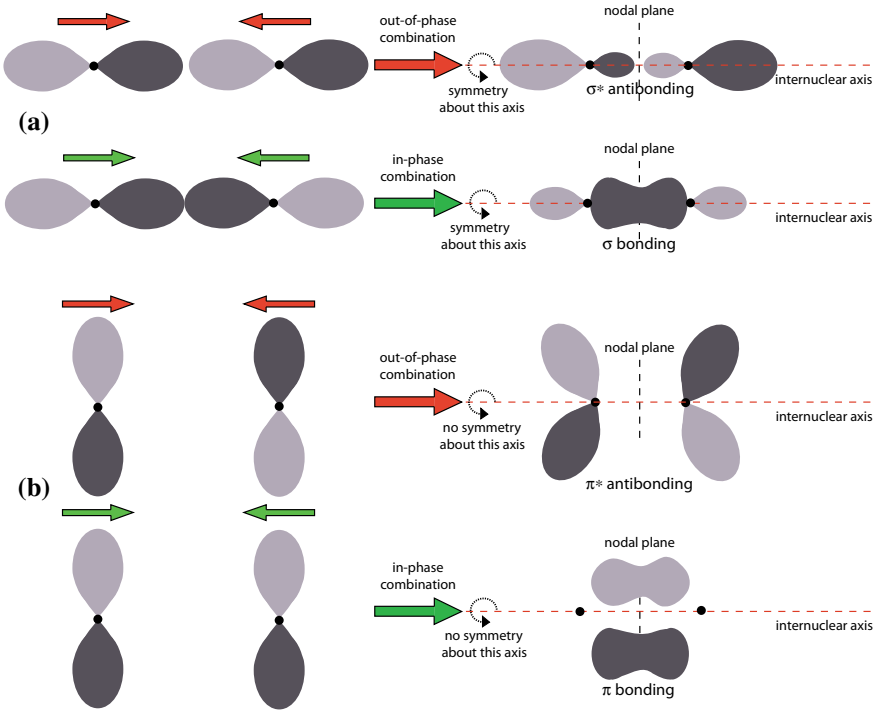


Fig. 7.35 **a** Constitution of antibonding (σ^*, first row) and bonding (σ, second row) combinations along the symmetrical axis (the two colours mean position probabilities of electrons being in different phases). **b** Constitution of antibonding (π^*, third row) and bonding (π, fourth row) combinations, which are orthogonal to the symmetrical axis (this means along the p_x or p_y axes). Adapted from Auletta (2011a, p. 168)

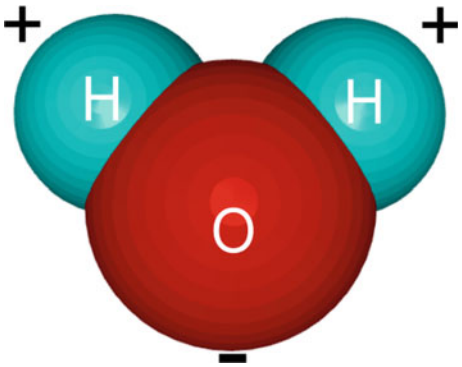


Fig. 7.36 Representation of the water molecule. Note that it is positively polarised on the "hydrogen side" and negatively on the opposite side. Molecular symmetry is the most important factor for having molecular polarity. According to the structure of the orbitals, we should expect that the hydrogen atoms make a 90° angle. In fact, the configuration is different

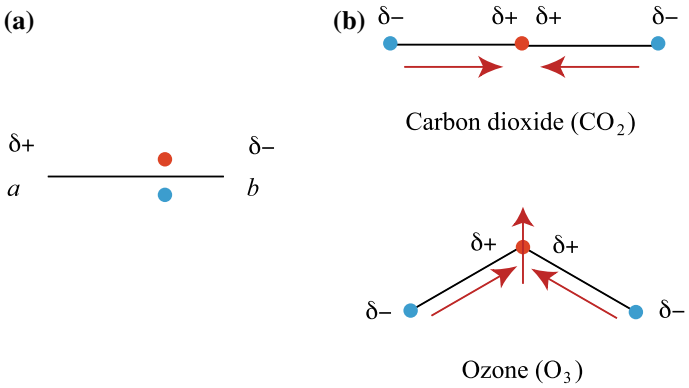


Fig. 7.37 **a** We say that an atom b is slightly more electronegative than an atom a, if the atom b will attract the electron pair (in red and blue) more than a does (Fig. 7.34). This means that the b end of the bond has more than its fair share of electron density and so becomes slightly negative. At the same time, the a end becomes slightly positive. The symbol δ_- means slightly negative while δ_+ means slightly positive. **b** An example of neutral molecule (carbon dioxide) and of a polar one (ozone). Adapted from Auletta (2011a, p. 1700)

pairs. In accordance with the VSEPR (valence–shell electron pair repulsion theory), here, the bond angles between the electron bonds are $\arccos(1/3) = 109.47°$. For example, methane (CH4) is a tetrahedral molecule. Octahedral: the bond angle is 90 degrees. For example, sulphur hexafluoride (SF6) is an octahedral molecule. Trigonal pyramidal: a trigonal pyramidal molecule has a pyramid-like shape with a triangular base.

Now, what is interesting is that such structures represent a new kind of order relatively to quantum-mechanical systems. First of all, let us remark that these geometric configurations have no quantum-mechanical significance, although they have originated from quantum-mechanical processes to which small perturbations (often due to the interaction with the environment) have contributed.[132] In such a way, this mechanics is basically an extension of decoherence. Moreover, quantum features are intrinsic self-correlations that display maximal order since they allow a combinatorics of any possible component of the state (Sect. 5.1.2). Here, at the opposite, we have a combinatorics of elements that pertain to different systems (as in any classical correlation) and are actually existent tokens (and not only possible). This means that such structures represent only *one* specific and typical combination of elements and thus necessarily instantiate a lower level of order than quantum-mechanical systems.

This geometric characteristic is due to a physical one: there is only *one way* to build any of the atoms and again *one way* to build molecules, although there is still some flexibility in certain regards: I recall here that in an ice crystal (where more oxygen and hydrogen atoms are considered) there are six possible different

[132]Claverie and Jona-Lasinio (1986). Note that the interaction between a molecular system and its environment is ruled by a non-linear Schrödinger equation.

arrangements of the hydrogen atoms relatively to the oxygen atoms, what denotes a certain amount of entropy. However, the fundamental factors (involved atoms, kind of bonds, number of electrons that are shared, electromagnetic characters, general geometric configuration of the molecule) contributing to the structure is always the same for each kind of atom or molecular configuration. I recall that molecules that share the same chemical formula but have different geometries are called *isomers*.

This is a crucial difference since it shows that such a combinatorics cannot represent a linear code. In fact, a linear code is composed of a set (of a finite number) of basic elements that can be combined in different ways producing different sequences (Sect. 3.2.3). In other words, a linear code is a code only because the codewords can give rise to different combinations, which implies a space of possibilities in advance. Here, at the opposite, we have only one possible combination. It is only when looking at the whole table of the elements that we can find that it is in fact a code, since we can understand electrons (and protons) as different instances of the same "characters" that, when appropriately combined, can give rise to different "sequences" (the elements, characterised by the atomic number, as displayed by Fig. 7.28). However, this is no longer true for single atoms and single molecules (although the same atom can contribute to very different molecules). Therefore, there is a crucial difference between the quantum-mechanical world and the world of atoms and molecules: the former instantiates linear codes, the later does not. We can still treat an atom (or a molecule) as instantiating a code (e.g. ground/excited state), but in such a case we are considering it as a pure quantum-mechanical system (that can be in a superposition state of these two possibilities) as for the experiment shown in Fig. 6.3, Sect. 6.2.1, and none of the particular characters of the atom (or the molecule) play a role here. Since such mesoscopic region is 'located' between QM and more complex systems that are characterised by important non-linear factors, we can understand this combinatorics without code as one of the distinctive marks of this part of physics (a kind of bridge to different regions of physics). Obviously, non-linear factors (as the mentioned perturbations for establishing the structures) are already important for the level we are considering now.[133] Nevertheless, they become dominant when several molecules are aggregated giving rise to, e.g. fluids, which are, for example characterised by different phenomena of turbulence.

It is very interesting to remark that geometry enters at many levels of our study: in the form of correlations, in the phase shifts, in the structure of space-time, in the structure of the environment, in the structure of molecules, in the different shapes of macroscopic bodies, in planetary systems and galaxies.

Classical Statistical Mechanics

Let us now consider large populations of molecules. Statistical mechanics is a fundamental discipline for considering what happens at the threshold between quantum-mechanical and classical processes. In fact, there are both a classical and a quantum-

[133] Wu (2005).

mechanical statistical mechanics. The important point is that the basic equations of classical and quantum statistical mechanics are the same. I shall start by the classical case for the sake of simplicity.

Let us consider an isolated system composed of two parts, one (Part 1) much smaller than the other (Part 2), whose (classical) Hamiltonians are $H_1(p_1, q_1)$ and $H_2(p_2, q_2)$ and with number of particles (molecules) N_1, N_2, respectively.[134] Thus, we assume that $N_2 \gg N_1$. Both systems are macroscopically large and here we are interested only in system 1. Consider a microcanonical ensemble of the composite system defined by having total energy between E and $E + \Delta E$, so that the energies E_1, E_2 of the two subsystems satisfy

$$E < (E_1 + E_2) < E + \Delta E, \tag{7.301}$$

where ΔE represents some increment of energy. Although this includes a range of values of E_1, E_2, it can be shown that only the set of mean values $\{\overline{E}_1, \overline{E}_2\}$ is important. Moreover, we assume that $\overline{E}_2 \gg \overline{E}_1$. If $W_2(E_2)$ be the volume occupied by system 2 in its own phase space, the probability to find system 1 in a state within the volume $dp_1 dq_1$ regardless of the state of system 2, is proportional to $dp_1 dq_1 W_2(E_2)$, where $E_2 = E - E_1$. Note that any volume element $dqdp$ of the phase space corresponds to $dqdp/N!$ states of the system (the correct Boltzmann counting), where N is the number of molecules of the system. Thus, the density in the phase space for system 1 is

$$\rho(p_1, q_1) \propto W_2(E - E_1). \tag{7.302}$$

Since only the values $\overline{E}_1$ are expected to be important, and $\overline{E}_1 \ll E$, in accordance with the expression (3.79) for Boltzmann entropy, we can perform the Taylor expansion (1.22) about E_2

$$\begin{aligned} k_{\mathrm{B}} \ln W_2(E - E_1) &= S_2(E) - E_1 \left[\frac{\partial S_2(E_2)}{\partial E_2}\right]_{E_2=E} + \cdots \\ &= S_2(E) - \frac{E_1}{T}, \end{aligned} \tag{7.303}$$

where I have suppressed the subscript B in the entropy symbol S for simplicity and T is the temperature of the system. Note that only the first two terms of the expansion have been considered and in the last step I have used the classical equation

$$\frac{\partial S}{\partial E} = \frac{1}{T}. \tag{7.304}$$

Therefore, taking the exponential of both sides of Eq. (7.303), we get

[134] Huang (1963, Sect. 7.1).

$$W_2(E - E_1) \approx e^{k_B^{-1} S_2(E)} e^{-\beta E_1}, \tag{7.305}$$

with $\beta = (k_B T)^{-1}$. Being the first exponential independent of E_1, it is a constant factor as far as system 1 is concerned. Since $E_1 = H_1(p_1, q_1)$, we take the ensemble density (7.302) for the small subsystem to be

$$\rho(p, q) = e^{-\beta H(p,q)}, \tag{7.306}$$

where the subscript denoting the subsystem has been omitted due to the extrapolation. The presence of the factor $\beta = (k_B T)^{-1}$ explains why $k_B T$ is a scale factor in Eq. (7.223).

The volume in the phase space occupied by the canonical ensemble is called the classical partition function:

$$Z(\beta) = \int \frac{d^{3N} p d^{3N} q}{N! \mathbf{h}^{3N}} e^{-\beta H(p,q)}, \tag{7.307}$$

where the function **h** has the dimension of momentum × distance.

Quantum Statistical Mechanics

The basic assumption is that we deal here with an ensemble of quantum systems that show phase randomness relative to each other in order to avoid probability interference.[135] In such a case, we deal with a mixed density matrix of the kind (1.365)

$$\hat{\tilde{\rho}} = |c_n|^2 \, |\psi_n\rangle \, \langle \psi_n| \, , \tag{7.308}$$

representing an incoherent superposition of eigenstates $|\psi_n\rangle$ of energy (Eq. (1.197)), so that the matrix elements are

$$\rho_{mn} = \left\langle \psi_n \left| \hat{\tilde{\rho}} \right| \psi_n \right\rangle = \delta_{mn} \, |c_n|^2 \, , \tag{7.309}$$

where

$$|c_n|^2 = \begin{cases} \text{Const.} & (E < E_n < E + \Delta E) \\ 0 & (\text{otherwise}) \end{cases} \tag{7.310}$$

where E_n are the eigenvalues of the system's Hamiltonian $\hat{H}$. Apart from the postulated phase randomness, quantum statistical mechanics assumes, as its classical counterpart, the equal a priori probability to get any of the eigenvalues of the energy. This is a crucial point: although we assume that there is no probability interference, nevertheless the single involved atoms or molecules (and therefore the single compo-

[135] Huang (1963, Chap. 8).

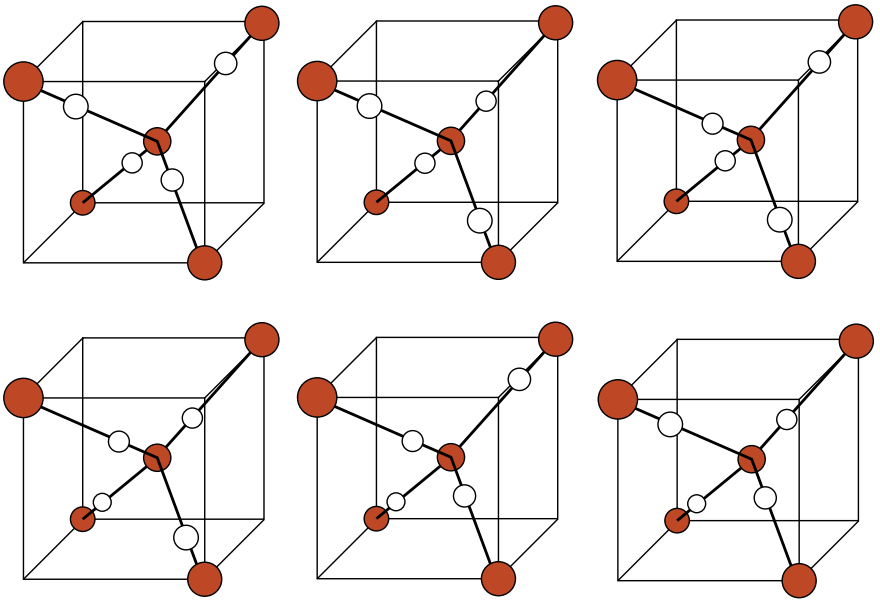


Fig. 7.38 The six possible locations of hydrogen atoms (white circles) relative to oxygen atoms (grey circles) in a ice crystal. Two of the hydrogen atoms must be near and too far away from the central oxygen atom. This incertitude about the location of the hydrogen atoms means that there is entropy even at absolute zero temperature. Moreover, since the six configurations are equiprobable, this is a maximal entropy state. Adapted from Auletta (2011a, p. 51)

nents of the mixed density matrix) still obey quantum-mechanical laws, what implies that we only have a probability that a particular element be in a specific energy eigenstate. This is in agreement with the fact that there is an irreducible indeterminacy in quantum statistical mechanics, whose consequence is that the (macroscopic) thermodynamic entropy, obeying the third law of thermodynamics, can be zero at absolute zero (because we assume that there is a single state in those conditions) while the *statistical entropy* is not.

In fact, I have already recalled the example of the ice crystal, showing that there can be alternative configurations even in the absence of thermodynamical processes (Fig. 7.38).[136] This proves that Boltzmann entropy deals with configurations (combinatorics) and is therefore a pure statistical–mechanical notion, and as a consequence, the notion of entropy makes physical sense only if we assume that the physical world is characterised by a *basic incertitude*, whose root must be in quantum-mechanical processes. This makes classical, and in particular statistical, mechanics and QM much closer than often assumed. Further evidence of this can be found in those physical fields, like non-linear fibre optics and plasma physics, where again probability plays a key role and analogues of the Schrödinger equation with an additional non-linear

[136] Atkins and De Paula (2006, pp. 609–610).

potential term are used.[137] Recently, a surprising application of these methods to the astronomical problem of the long-term evolution of self-gravitating disks of matter and gas forming at different scales (from galaxies to protoplanetary nebulae and circumplanetary rings) seems very promising.[138]

There is also a further aspect of the problem. In Sect. 2.4.2 I have recalled that a state of maximal entropy can be considered as symmetric, and there as well as in Sect. 6.2.5, I have stressed that the laws of Nature are grounded on symmetries. Thus, we can say that the second law expresses an expectation of what will be tendentially the case in the long-time run. In other words, the symmetry of a maximal entropic state is the ground of the statistical laws of statistical mechanics in the same sense in which the symmetries of QM (Sect. 1.2.5) are the basis of the probabilistic determinism of quantum laws. This confirms that the Nature's laws can coexist with disorder.

The further consequence is even that statistical mechanics is a fundamental science that cannot be reduced to some kind of classical deterministic theory, as still desired by Einstein (Sect. 3.3.4). Thus, although the postulates of quantum statistical mechanics may be considered as phenomenological statements, they are in fact more fundamental, and the reason is twofold:

- First, from the postulates of quantum statistical mechanics, we can not only derive the laws of thermodynamics but they lead to definite formulas for all the thermodynamic functions of a given system and
- Second, they are more directly related to molecular dynamics than are the laws of thermodynamics.

Coming back to the previous examination, the trace of the density matrix (7.308) is equal to the number of states whose energy lies between E and $E + \Delta E$:

$$\mathrm{Tr}\,\hat{\hat{\rho}} = \sum_n \rho_{nn} = W(E), \tag{7.311}$$

since for macroscopic systems, the spectrum $\{E_n\}$ almost forms a continuum. For $\Delta E \ll E$, we may take

$$W(E) = \omega(E)\Delta E, \tag{7.312}$$

where $\omega(E)$ is the density of states at energy E. This allows us to establish the identification

$$S = k_\mathrm{B} \ln W(E), \tag{7.313}$$

in full agreement with Eq. (3.79). Thus, in analogy with the classical case (7.305), we can write

$$\rho_{mn} = \delta_{mn} e^{-\beta E_n}. \tag{7.314}$$

[137] Agrawal (2013, Sect. 2.3), Bellan (2006, Sect. 15.5).

[138] Batygin (2018).

Then, the quantum partition function (which is the discontinuous counterpart of its classical analogue (7.307)) is given by

$$Z(\beta) = \mathrm{Tr}\left(\hat{\hat{\rho}}\right) = \sum_n e^{-\beta E_n}, \tag{7.315}$$

where the sum on the RHS is a sum over states and not on energy eigenvalues. Thus, $e^{-\beta E_n}$ represents the probability for the system to have the energy eigenvalue E_n. Since, according to the definition (7.308) and Eq. (7.314), we have

$$\hat{\hat{\rho}} = \sum_n e^{-\beta E_n} \left|\psi_n\right\rangle \left\langle\psi_n\right| = e^{-\beta \hat{H}} \sum_n \left|\psi_n\right\rangle \left\langle\psi_n\right| = e^{-\beta \hat{H}}, \tag{7.316}$$

we can reformulate the partition function as

$$Z(\beta) = \mathrm{Tr}\left(\hat{\hat{\rho}}\right) = \mathrm{Tr}\left(e^{-\beta \hat{H}}\right) . \tag{7.317}$$

This allows us to back express the energy and the entropy as

$$E = \langle \hat{H} \rangle = -\frac{\partial}{\partial \beta} \ln Z , \tag{7.318a}$$

$$S = k_\mathrm{B} \left(\ln Z - \beta \frac{\partial}{\partial \beta} \ln Z \right) , \tag{7.318b}$$

respectively.

Fluids

Here, we come back to the issue of critical limits discussed at the beginning of the present subsection but now for the case in which many molecules interact. Both the mesoscopic and the macroscopic level are dominated by structures (rooted in discontinuities) and fluxes (rooted in continuities), the former more by structures, the latter more by fluxes. I shall focus on fluids. The average relative locations of the particles in a liquid are expressed by the radial distribution function $g(r)$, defined in such a way that $g(r)r^2 dr$ gives the probability that a molecule will be found in the range dr at a distance r from a reference molecule. However, again in accordance with our understanding of emergence, these distributions can only be simulated by knowing the specific macroscopic characters of the fluid in question[139]. For instance, water is not constituted by $n\mathrm{H_2O}$ molecules but rather by dynamic aggregates $\mathrm{H_{2}}_n\mathrm{O}_n$, where n is often much larger than unity.[140] Again, general laws do not determine single results.

[139] Atkins and De Paula (2006, pp. 606–607).

[140] Earley (2006).

Thus, many macroscopic properties of, for example water (but this is true of many chemical compounds) cannot be predicted by observing the structure of a single water molecule (Fig. 7.36). In particular, the dynamical aspects of this and many other fluids (including turbulence phenomena) require treatments that go much further than the molecular analysis. In fact, fluids are characterised by continuous parameters (best described through fields) that are very different in comparison with the discontinuous properties still characterising atoms and single molecules.[141] At such a level, turbulence phenomena that already played a role in the formation of molecular structures, become even more important, without necessarily giving rise to self-organising processes (like the Bénard cells: see Fig. 5.2, Sect. 5.1.2). Up to the mid of the last century, the very complicated behaviour of a fluid was thought by most physicists (like the great Russian physicist Lev Landau (1908–1968)) to result from the superposition of many simple harmonic oscillations (see Fig. 6.3, Sect. 6.2.1). However, later on it was shown that it is not so.[142] This suggests us that when dealing with macroscopic objects or substances, combinatorics is no longer possible (as, to a certain extent, it was still the case for atoms and molecules).

A good test is represented by the difference between Brownian motion (at a molecular scale) and fluids (at a macroscopic scale).[143] The mesoscopic Brownian motion, after the name of the Scottish botanist and palaeobotanist Robert Brown (1773–1858), is random motion of particles (like pollen) suspended in a fluid due to their collisions with the molecules composing the fluid that move relatively quick.[144] The motion of a spherical particle of mass m and radius R in a liquid with viscosity η (which is a measure of the fluid's resistance to gradual deformation by shear stress or tensile stress that can be thought of as its "thickness") is ruled by the equation

$$\frac{d\mathbf{v}}{dt} + \frac{6\pi R\eta}{m}\mathbf{v} = 0, \tag{7.319}$$

where $md\mathbf{v}/dt$ is the frictional force. The previous equation, known as Stokes' law (after the name of G. G. Stokes), depicts the exponentially decaying behaviour

$$\mathbf{v}(t) = \mathbf{v}(0)e^{-\frac{t}{\tau}}, \text{ with } \tau = \frac{m}{6\pi R\eta}, \tag{7.320}$$

where $\mathbf{v}(0)$ is the velocity at initial time $t_0 = 0$ and $\mathbf{v}(t)$ is the subsequent velocity at time t. Both t and τ depend on the characters of the pollen and of the water. At the beginning of the twentieth century, A. Einstein and the Polish physicist Marian Smoluchowski (1872–1917), followed by the French physicist Paul Langevin (1872–1946), proposed to consider the motion of pollen as determined by two forces: (i) the deterministic viscous force obtained from Stokes law (7.319), (ii) a stochastic

[141] Chibbaro et al. (2014, Sect. 4.3). Fluids are best described by the Navier–Stokes equation, which presents non-linear terms.

[142] Chibbaro et al. (2014, p. 7).

[143] Chibbaro et al. (2014, pp. 58–61).

[144] Parisi (2005).

force due to the collisions with water molecules, which bears no memory of events occurring at different times. This implies that Eq. (7.319) need to be modifies as follows:

$$\frac{d\mathbf{v}}{dt} + \frac{6\pi R\eta}{m}\mathbf{v} = f_R(t), \tag{7.321}$$

where $f_R(t)$ is a random force representing the action of the water molecules on the pollen grains which, in agreement with kinetic theory, is more energetic at higher temperatures. Now, it turns out that the time scale of the impacts of the water molecules on the pollen grain is much smaller than the macroscopic viscous damping. Then, from these premises, Einstein and Smoluchowski derived the equation called after their names:

$$\langle \mathbf{x}^2(t)\rangle \approx 6Dt, \text{ with } D = \frac{k_B T}{6\pi R\eta}, \tag{7.322}$$

where $\mathbf{x}(t)$ is the displacement of a pollen grain at time t from its initial position, $\mathbf{x}(0)$, and T is the common temperature of water and pollen. The constant D is known as the diffusion coefficient.

We can now envisage two limiting situations involving the mass of pollen particles. The first is the limit of large mass of the pollen particles, which makes the effect of the molecular impacts (of water molecules) negligible, compared to that of the fluid viscosity. The second is that of small pollen mass, which makes the effect of the molecular impacts dominant, with respect to the viscous forces. These two situations seem to delimit all possibilities, ranging from suspended particles of the same size as the fluid molecules up to macroscopic objects floating in a liquid, which are insensitive to molecular impacts and only feel the viscosity of the fluid. Therefore, Brownian motion seems to lie at the border of the microscopic and the macroscopic worlds, and the crucial quantity ruling the transition from one level to another is the ratio of the mass of water molecules and of that of pollen grains. When it is zero (or anyway negligible), we have a singular limit. However, as long as this quantity does not vanish, Brownian motion takes place, although the observation time scales will increase with the mass of pollen.

7.4 Summary of the Chapter

- We have reviewed the classical and quantum electromagnetic theory and then introduced the general treatment of quantum-relativistic fields.
- The latter formalism is crucial for the standard model of particles. This, and especially recent research with the Higgs boson seems to suggest that everything can be explained in terms of fields, particles and carriers of the force.
- However, general relativity simply drops the notion of a gravitational field or force and treats gravitation in terms of the geometry of space-time.

- This has raised many different attempts at overcoming such a duality. I have proposed an explanation that is based on two basic assumptions: (i) the distinction between event and observed event, (ii) the distinction between laws and causal processes. As a consequence, I have proposed to consider quantum events as singularities generating the continuous space-time.
- Moreover, once that masses arise (likely through a mechanism of the kind of the Higgs boson), the effect of the masses is precisely a distortion of space-time as predicted by general relativity. The reason is that signals cannot be made reversible. This kind of solution takes the three fundamental theories involved here (QM, SR, GR) as describing each a different aspect of our physical world. Both entropic gravity and LQG go into this direction.
- The rest of the chapter is devoted to the emergence of the mesoscopic–macroscopic world. First, it has ben shown the considerable number of constraints that are necessary when building the hydrogen atom. Then, the combinatorics of elements has been sketched.
- Once we enter the world of larger atoms and the molecular world, approximation methods are the only available ones due to the problem of the three bodies. In particular, the LCA and VBM methods have been presented.
- Moreover, from the molecular level onwards the notion of emergence becomes crucial. This concept can be understood by taking into account the selection mechanisms at work and the typical behaviour of the emergent level that needs to represent an equivalence class of lower level behaviours and operations. In particular, the three major transitions of our physical universe have been sketched.
- When dealing with molecules, geometric considerations becomes relevant. Moreover, phase transition and formation of macroscopic stuff like liquids cannot be predicted on the basis of QM only and are true emergent phenomena. The problem of the limits has been also discussed.
- Statistical mechanics assumes a basic incertitude in our world and in this sense, it perfectly agrees with QM, although it deals with systems displaying positive entropy.

References

Agrawal, Govind P. 2013. *Nonlinear Fiber Optics*, 5th ed. San Diego: Academic.

Atkins, P.W., and Friedman, R. 2005. *Molecular Quantum Mechanics*, 4th ed. Oxford: Oxford University Press.

Atkins, P.W., and De Paula. 2006. *Journal of Physical Chemistry*. Oxford: Oxford University Press, 2002; 2nd ed. 2006.

Auletta, Gennaro. 2011a. *Cognitive Biology: Dealing with Information from Bacteria to Minds*. Oxford: Oxford University Press.

Auletta, Gennaro. 2011b. A Mathematical Model of Complexity and Its Application to Chemotaxis. *British Journal of Mathematics and Computer Science* 1: 204–227. http://www.sciencedomain.org/issue.php?iid=77&id=6.

Auletta, Gennaro. 2011c. (in collaboration with I. Colagè, P. D'ambrosio, and L. Torcal) *Integrated Cognitive Strategies in a Changing World*. Rome: G and B Press.
Auletta, Gennaro. 2015a. Brevi note sul Collegio Romano e il suo contributo alla fisica. In *Cristoforo Clavio e la cultura scientifica del suo tempo*, ed. Paola Vasconi, 103–111. Roma: Gangemi editore.
Auletta, Gennaro. 2015b. Emergence: Selection, Allowed Operations, and Conserved Quantities. *South-African Journal of Philosophy* 34: 93–105. https://doi.org/10.1080/02580136.2014.992156.
Auletta, G., and Wang, S.-Y. 2014. *Quantum Mechanics for Thinkers*. Singapore: Pan Stanford Pub. https://doi.org/10.4032/9789814411721.
Auletta, G., Ellis, G., and Jaeger, L. 2008. 1Top-Down Causation by Information Control: From a Philosophical Problem to a Scientific Research Program. *Journal of the Royal Society: Interface*. 5: 1159–1172. https://doi.org/10.1098/rsif.2008.0018.
Auletta, G., Fortunato, M., and Parisi, G. 2009. *Quantum Mechanics*. Cambridge: Cambridge University Press; rev. paperback ed. 2013.
Barrow-Green, June. 1997. *Poincaré and the Three Body Problem*. American Mathematical Society
Batterman, Robert W. 2002. *The Devil in the Details: Asymptotic Reasoning in Explanation, Reduction, and Emergence*. Oxford: Oxford University Press.
Batygin, Konstantin. 2018. Schrödinger Evolution of Self–Gravitating Disks. *Monthly Notices of the Royal Astronomical Society* 4: 5070–5084. https://doi.org/10.1093/mnras/sty162.
Bedau, M.A., and Humphreys, P. (eds.). 2008. *Emergence*. Cambridge: MIT Press.
Bekenstein, Jacob D. 1973. Black Holes and Entropy. *Physical Review* D7: 2333–2346.
Bekenstein, Jacob D. 2003. Information in the Holographic Universe. *Scientific American* 289 (2): 58–65. https://doi.org/10.1038/scientificamerican0803-58.
Bellan, Paul M. 2006. *Fundamentals of Plasma Physics*. Cambridge: Cambridge University Press.
Berry, Michael V. 1994. Asymptotics, Singularities and the Reduction of Theories. *Logic, Methodology and Philosophy of Science* 9: 597–607.
Berry, Michael V. 2001. Chaos and the Semiclassical Limit of Quantum Mechanics (Is the Moon There When Somebody Looks?), in [Russell et al. 2001, 41–54].
Berry, Michael V. 2002. Singular Limits. *Physics Today* 5510–5511.
Bjorken, J.D., and Drell, S.D. 1964. *Relativistic Quantum Mechanics*. New York: McGraw-Hill.
Bohr, Niels. 1949. Discussion With Einstein on Epistemological Problems in Atomic Physics, in [Schilpp 1949, 201–241].
Boscovich, Rogerius J. 1745. *De viribus vivis*. Roma: Collegio Romano.
Boscovich, Rogerius J. 1763. *Theoria philosophiæ naturalis redacta ad unicam legem virium in natura existentium*. Vienna, 1758; Venezia, 1763.
Brouwer, M. M., Visser, M. R., Dvornik, A., Hoekstra, H,, Kuijken, K., Valentijn, E. A., Bilicki, M., Blake, C., Brough, S., Buddelmeijer, H., Erben, T., Heymans, C., Hildebrandt, H., Holwerda, B. W., Hopkins, A. M., Klaes, D., Liske, J., Loveday, J., McFarland, J., Nakajima, R., Sifón, C., and Taylor, E. N. 2017. First Test of Verlinde's Theory of Emergent Gravity Using Weak Gravitational Lensing Measurements. *Monthly Notices of the Royal Astronomical Society* 466: 2547–2559. https://doi.org/10.1093/mnras/stw3192.
Brown, Harvey R. 2005. *Physical Relativity: Space–Time Structure from a Dynamical Perspective*. Oxford: Oxford University Press.
Byron Jr., F.W., and R.W. Fuller. 1969–70. *Mathematics of Classical and Quantum Physics, 1969–1970*. New York: Dover.
Chibbaro, S., Rondoni, L., and Vulpiani, A. 2014. *Reductionism, Emergence and Levels of Reality: The Importance of Being Borderline*. Heidelberg: Springer. https://doi.org/10.1007/978-3-319-06361-4
Claverie, P., and Jona-Lasinio, G. 1986. Instability of tunneling and the concept of molecular structure in quantum mechanics: The case of pyramidal molecules and the enantiomer problem. *Physical Review* A33: 2245–2253.
Clayden, J., Greeves, N., Warren, S., Wothers, P. 2008. *Organic Chemistry*, vol. 2001. Oxford: Oxford University Press.

Clayton, P., and Davies, P.C.W. (eds.). 2006. *The Re-Emergence of Emergence*. Oxford: Oxford University Press.
Cohen-Tannoudji, Gilles. 1991. *Les constantes universelles*. Paris: Hachette.
Cottingham, W.N., and Greenwood, D.A. 2007. *An Introduction to the Standard Model of Particle Physics*, 2nd ed. Cambridge: Cambridge University Press.
Cramer, John G. 1986. The Transactional Interpretation of Quantum Mechanics. *Review of Modern Physics* 58: 647–687.
D'Ariano, G.M., and Perinotti, P. 2014. Derivation of the Dirac Equation from Principles of Information Processing. *Physical Review* A90: 062106. https://doi.org/10.1103/PhysRevA.90.062106.
D'Ariano, M., Chiribella, G., and Perinotti, P. 2017. *Quantum Theory from First Principles: An informational approach*. Cambridge: Cambridge University Press.
D'Espagnat, Bernard (ed.). 2011. *The Beginning of Infinity: Explanations that Transform the World*. London: Penguin Books.
Deutsch, David. 1997. *The Fabric of Reality: The Science of Parallel Universes and its Implications*. London: Penguin Books.
DeWitt, Bryce S. 1967. Quantum Theory of Gravity. I. The Canonical Theory. *Physical Review* 160: 1113–1148. https://doi.org/10.1103/PhysRev.160.1113.
Dirac, Paul A.M. 1927. The Quantum Theory of the Emission and Absorption of Radiation. *Proceedings of the Royal Society of London* A114: 243–265. https://doi.org/10.1098/rspa.1927.0039.
Dirac, Paul A.M. 1928a. The Quantum Theory of the Electron I. *Proceedings of the Royal Society of London* A117: 610–624.
Dirac, Paul A.M. 1928b. The Quantum Theory of the Electron II. *Proceedings of the Royal Society of London* A118: 351–361.
Dirac, Paul A.M. 1930a. The Proton. *Nature* 126: 605–606.
Dirac, Paul A.M. 1930b. A Theory of Electrons and Protons. *Proceedings of the Royal Society of London* 126: 360–365.
Earley, Joseph E. 2006. Chemical "Substances" That Are Not "Chemical Substances". *Philosophy of Science* 73: 841–852.
Poincaré, Henri. 1928. *OEuvres*. Paris: Gauthier–Villars.
Einstein, Albert. 1949. Autobiographisches, in [Schilpp 1949, 2–94].
Elitzur, A.C., and Dolev, S. 2005. Quantum Phenomena Within a New Theory of Time, in [Elitzur et al. 2005, 325–349]
Elitzur, A.C., S. Dolev, and N. Kolenda (eds.). 2005. *Quo Vadis Quantum Mechanics?* Berlin: Springer.
Feynman, Richard P. 1949. The Theory of Positrons. *Physical Review* 76: 749–59.
Friedman, Michael. 1983. *Foundations of Space–Time Theories. Relativistic Physics and Philosophy of Science*. Princeton: Princeton University Press.
Fuchs, J., and C. Schweigert. 1997. *Symmetries, lie algebras and representations*. Cambridge: Cambridge University Press.
Gell-Mann, Murray. 1964. A Schematic Model of Baryons and Mesons. *Physics Letters* 8214–8215. https://doi.org/10.1016/S0031-9163(64)92001-3.
Geroch, Robert. 1978. *General Relativity from A to B*. Chicago: University of Chicago Press.
Greiner, W., Schramm. S., and Stein, E. 1994. *Quantum Chromodynamics*. Berlin, Springer, 1994; 2nd ed. 2002; 3rd ed. 2007.
Halliwell, Jonathan J. 2010. Macroscopic Superpositions, Decoherent Histories, and the Emergence of Hydrodynamic Behavior. [Saunders et al. 2010, 99–117].
Halvorson, H., and Clifton, R. 2002. No Place for Particles in Relativistic Quantum Theories? *Philosophy of Science* 69: 1–28.
Hempel, Carl G. 1965. *Aspects of Scientific Explanation and other Essays in the Philosophy of Science*. New York: Free Press.
Hempel, C.G., and Oppenheim, P. 1965. On the Idea of Emergence, in [Hempel 1965, 61–67].
Higgs, Peter. 1964. Broken Symmetries and the Masses of Gauge Bosons. *Physical Review Letters* 13508–13509.

Huang, Kerson. 1963. *Statistical Mechanics*. New York: Wiley, 1963; 2nd ed. 1987.
Jackson, John D. 1962. *Classical Electrodynamics*. New York: Wiley; 2nd ed. 1975; 3rd ed. 1998.
Jordan, P., and Klein, O. 1927. Zum Mehrkörperproblem der Quantentheorie. *Zeitschrift für Physik* 45: 751–756.
Jordan, P., and Wigner, E.P. 1928. Über das Paulische Äquivalenzverbot. *Zeitschrift für Physik* 47: 631–651.
Kastner, Ruth E. 2013. *The Transactional Interpretation of Quantum Mechanics: The Reality of Possibility*. Cambridge: Cambridge University Press.
Kiefer, Claus. 2007. *Quantum Gravity*. Oxford: Oxford University Press.
Kircher, Athanasius. 1643. *Magnes sive De arte magnetica Opus tripartitum*. Kalcoven: Cologne.
Kircher, Athanasius. 1667. *Magneticum Naturæ Regnum sive Disceptatio physiologica*. Rome, de Lazaris.
Klauber, Robert D. 2013. *Student Friendly Quantum Field Theory: Basic Principles and Quantum Electrodynamics*. Sandtrove Press: Fairfield (Iowa).
Landsberg, Peter T. 1984a. Is Equilibrium Always an Entropy Maximum? *Journal of Statistical Physics* 35: 159–69.
Landsberg, Peter T. 1984b. Can Entropy and 'Order' Increase Together? *Physics Letters* 102A: 171–173.
Laughlin, Robert B. 2005. *A Different Universe (Reinventing Physics from the Bottom Down)*. New York: Basic Books.
Leibniz, Gottfried W. 1686. *Discours de Metaphysique*, in [Leibniz PS, IV, 427–463].
Leibniz, Gottfried W. 1875. In *Philosophische Schriften*, ed. Gerhardt, Halle, 1875; rep. Hildesheim, Olms, 1978.
Lemaître, Georges. 1946. L'Hypothèse de l'atome primitif1946; eng. tr. *The Primeval Atom: An Essay on Cosmogony*. D. Van Nostrand Co, 1950.
Lewes, George H. 1875. *Problems of Life and Mind, First Series: The Foundations of a Creed*. London: Trübner.
Lloyd Morgan. 1923. Conwy. *Emergent Evolution*. London: Williams and Norgate, 1923, 1927.
Ma, X.-S., Kofler, J., Qarry, A., Tetik, N., Scheidl, T., Ursin, R., Ramelow, S., Herbst, T., Ratschbacher, L., Fedrizzi, A., Jennewein, T., and Zeilinger, A. 2013. Quantum Erasure with Causally Disconnected Choice. *Proceedings of the National Academy of Sciences USA* 110: 1221–1226. https://doi.org/10.1073/pnas.1213201110.
Malament, David B. 1977. The Class of Continuous Timelike Curves Determines the Topology of Spacetime. *Journal of Mathematical Physics* 18: 1399–1404. https://doi.org/10.1063/1.523436.
Malament, David B. 2012. *Topics in the Foundations of General Relativity and Newtonian Gravitation Theory*. Chicago: Chicago University Press.
Maxwell, James C. 1873. *A Treatise on Electricity and Magnetism*. Oxford: Clarendon Press.
Maynard Smith, J., and Szathmáry, E. 1995. *The Major Transitions in Evolution*. Oxford: Freeman/Spektrum, 1995; Oxford: Oxford University Press, 1997, 2002.
Mill, John Stuart. 1843. *System of Logic Ratiocinative and Inductive: Being a Connected View of the Principles of Evidence and the Methods of Scientific Investigation* 1st ed. London: J. Parker 1843; 8th ed. London: Longmans, Green, and Co., 1919.
Misner, C.W., Thorne, K.S., and Wheeler, J.A. 1970. *Gravitation*. San Francisco: Freeman and Co., 1973.
Mitchell, Sandra D. 2009. *Unsimple Truths: Science, Complexity, and Policy*. Chicago: The University of Chicago Press.
Parisi, Giorgio. 2005. Brownian Motion. *Nature* 433: 221.
Pauli, Wolfgang. 1927. Zur Quantenmechanik des magnetischen Elektrons. *Zeitschrift für Physik* 43: 601–623.
Peirce, Charles S. 1878. The Order of Nature. *Popular Science Monthly* 13: 203–217; in [Peirce W, III, 306–322].
Penrose, Roger. 1971. Angular momentum: An Approach to Combinatorial Spacetime. In *Quantum Theory and Beyond*, ed. Ted Bastin. Cambridge: Cambridge University Press.

Penrose, Roger. 2004. *The Road to Reality: A Complete Guide to the Laws of the Universe*. London: J. Cape; London, Vintage.
Poincaré, Henri. 1890. Sur le problème des trois corps et les équations de la dynamique. *Acta Mathematica* 131–270; rep. in [Poincaré O, VII, 262–490].
Poincaré, Henri. 1893. Le mécanisme et l'expérience. *Revue Générale de Métaphysique et de Morale* 1: 534–537; rep. in [Poincaré 1991, 87–91].
Poincaré, Henri. 1897. Les idées de Hertz surla mécanique. *Revue Générale des Sciences Pures et Appliqués* 8: 734–743; rep. in [Poincaré 1991, 62–86].
Poincaré, Henri. 1902. *La science et l'hypothèse, Paris, 1968*. Flammarion.
Poincaré, Henri. 1991. *L'analyse et la recherche*. Paris: Hermann.
Rindler, Wolfgang. 2001. *Relativity: Special, General, and Cosmological*. Oxford: Oxford University Press.
Rovelli, Carlo. 2002. Partial Observables. *Physical Review* D65: 124013. https://doi.org/10.1103/PhysRevD.65.124013.
Rovelli, Carlo. 2004. *Quantum Gravity*. Cambridge: Cambridge University Press.
Russell, R.J., P. Clayton, K. Wegter–McNelly, and J. Polkinghorne (eds.). 2001. *Quantum Mechanics: Scientific Perspectives on Divine Action*. Vatican Observatory.
Saunders, S., J. Barrett, A. Kent, and D. Wallace (eds.). 2010. *Many Worlds?* Oxford: Oxford University Press.
Schilpp, Arthur (ed.). 1949. *Albert Einstein. Philosopher-Scientist*. La Salle (Illinois): Open Court, 1949; 3rd ed. 1988.
Schwartz, Matthew D. 2014. *Quantum Field Theory and the Standard Model*. Cambridge: Cambridge University Press.
Serjeant, Stephen. 2010. *Observational Cosmology*. Cambridge: Cambridge University Press.
Susskind, Leonard. 1995. The World as a Hologram. *Journal of Mathematical Physics* 36: 6377–6396. https://doi.org/10.1063/1.531249.
Susskind, Leonard. 1997. Black Holes and the Information Paradox. *Scientific American* 276 (4): 52–57.
Synge, John. 1960. *Light on Relativity: The General Theory*. Amsterdam: North–Hollande.
't Hooft, Gerard. 1993. Dimensional Reduction in Quantum Gravity. arXiv:gr-qc/9310026.
't Hooft, Gerard. 2016. *The Cellular Automaton Interpretation of Quantum Mechanics*. Springer.
Tegmark, M., Aguirre, A., Rees, M.J., and Wilczek, F. 2006. Dimensionless Constants, Cosmology, and Other Dark Matters. *Physical Review* D73: 023–505.
Verlinde, Erik. 2011. On the Origin of Gravity and the Laws of Newton. arXiv:1001.0785v1
Watson, Andrew. 2004. *The Quantum Quark*. Cambridge: Cambridge University Press.
Wheeler, J.A., and Feynman, R.P. 1945. Interaction with Absorber as the Mechanism of Radiation. *Review of Modern Physics* 17: 157–181.
Wheeler, J.A., and Feynman, R.P. 1949. Classical Electrodynamics in Terms of Direct Interparticle Action. *Review of Modern Physics* 21: 425–433.
Wilczek, Frank. 2006. *Fantastic Realities: 49 Mind Journeys and a Trip to Stockholm*. Singapore: World Scientific.
Wu, Guozhen. 2005. *Nonlinearity and Chaos in Molecular Vibrations*. Amsterdam: Elsevier.
Zurek, Wojciech H. 2004. Quantum Darwinism and Envariance 121–137; Barrow-et al. 2004.
Zurek, Wojciech H. 2009. Quantum Darwinism. *Nature Physics* 5: 181–188. https://doi.org/10.1038/nphys1202.
Zwiebach, Barton. 2004. *A First Course in String Theory*. Cambridge: Cambridge University Press.

Chapter 8
Category Theory and Quantum Mechanics



In the present chapter, we shall deal with the general logical–epistemological foundations of the quantum theory, where the stress is on categorisation. This aspect and the physical–ontological ones previously discussed need finally to agree. I first give a general account of category theory. Then, we shall see its applications to QM and especially to quantum information. Then a logical and epistemological assessment of the theory follows.

8.1 Category Theory

8.1.1 Introduction

Category theory represents the highest level attained by humanity as far in unifying mathematics broadly understood in a single formal theory. Already for this reason, it presents great interest for both the scientist and the philosopher. In fact, category theory allows a powerful connection with both logic and computation, as displayed in Fig. 8.1. This connection has been known since the 1970s, and is widely used in computer science—it is also beginning to be used in quantum informatics.[1] However, as a consequence of this unification power, what is really amazing with category theory is rather its application range, going from information technology to physics, from biology to neuroscience. Indeed, everything is finally a problem of categorisation! And, as we shall see, any categorisation problem is ultimately a problem of logic.[2]

[1] Abramsky and Tzevelekos (2011).

[2] An insight that was already of Aristotle but forgotten by many: see Aristotle Cat. (2019c). See also Auletta (2013d).

© Springer Nature Switzerland AG 2019

G. Auletta, *The Quantum Mechanics Conundrum*,
https://doi.org/10.1007/978-3-030-16649-6_8

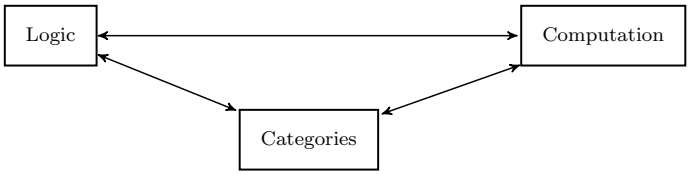


Fig. 8.1 Connection categories–logic–computation. Adapted from Abramsky and Tzevelekos (2011)

Unification of mathematics was already the dream of the German mathematician David Hilbert and logician Gottlob Frege (1848–1925).[3] This attempt was based on axiomatisation, i.e. in setting basic axioms from which all mathematical truths could be derived. Hilbert's formal axiomatic system is composed of alphabet, grammar, axioms, rules of inference and proof-checking algorithms. I note that what is problematic of this list is precisely the notion of axiom. Such a weak point determined a failure of this research program at that time. The main difference between such unification program and the current research based on category theory is twofold: the current success of category theory lies (1) in the renouncement to a deductive–foundational approach as such and (2) in the adoption of the concept of class. The issues (1) and (2) are interrelated. In fact, Frege and Hilbert tried to cast mathematics in an axiomatic framework in which all mathematical truths could be deduced from logic and in particular from *set theory*.[4] This seems quite natural since, according to a result of the American philosopher Patrick Suppes (1922–2014), to axiomatise a theory is to place it within set theory.[5] At the opposite, as already recalled in the Introduction to the present book, the approach of category theory, although very formal, is more pragmatic and focuses rather on the *results* (or consequences) than on the premises of reasoning. As pointed out by the British mathematician Tom Leinster, instead of being centred on definitions of basic notions, category theory is more focused on *what we can do* with its objects, or, more in general, what are the consequences in introducing those objects with those characters. Moreover, instead of insisting on the notion of true statement as it was typical of the previous foundational issues, Leinster points out that the issue at stake is *how convenient* is an assumption.[6]

Thus, while the previous axiomatisation (whose final result is the so-called Zermelo–Fraenkel theory with choice, after the names of the German logician and mathematician Ernst F.F. Zermelo (1871–1953) and the German–Israeli mathematician Abraham Halevi H. Fraenkel (1891–1965)) was still based on set theory and thus considered everything to be a set and insisted on the notion of set membership,[7] category theory relies on the more general notions of *class* and category, which are

[3] See Chaitin (1998, Sect. 1.1).

[4] Frege (1884), Frege (1893), Hilbert (1903).

[5] Suppes (1960). Newton da Costa and the Chilean logician Rolando Chuaqi split this requirement into two distinct sets of axioms (Chaitin et al. 2011, pp. 68–69).

[6] Leinster (2014, pp. 71–73).

[7] Zermelo (1908), Zermelo (1930), Fraenkel et al. (1958).

far richer than that of set and more suitable for being used in different contexts.[8] A class is a generalisation of set since its members can be sets: the common view is that a class is a collection of sets defined by a formula whose quantifiers range only over sets. Thus, if X and Y are sets, $X \cup Y$ (the set union of X and Y) is a class (whose members are X and Y). This shows that, while one can build arbitrary sets by using a single property (like red, number, abstract, and so on), and thus basic sets are defined by such a single property, classes put together different properties generating interconnections among those properties. What is crucial is that, while sets, at least when basically understood, represent disordered collections of objects, classes can be understood as collections of objects displaying an internal structure. Thus, also objects like $X \cup (Y \cap Z)$ are classes. I recall that, here, the notion of "object" has pure formal connotations. This allows us to treat also subsets like $X \cup Y$ as (improper) classes: while all sets are classes, not all classes are sets. In fact, we can interpret the previous expression as telling us either that it is about individuals being members of both set X and set Y or that the sets X and Y have a positive intersection, what is of course a relation between sets.

8.1.2 Preliminary Notions: Sets and Maps

The objects of a category are connected through morphisms, also called maps, that need to satisfy some requirements, while different categories can be connected by *functors*. These functors provide powerful communication of ideas in the sense that facts and theorems proven in one category can be transferred through a connecting functor to yield proofs of analogous theorems in another category. A functor is like a conductor of mathematical truth.

Sets and Functions

Let us first establish some basic and abstract notions.[9] One of the most basic concepts is that of *set*. We can think of a set X as a collection of things $x \in X$, each of which is recognisable as being in X and such that for each pair of named elements $x, x' \in X$, we can tell if $x = x'$ or not. I recall that the elements of a *set* are not (or need not to be) related to one another in any way, i.e. do not show an (algebraic) structure. Although sets have this constitution, they can of course display interesting relations *among* them.

Very elementary examples of sets are the set with no elements, the empty set $\varnothing = \{\}$ and the set with just one element, $\{1\}$, symbolised by "1". It can be proved that for any set X, there exists a unique map from X to 1, symbolised by $X \longrightarrow 1$,

[8] Leinster (2014, pp. 79–80).

[9] For the subject of the present section, I shall mainly rely on two recent textbooks: (Spivak 2013; Leinster 2014) The former is more approachable by non-mathematicians (since it also shows some applications), the latter is more abstract and formal.

because each element $x \in X$ maps to the single element of 1: $\forall X, \forall x \in X, x \longmapsto 1$. Note that the symbol $\longmapsto$ denotes a mapping between elements while the symbol $\longrightarrow$ mapping between collections of elements.[10] Of course, there are also sets with more elements. Let us define a set $\underline{n} := \{1, 2, 3, \ldots n\}$ for any natural number n. We denote by $|X| = n$ a set X that has cardinality n.

All functions are also maps, and the function f from X to Y is denoted $f : X \longrightarrow Y$. X is called the domain of the function f while Y is called the codomain. All elements of Y that have at least one arrow pointing to them are said to be in the *image* of f:

$$\text{Im}(f) := \{y \in Y | \exists x \in X \quad \text{such that} \quad f(x) = y\}, \tag{8.1}$$

where the symbol $\exists x$ denotes that "there is at least one element x such that …"

For any set X, we define the identity map on X, denoted by $1_X : X \longrightarrow X$, as the function such that $\forall x \in X$ we have $1_X(x) = x$. Given two functions, $f : X \longrightarrow Y$ and $g : Y \longrightarrow Z$ (where it may be noted that the codomain of f is the domain of g), we say that they are composable and write $X \overset{f}{\to} Y \overset{g}{\to} Z$, denoting the *composition* of these two maps as

$$g \circ f : X \longrightarrow Z, \tag{8.2}$$

where I stress that sequences of maps follow the same convention used for operations. Composition of maps can be represented as

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ & \searrow^{g\circ f} & \downarrow g \\ & & Z \end{array} \tag{8.3}$$

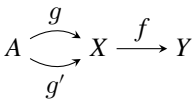
As said, the maps are called *morphisms* (although I shall occasionally use the term "map" for simplicity) and usually we denote the set of all morphisms $X \longrightarrow Y$ between sets X, Y by $\text{Hom}_{\mathbf{Set}}(X, Y)$: this is called the *hom-set*. Thus, functions in set theory are a particular kind of morphism.

Some Basic Morphisms

Let us explore now some very general mathematical notions that will be helpful for the following. Let $f : X \longrightarrow Y$ be a function. We say that f is *surjective* if, $\forall y \in Y$ there exists some $x \in X$ such that $f(x) = y$. I recall (see Sect. 6.3.2) that we say that the mapping f is *injective* if, $\forall x \in X$ and $\forall x' \in X$ with $f(x) = f(x')$, we have $x = x'$. A function that is both injective and surjective is called *bijective*.

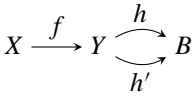
[10] I have previously used the symbol $\longmapsto$ for some mappings between physical systems. In this sense, I have treated them as individual systems, what is in general correct. There is some ambiguity when we deal with information, since in that case we abstract from the particular physical characters of the system. Nevertheless, also in this case, we often deal with a qubit in a well definite state.

Let $f : X \longrightarrow Y$ be a function. We say that f is a *monomorphism* if, for all sets A and pair of functions $g, g' : A \longrightarrow X$, when $f \circ g = f \circ g'$ then $g = g'$. This can be represented by means of the following graph or diagram:



(8.4)

We say that f is an *epimorphism* if for all sets B and pairs of functions $h, h' : Y \longrightarrow B$, if $h \circ f = h' \circ f$ then $h = h'$, as shown by the following diagram:



(8.5)

Let $f : X \longrightarrow Y$ be a function. Then f is injective iff it is a monomorphism, whereas it is surjective iff it is an epimorphism. Note that the notion of monomorphism and epimorphism is very general.

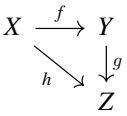
Another very important notion in mathematics is that of isomorphism. A function $f : X \longrightarrow Y$ is an *isomorphism*, denoted by $X \overset{\cong}{\rightarrow} Y$, if there exists a function $g : Y \longrightarrow X$ such that

$$g \circ f = 1_X \text{ and } f \circ g = 1_Y. \tag{8.6}$$

In such a case, we say that f is invertible and that g is the inverse of f. If there exists such an isomorphism, we say that X and Y are *isomorphic* and write $X \cong Y$. Note that isomorphism (like equivalence) satisfies

- *Reflexivity*: $\forall X, X \cong X$,
- *Symmetry*: $\forall X, Y$, if $X \cong Y$, we also have $Y \cong X$, and
- *Transitivity*: $\forall X, Y, Z$, if $X \cong Y$ and $Y \cong Z$, we have also $X \cong Z$.

Note, however, that isomorphism is stronger than (material) equivalence: both isomorphism and homomorphisms, on which I shall say more in the next subsection, in general, are structure-preserving maps while equivalence is not necessarily so. However, this means that equivalence is sufficient for sets. Another important concept is that of *commutative diagrams*. Consider the following diagram or graph of sets:



(8.7)

The diagram is said to be commutative if $g \circ f = h$. Similarly, we have also commutative square diagrams of sets:

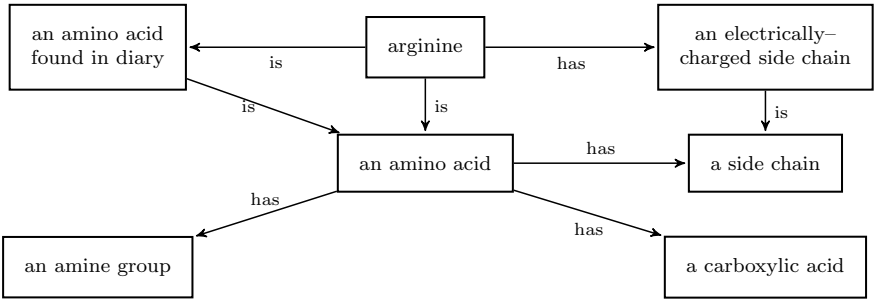


Fig. 8.2 Example of olog of the amino acid arginine. Note that the two closed circuits commute. Adapted from Spivak (2013, p. 23)

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ {\scriptstyle h}\downarrow & & \downarrow{\scriptstyle g} \\ C & \xrightarrow{i} & D \end{array} \tag{8.8}$$

We say that this is commutative if $g \circ f = i \circ h$. This establishes the important notion of *path equivalence*: we can reach D from A through two equivalent paths, passing by either B or C.

Ologs and Aspects

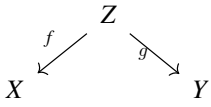
I shall introduce now another kind of diagram called *olog*. First of all, ologs deal with types and not tokens (while most sets are collections of tokens), i.e. with entire classes of things or objects and especially with equivalence classes of them, as far as all the elements or objects being of the same type share a complex of properties (or aspects, as we shall see soon) that make any of these elements or objects a representative of the same type on the same foot of, i.e. equivalent to, any item in the collection. For instance, the type of $\sqrt{2}$ is "real number", but "$\sqrt{2}$" itself can be taken as a type, i.e. abstract mathematical notion of square two. Note that in the axiomatisation based on set theory, we lose the notion of type since everything is member of a set. We can represent each type as a box containing a singular indefinite noun phrase, as displayed in Fig. 8.2. Ologs are very important when dealing with categorial problems and their use is practically universal.

One of the main lessons of our examination of QM is that properties are not intrinsic but are inferred given certain contexts (Sect. 5.2.4). This suggests a helpful generalisation of the notion of property: the notion of aspect. An *aspect* of a thing x is a way of viewing it, a particular way in which x can be regarded or measured. It represents our partial access to the information that, physically, is carried by the signal propagating from the physical object that is observed (Sects. 6.2.5 and 7.2.3), like light denoting that the object is yellow. In other words, by aspect we simply mean a function. The domain A of the function $f : A \longrightarrow B$ is the thing we are measuring or observing, and the codomain B is the set of possible 'answers' or results of the

measurement. In fact an aspect is a *relation*, which is on the line of a 'relational' view of QM (Sects. 3.2.3, 5.2.6, and 7.2.1). Having clarified this, in the following, for simplicity I shall sometimes also use the term "property" but always as denoting the notion of aspect defined here.

Product and Pullback

Given sets X and Y, a *span* on X and Y is a set Z together with functions $f : Z \longrightarrow X$ and $g : Z \longrightarrow Y$:

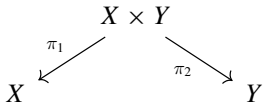


(8.9)

Spans can be hierarchically ordered in subspans. This notion will help us in dealing with products. Given two *projection* functions (where the term here has a pure abstract mathematical meaning)

$$\pi_1 : X \times Y \longrightarrow X \text{ and } \pi_2 : X \times Y \longrightarrow Y, \tag{8.10}$$

the *product* (or intersection) of sets X and Y denoted $X \times Y$ is defined as the set of the ordered pairs (x, y), where $x \in X$, $y \in Y$:



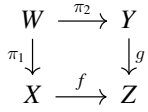
(8.11)

We can thus understand $X \times Y$ as a matrix formed by the direct product of the columnar vector X and the row vector Y. Then, the two projections project from the matrix to these vectors.

The *fibre product* is the set

$$X \times_Z Y := \{(x \in X, z \in Z, y \in Y) | f(x) = z = g(y)\}. \tag{8.12}$$

In other words, a fibre product deals not only with a 'source' product but also with a terminal object Z, as shown by the commutative diagram:



(8.13)

Thus, the fibre product of X, Y is such that there exists a set Z of the z's that are both $f(x)$ and $g(y)$.

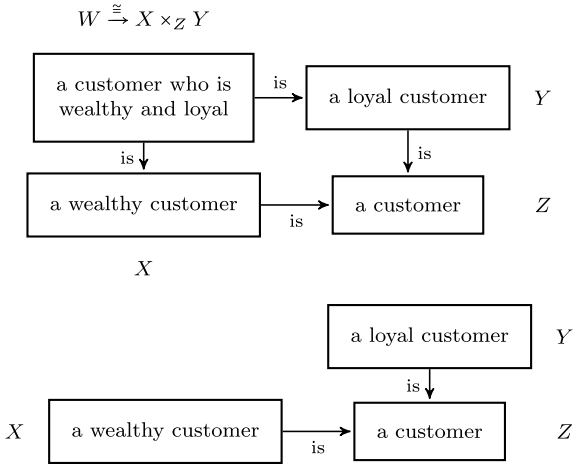


Fig. 8.3 Consider the two ologs above. The one on the top is the pullback of the one on the bottom. Adapted from Spivak (2013, p. 43)

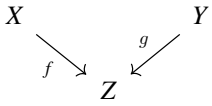
The *pullback* of X and Y over Z is any set W for which we have an isomorphism

$$W \xrightarrow{\cong} X \times_Z Y. \tag{8.14}$$

An example is shown in Fig. 8.3.

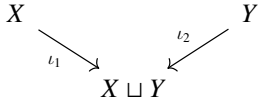
Coproduct and Pushout

Let us now define the inverse mappings and notions. First let us introduce the notion of *cospan* for some sets X, Y, Z and mappings f, g:



(8.15)

Also cospans can be hierarchically ordered. If X and Y are sets, the *coproduct* (or the union) of X and Y, denoted $X \sqcup Y$, is defined as the "disjoint union" of X and Y, i.e. the set for which an element is an element of X or Y (if something is an element of both X and Y then we include both copies and distinguish between them, in $X \sqcup Y$): e.g. if $X = \{a, b, c, d\}$ and $Y = \{d, e, f\}$, then $X \sqcup Y = \{a, b, c, d, d, e, f\}$. In general graphical terms, we have



(8.16)

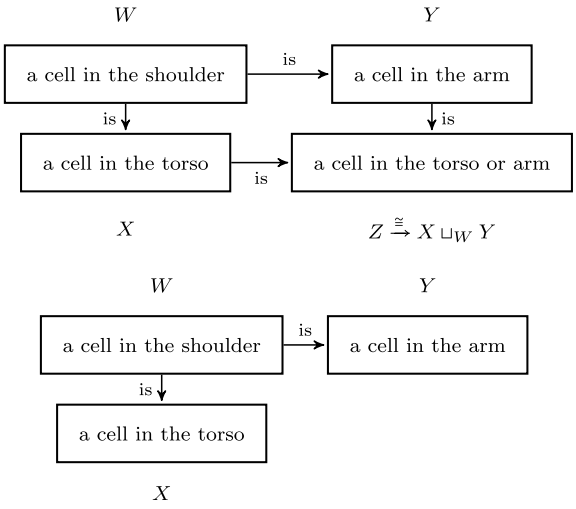


Fig. 8.4 Example of pushout. In each example above, the diagram on the top is intended to be a pushout of the diagram on the bottom. The new object, $Z \xrightarrow{\cong} X \sqcup_W Y$, is the union of X and Y, but instances of W are equated to their X and Y aspects. In the olog in the bottom, the two arrows are inclusions: every cell in the shoulder is here considered to be in the arm or in the torso. The pushout is then just the union, where cells in the shoulder are not double-counted. Adapted from Spivak (2013, p. 53)

where i_1, i_2 are called *injections*. Spans and cospans are meets and joins, respectively: the meet of X and Y is the biggest thing smaller than both, i.e. a greatest lower bound, and the join of X and Y is the smallest thing bigger than both, i.e. a least upper bound. In other words, the logical product is the ground of spans (and products) as well as the logical sum is the ground of cospans (and coproducts) (see Tables 3.2, and 3.3, Sect. 3.3.1).

As we have the fibre product (8.12), we have also a fibre sum. The *fibre sum*, denoted $X \sqcup_W Y$, is defined as the equivalence relations

$$x \sim f(w) \text{ and } y \sim g(w), \tag{8.17}$$

$\forall x \in X, y \in Y$. In other words, instead of a terminal object, we have a coproduct with an initial object W. The diagram

$$\begin{array}{ccc} W & \xrightarrow{g} & Y \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle \iota_2} \\ X & \xrightarrow{\iota_1} & Z \end{array} \tag{8.18}$$

commutes when the *pushout* of X and Y over W is any set Z such that we have an isomorphism

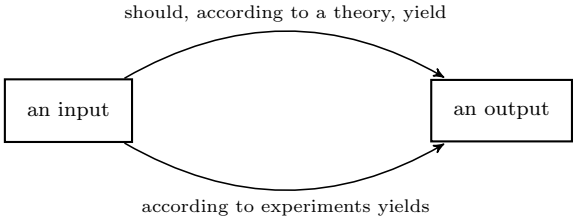


Fig. 8.5 Example of equaliser. An often raised question is: "when does the theory match the experiment?" The rough answer is given by the equaliser of the above diagram. Thus, this equaliser is the set of all inputs for which the theory and the experiment yield the same output. Adapted from Spivak (2013, p. 49)

$$Z \xrightarrow{\cong} X \sqcup_W Y, \tag{8.19}$$

which is a kind of inverse notion relative to the pullback (8.14). An example is shown in Fig. 8.4. To understand the difference between pullback and pushout, we have that the pullback in Fig. 8.3 is a customer that is *both* wealthy and loyal, while the pushout shown in Fig. 8.4 is a cell that is in the torso *or* arm.

Equaliser and Coequaliser

Let us introduce the important notion of *equaliser*. Suppose the two parallel arrows between objects X, Y displayed by the following diagram:

$$X \underset{g}{\overset{f}{\rightrightarrows}} Y \quad Eq(f,g) \xrightarrow{p} X \underset{g}{\overset{f}{\rightrightarrows}} Y. \tag{8.20}$$

The *equaliser* of f and g is the commutative diagram on the right that is defined as

$$Eq(f,g) := \{x \in X | f(x) = g(x)\}. \tag{8.21}$$

The example of a theory matching experiments is shown in Fig. 8.5. The inverse concept is that of coequaliser. Suppose two parallel arrows between objects X, Y:

$$X \underset{g}{\overset{f}{\rightrightarrows}} Y \quad X \underset{g}{\overset{f}{\rightrightarrows}} Y \xrightarrow{q} Coeq(f,g). \tag{8.22}$$

The *coequaliser* of f and g is the commutative diagram on the right defined as

$$Coeq(f,g) := \{y \in Y | f(x) = g(x)\}. \tag{8.23}$$

8.1.3 Categories and Functors

Definition of Category

A *category* consists of a collection (class) or more collections of objects, all of which are related in some way.[11] In everyday speech, we think of a category as a kind of thing. In mathematics, a category can also be construed as a collection of objects and a type of relationship between pairs of such objects. For this kind of object–relationship duo to count as a category, two rules need to be satisfied:

(i) Everything must be related to itself by simply being itself, and
(ii) If one object is related to another one through a morphism and the second is related to a third through a morphism, then the first is also related to the third through a morphism.

One can think of a category as a graph (a notion that we shall introduce formally later on) in which certain paths have been declared equivalent, although the formal definition of category does not reduce to the "graph+path equivalences" notion. Formally, a *category* $\mathcal{C}$ consists of:

- A collection $\mathrm{Ob}(\mathcal{C})$ of objects;
- For each $A, B \in \mathrm{Ob}(\mathcal{C})$, a collection $\mathcal{C}(A, B)$ of maps or arrows or morphisms from A to B, denoted also $\mathrm{Hom}_{\mathcal{C}}(A, B)$;
- For all objects $A, B, C \in \mathrm{Ob}(\mathcal{C})$ and morphisms $f \in \mathcal{C}(A, B)$, $g \in \mathcal{C}(B, C)$, a function

$$\mathcal{C}(B, C) \times \mathcal{C}(A, B) \longrightarrow \mathcal{C}(A, C) \tag{8.24a}$$

$$(g, f) \longmapsto g \circ f, \tag{8.24b}$$

  called *composition*: we can capture these fundamental elements in the definition of the notion of category by means of the olog shown in Fig. 8.6.;
- Moreover, for each $f \in \mathcal{C}(A, B)$, $g \in \mathcal{C}(B, C)$ and $h \in (C, D)$, we have

$$(h \circ g) \circ f = h \circ (g \circ f); \tag{8.25}$$

  This property is called *associativity*.
- Finally, for each $A \in \mathrm{Ob}(\mathcal{C})$, there is an element 1_A of $\mathcal{C}(A, A)$, called the *identity* on A, satisfying the *identity laws*: for each $f \in \mathcal{C}(A, B)$, we have

$$f \circ 1_A = f = 1_B \circ f. \tag{8.26}$$

[11] Spivak (2013, Chap. 3 and Sects. 4.1, 4.2).

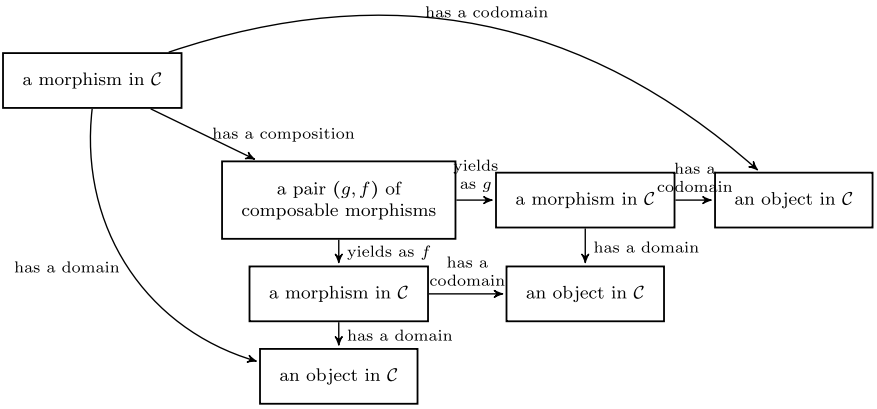


Fig. 8.6 The basic structure (type) of category $\mathcal{C}$ as an olog: it is made of objects, morphisms and composition of the latter. Note that the three closed paths commute. Adapted from Spivak (2013, p. 120)

Sets

Let us now consider some examples of categories. Sets satisfy this general notion of category. So they are ground-level categories. Since some classes are too big to be sets, we introduce the following terminology. A class is called *small* if it is a set and *large* otherwise. A category $\mathcal{C}$ is small if the class or collections of all maps in $\mathcal{C}$ are small and large otherwise. If $\mathcal{C}$ is small, also the class of objects of $\mathcal{C}$ is small, since objects correspond one-to-one to identity maps. If only classes of specific objects pertaining to a category are small, we say that the category is locally small.

Preorders

We pass to other interesting examples of categories by introducing some order. Note that, although many kinds of orders are possible (and some will be explored), the most basic (and universal) one is preorder, and a Partially Ordered Set (POSet) is simply a preorder with the antisymmetry property (that is crucial for giving rise to relations that are fundamental for category theory in general, like hierarchical spans and cospans). It is helpful to recall here the essential properties of preorders. Let S be a set and $R \subseteq S \times S$ a binary relation on S symbolised with $\leq$. Then, we say that R is a *preorder* if $\forall s, s', s'' \in S$ we have

Reflexivity: $s \leq s$, and
Transitivity: If $s \leq s'$ and $s' \leq s''$, then $s \leq s''$.

Moreover, we say that R is a *partial order* (POSet) if it is a preorder and, in addition, $\forall s, s' \in S$ we have

Antisymmetry: If $s \leq s'$ and $s' \leq s$, then $s = s'$.

We say that R is a *linear order* if it is a partial order and, in addition, $\forall s, s' \in S$ we have:

Comparability: Either $s \leq s'$ or $s' \leq s$.

Preorders, partial and linear orders are denoted by $(S, \leq)$, where S is a set. Note that every equivalence relation is a preorder but rarely are they partial orders. For example, if $S = \{1, 2\}$ and we put $R = S \times S$, then this is an equivalence relation. This is a preorder but not a partial order (because $1 \leq 2$ and $2 \leq 1$, but $1 \neq 2$).

Monoids

Mathematical objects called *monoids* (and groups) are tasked with encoding the agent's perspective, i.e. what an agent can do, and what happens when different actions are done in succession. Note that *agent* is understood here in the broadest sense of whatever system being capable to give rise to some action. A monoid can be construed as a set of actions, together with a formula that encodes how a sequence of actions is itself considered an action. A group is the same as a monoid, except that every action is required to be reversible (Sect. 1.2.5).

A *monoid* is a sequence $(M, e, *)$, where M is a set, $e \in M$ is the identity element, and the symbol $*$ denotes the multiplication symbol for monoids, so that

$$* : M \times M \longrightarrow M \tag{8.27}$$

is function such that $\forall m, n, p \in M$ the conditions (8.24)–(8.26) hold:

$$m * e = m = e * m, \tag{8.28a}$$

$$(m * n) * p = m * (n * p). \tag{8.28b}$$

The last formula is the associativity law for monoids, which allows us to drop parentheses when no ambiguity arises.

Lists

Let X be a set. A *list* in X, denoted $\text{List}(X)$, is a pair (n, f), where $n \in \mathbb{N}$ and is called the length of the list, while

$$f : \underline{n} \longrightarrow X \tag{8.29}$$

is a function, where I recall that $\underline{n} = \{1, 2, \ldots, n\}$. This list can be denoted by

$$(n, f) = [f(1), f(2), \ldots, f(n)]. \tag{8.30}$$

The empty list is the unique list in which $n = 0$ and can be denoted by []. Given an element $x \in X$, the singleton list on x is $[x]$. Given two list $L = (n, f)$ and

$L' = (n',\, f')$, the concatenation of L and L' (Eq. (8.24)), denoted by $L + +L'$ is the list $(n + n',\, f + +f')$, where

$$f + +f' : \underline{n + n'} \longrightarrow X \tag{8.31}$$

is given on $i \leq n + n'$ by

$$(f + +f')(i) := \begin{cases} f(i) & \text{if } \ i \leq n \\ f'(i - n) & \text{if } \ i \geq n + 1 \end{cases} \tag{8.32}$$

Practically, the new list is simply a sequence of the lists L and L'.

Free Monoids

The *free monoid* generated by the set X is the sequence $\mathcal{M} := (\text{List}(X), [], ++)$, and X is the set of generators of monoid $\mathcal{M}$. Lists of generators provide us all the possible ways to write elements of $\mathcal{M}$. Let now (i) G be a finite set, (ii) $n \in \mathbb{N}$, and, for each $1 \leq i \leq n$, (iii) m_i, m_i' be elements of $\text{List}(G)$. The *monoid presented* by generators G and relations $\{(m_i, m_i') | 1 \leq i \leq n\}$ is the monoid $\mathcal{M} = (M, e, *)$ defined as follows. Let $\sim$ denote the equivalence relation on $\text{List}(G)$ generated by

$$\left\{(x m_i y \sim x m_i' y) | x, y \in \text{List}(G), 1 \leq i \leq n\right\}, \tag{8.33}$$

and define $M = \text{List}(G)/\sim$ (i.e. we drop equivalent alternatives). Let $e = []$ and $a * b$ obtained by concatenating representing lists. Note that every free monoid is a presented monoid because we can just take the set of relations to be empty.

Let us consider an example. Let $G = \{a, b, c, d\}$, where the letters denote buttons that can be pressed for performing some operations. The free monoid $\text{List}(G)$, presented by G, is the set of all ways to press the buttons subsequently. For instance, we can have the list $[a, a, c, c, d]$. The idea of presented monoid comes from the fact that e.g. it turns out that pressing $[a, a, c]$ always gives an equivalent result as pressing $[d, d]$, while, pressing $[c, a, c, a]$ is equivalent to doing nothing. In this case, we would have $m_1 = [a, a, c]$, $m_1' = [d, d]$, and $m_2 = [c, a, c, a]$, $m_2' = []$, as well as relations $\{(m_1, m_1'), (m_2, m_2')\}$. A calculation on $M = \text{List}(G)/\sim$ takes into account equivalences for dropping terms and gives, e.g.

$$\begin{aligned} [b, c, b, d, d, a, c, a, a, c, d] &= [b, c, b, a, a, c, a, c, a, a, c, d] = [b, c, b, a, a, a, c, d] \\ &= [b, c, b, a, d, d, d]. \end{aligned} \tag{8.34}$$

A monoid is called *cyclic* if it has a presentation involving only one generator. For instance, suppose that this generator is $\{Q\}$. Then, we would have underlying sets $\{[], [Q], [Q, Q], [Q, Q, Q], \ldots\}$, with identity element $[]$ and multiplication given by the concatenations like $[Q, Q, Q] + +[Q, Q] = [Q, Q, Q, Q, Q]$. This turns out to be just $\mathbb{N}$, i.e. the additive monoid of natural numbers.

Table 8.1 Table of actions of a finite-state machine

ID	a	b
State 0	State 1	State 2
State 1	State 2	State 1
State 2	State 0	State 0

Action

Let $\mathcal{M} = (M, e, *)$ be a monoid and let S be a set. An *action* of $(M, e, *)$ on S, or simply an action of M on S or an M-action on S, is a function

$$\circlearrowright : M \times S \longrightarrow S, \tag{8.35}$$

such that the following conditions hold $\forall m, n \in M$ and $e \in M$ as well as $\forall s \in S$ (see also Eq. (8.28)):

$$e \circlearrowright s = s, \tag{8.36a}$$

$$m \circlearrowright (n \circlearrowright s) = (m * n) \circlearrowright s. \tag{8.36b}$$

Example

As said, a monoid helps us to understand how an agent acts on the set of states of an object. If monoids are understood in terms of how they act on sets, then it is reasonable to think of them in terms of ologs. Let us consider the example of a *finite-state machine*. A deterministic finite-state machine is a quintuple $(\Sigma, S, s_0, \delta, F)$, where

1. Σ is a finite non-empty set of symbols, called the *input alphabet*,
2. S is a finite, non-empty set, called the *state set*,
3. $\delta : \Sigma \times S \longrightarrow S$ is a function, called the *state transition function* (corresponding, e.g. to non-necessarily reversible physical transformations), and
4. $s_0 \in S$ is an element, called the *initial state*,
5. $F \subseteq S$ is a subset, called the *set of final states*.

Note that a Turing machine is a more complex object than, but satisfying the fundamental requisites of, a finite-state machine (Sect. 3.2.4). In fact, also finite-state machines in general can run indefinitely, displaying the halting problem. We can express this in the language of free monoids and their actions on finite sets. Let Σ, S be finite non-empty sets. Giving a function $\delta : \Sigma \times S \longrightarrow S$ is equivalent to giving an action of the free monoid $\mathrm{List}(\Sigma)$ on S. If Σ and S are the sets shown in Fig. 8.7, the displayed action of $\mathrm{List}(\Sigma)$ on S would be given by the action Table 8.1.

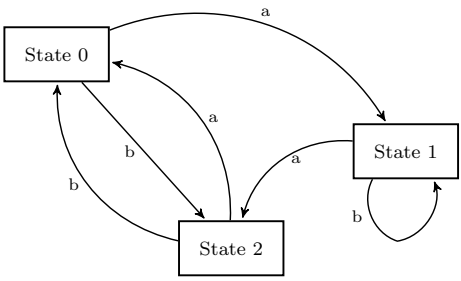


Fig. 8.7 A deterministic finite-state machine with alphabet $\Sigma = \{a, b\}$ and state set $S =$ {State 0, State 1, State 2}. If pressed, we will make State 0 the initial state and {State 0, State 1, State 2} the set of final states. Adapted from Spivak (2013, p. 77)

Ordering Data

The necessity to order things (in different ways) is crucial not only for practical purposes but also to knowledge. Data, in particular, the set of observations made during experiment, plays a primary role in science of any kind. To be useful, data must be organised, often in a row-and-column display called a table. A database is a collection of tables, each table T of which consists of a set of columns and a set of rows. The existence of table T suggests the existence of a fixed methodology for observing objects or events of a certain type. Each column c in T prescribes a single kind or method of observation, so that the datum inhabiting any cell in column c ideally refers to an observation of that kind. Each row r in T has a fixed source event or object, which can be observed using the methods prescribed by the columns. The cell (r, c) refers to the observation of kind c made on event r. All of the rows in T should refer to uniquely identifiable objects or events of a single type. In other words, as we shall, they need to be organised in equivalence classes.

Homomorphisms Among Monoids

We have seen that a monoid $\mathcal{M} = (M, e, *)$ involves a set, an identity element and a multiplication formula. Thus, for two monoids to be comparable, their sets, their identity elements and their multiplication formulas should be appropriately comparable. For example, the additive monoids $\mathbb{N}$ and $\mathbb{Z}$ should be comparable because $\mathbb{N} \subseteq \mathbb{Z}$, the identity elements in both cases are the same $e = 0$ and the multiplication formulas are both integer addition. Therefore, if $\mathcal{M} := (M, e, *)$ and $\mathcal{M}' := (M', e', *')$ are monoids, we define a monoid *homomorphism* f from $\mathcal{M}$ to $\mathcal{M}'$, denoted $f : \mathcal{M} \longrightarrow \mathcal{M}'$ to be a function

$$f : M \longrightarrow M', \tag{8.37}$$

satisfying two conditions:

$$f(e) = e' \text{ and } f(m_1 * m_2) = f(m_1) *' f(m_2), \ \forall m_1, m_2 \in M, \tag{8.38}$$

that is, given two elements of M, one gets the same element of M' whether one (i) first takes the product of these elements in M and sends the result to M' through f or (ii) first sends these elements to M' thanks to f and then takes the product in M'. Thus, a homomorphism is a particular instance of morphism being a structure-preserving map between two sets.

The set of all monoid homomorphisms from $\mathcal{M}$ to $\mathcal{M}'$ is denoted $\mathrm{Hom}_{\mathbf{Mon}}(\mathcal{M}, \mathcal{M}')$. Example: if A is the set of amino acids and $\mathcal{A} = \mathrm{List}(A)$ the free monoid on A, the process of genetic translation can be represented by a monoid homomorphism $G : \mathcal{T} \longrightarrow \mathcal{A}$, turning the list $\mathcal{T} = \mathrm{List}(T)$ of RNA triplets $T = R^3$ (where $R = A, C, G, U$) into a polypeptide.

Let G be a set, let $F(G) := (\mathrm{List}(G), [], ++)$ be the free monoid on G, and let $\mathcal{M} := (M, e, *)$ be any monoid. There is a natural isomorphism

$$\mathrm{Hom}_{\mathbf{Mon}}(F(G), \mathcal{M}) \xrightarrow{\cong} \mathrm{Hom}_{\mathbf{Set}}(G, M). \tag{8.39}$$

Monoids are particularly relevant for the categorial structure of QM, which will turn out to be a monoid with a submonoid that is in fact a group.

Groups

As mentioned, a fourth fundamental example of category is represented by groups. A *group* is a monoid $(M, e, *)$ in which every element $m \in M$ has an inverse (Sect. 1.2.5). Note also that any element m has at most one inverse. An element $m \in M$ is said to have an inverse if there exists an $m' \in M$ such that $m * m' = e$ and $m' * m = e$. A subgroup $\mathcal{G}'$ of a group $\mathcal{G}$ is a group $\mathcal{G}'$ and a monomorphism (injective homomorphism) $\mathcal{G}' \longrightarrow \mathcal{G}$. It can be proved that every group is a subgroup of the permutation group on some set: each element of the group $\mathcal{G}$ defines, via "left multiplication by that element," a permutation on the set $\mathcal{G}$ itself.[12]

Let us now define the notion of *group action*. Let $\mathcal{G} = (G, e, *)$ be a group and S a set. An *action* of G on S is a function

$$\circlearrowright : G \times G \longrightarrow S \tag{8.40}$$

such that $\forall s \in S$ and $g, g' \in G$, we have (Eqs. (8.24)–(8.26))

$$e \circlearrowright s = s, \tag{8.41a}$$

$$g \circlearrowright (g' \circlearrowright s) = (g * g') \circlearrowright s. \tag{8.41b}$$

We recover here the definition of monoid action (8.36). A group homomorphism $f : \mathcal{G} \longrightarrow \mathcal{G}'$ is defined to be a monoid homomorphism $\mathcal{G} \longrightarrow \mathcal{G}'$, where the groups $\mathcal{G}, \mathcal{G}'$ are being regarded as monoids.

[12] Geroch (1985, pp. 18–19).

Graphs

Now I introduce a fifth example of category: graphs (see also Sects. 7.1.4 and 7.2.3). A *graph* Γ consists of a sequence $\Gamma := (V, A, \mathrm{src}, \mathrm{tgt})$, where

- V is a set, called the set of *vertices* of Γ,
- A is a set, called the set of *arrows* of Γ,
- $\mathrm{src} : A \longrightarrow V$ is a function, called the *source function* for Γ and
- $\mathrm{tgt} : V \longrightarrow A$ is a function, called the *target function* for Γ.

Given an arrow $a \in A$, we refer to $\mathrm{src}(a)$ as the source vertex of a and to $\mathrm{tgt}(a)$ as the target vertex of a. Note that up to now we have already used graphs (all the previous diagrams of the present section are in fact also graphs).

We all know that a path must always follow the direction of arrows. Let us formulate this in precise terms. Let $\Gamma := (V, A, \mathrm{src}, \mathrm{tgt})$ be a graph. A *path* of length n in Γ, denoted $p \in \mathrm{Path}_\Gamma^{(n)}$, is a head-to-tail sequence

$$p = (v_0 \overset{a_1}{\longmapsto} v_1 \overset{a_2}{\longmapsto} v_2 \overset{a_3}{\longmapsto} \ldots \overset{a_n}{\longmapsto} v_n) \tag{8.42}$$

of arrows in Γ, which we denote by $v_0 a_1 a_2 \ldots a_n$. Note that a path is predefinite and sequential (and so the order is inversed relative to that of quantum operations). In particular, we have canonical isomorphisms

$$\mathrm{Path}_\Gamma^{(1)} \cong A, \quad \mathrm{Path}_\Gamma^{(0)} \cong V, \tag{8.43}$$

that is, paths of length 1 are arrows and paths with length 0 are vertices. We refer to the path of length 0 on vertex v as the *trivial* path on v and denote it simply by v. We denote by Path_Γ the set of all paths in Γ:

$$\mathrm{Path}_\Gamma := \bigcup_{n \in \mathbb{N}} \mathrm{Path}_\Gamma^{(n)}. \tag{8.44}$$

If p is a path with source vertex $\overline{\mathrm{src}}(p) = v$ and target vertex $\overline{\mathrm{tgt}}(p) = w$, we may denote it by $p : v \longmapsto w$. Given two vertices $v, w \in V$, we write $\mathrm{Path}_\Gamma(v, w)$ the set of all paths $p : v \longmapsto w$. There is a concatenation operation on paths. Given paths $p : v \longmapsto w$ and $q : w \longmapsto z$, we denote their concatenation $pq : v \longmapsto z$ (see also Eq. (8.24)).

Let $\Gamma = (V, A, \mathrm{src}, \mathrm{tgt})$ and $\Gamma' = (V', A', \mathrm{src}', \mathrm{tgt}')$ be graphs. A *graph homomorphism* f from Γ to Γ', denoted $f : \Gamma \longrightarrow \Gamma'$, consists of two functions

$$f_0 : V \longrightarrow V' \text{ and } f_1 : A \longrightarrow A'$$

such that the two diagrams below commute:

$$\begin{array}{ccc} A & \xrightarrow{f_1} & A' \\ {\scriptstyle \mathrm{src}}\downarrow & & \downarrow{\scriptstyle \mathrm{src}'} \\ V & \xrightarrow{f_0} & V' \end{array} \qquad \begin{array}{ccc} A & \xrightarrow{f_1} & A' \\ {\scriptstyle \mathrm{tgt}}\uparrow & & \uparrow{\scriptstyle \mathrm{tgt}'} \\ V & \xrightarrow{f_0} & V' \end{array} \tag{8.45}$$

The idea, roughly stated, is: "arrows are bound to their vertices". Under a map of graphs $\Gamma \longrightarrow \Gamma'$, any arrow must still connect the vertices it connected before.

Path Equivalence Declaration

Let $\Gamma := (V, A, \mathrm{src}, \mathrm{tgt})$ be a graph and Path_Γ the set of all paths in Γ. A *path equivalence declaration* (PED) is an expression of the form $p \simeq q$, where $p, q \in \mathrm{Path}_\Gamma$ have the same source and the same target: $\mathrm{src}(p) = \mathrm{src}(q)$ and $\mathrm{tgt}(p) = \mathrm{tgt}(q)$. A *congruence* on Γ is a relation $\simeq$ on Path_Γ that has the following properties:

1. The relation $\sim$ is an equivalence relation.
2. If $p \sim q$ then $\mathrm{src}(p) = \mathrm{src}(q)$.
3. If $p \sim q$ then $\mathrm{tgt}(p) = \mathrm{tgt}(q)$.
4. Suppose $p, q : b \longmapsto c$ are paths, and $m : a \longmapsto b$ is an arrow. If $p \sim q$ then $mp \sim mq$.
5. Suppose $p, q : a \longmapsto b$ are paths, and $n : b \longmapsto c$ is an arrow. If $p \sim q$ then $pn \sim qn$.

Any set of PEDs generates a congruence. We can draw any preorder $(S, \leq)$ as a graph with vertices S and with an arrow $a \longmapsto b$ if $a \leq b$. These are precisely the graphs with the following two properties for any vertices $a, b \in S$:

1. There is at most one arrow $a \longmapsto b$, and
2. If there is a path from a to b, then there is an arrow $a \longmapsto b$

If $(S, \leq)$ is a partial order, then the associated graph has an additional "no loops" property,

3. If $n \in \mathbb{N}$ is an integer with $n \geq 2$, then there no paths of length n that start at a and end at a.

If $(S, \leq)$ is a linear order, then there is an additional "comparability" property,

4. For any two vertices $a, b \in S$ there is an arrow $a \longmapsto b$ or an arrow $b \longmapsto a$.

Thus, graphs give us a nice way to visualise orders. Let us build a definition of morphisms of orders. Let $\mathcal{S} := (S, \leq)$ and $\mathcal{S}' := (S', \leq')$ be preorders (respectively partial orders or linear orders). A *morphism of preorders* (respectively of partial orders or of linear orders) f from $\mathcal{S}$ to $\mathcal{S}'$, denoted $f : \mathcal{S} \longrightarrow \mathcal{S}'$, is a function $S \longrightarrow S'$ such that $\forall s_1, s_2 \in S$, and if $s_1 \leq s_2$ then $f(s_1) \leq' f(s_2)$.

Categories of Categories

Once that we have considered sets, preorders, monoids, groups and graphs as examples of categories, let us come back to the general issues concerning categories. Note that we can build several categories as general types of the categories as far considered: the category (and not the set!) of all sets (denoted by **Set**), the category of all monoids (denoted by **Mon**), the category of all groups (denoted by **Grp**), the category of all graphs (denoted by **Grph**) and the category of all preorders (denoted by **PrO**). Obviously, there are many others, and some of them will be explored in the following. It is also possible to build the category of all categories, denoted by **CAT**. As recalled in Sect. 8.1.1, while the notion of the set of all sets is not self-consistent and gives rise to well-known paradoxes that explain the failure of the foundational program based on sets,[13] categories, being structured collections of objects, do not undergo the same problem. Obviously, another question is if **CAT** encompasses everything. The question could be reformulated as follows: are problems or aspects of the world that cannot be framed in category theory? As we shall see, this question is crucial for the foundations of QM. If category theory is really so general, it should help us to deal also with this kind of problem.

As the previous examples show, the relations among these or other categories are very interesting. For instance, let **FLin** be the full subcategory of **PrO** spanned by the linear orders. That is, given linear orders X, Y, every morphism of preorders $X \longmapsto Y$ counts as a morphism in **FLin**.

Functors

A monoid is a category with just a single object (the set M). A group is just a monoid in which every element is invertible. Then, also a group must be a category with one object (the set G, while groupoids are like groups except a groupoid can have more than one object) and with an additional property having to do with invertibility, so that all maps are isomorphisms.

However, there are also other categories with more objects. In fact, I recall that a category $\mathcal{C} = (\mathrm{Ob}(\mathcal{C}), \mathrm{Hom}_{\mathcal{C}}, \mathrm{dom}, \mathrm{cod}, 1, \circ)$, involves a set of objects, a set of morphisms, a notion of domains and codomains, a notion of identity morphisms and a composition formula, as displayed by the olog Fig. 8.6. Moreover, we can be interested in connecting objects of different categories. For two categories to be comparable, these various components should be appropriately comparable with those of another category $\mathcal{C}'$. This is the purpose of the notion of functor. A *functor* F from $\mathcal{C}$ to $\mathcal{C}'$ is defined by announcing some constituents and asserting that they conform to some laws. Specifically, one announces

[13] The error was found by Russell and deals with the fact that both the definitions of "the set of the sets that are not members of themselves" and of "the set of the sets that are members of themselves" (and similar sets) lead to a contradiction that cannot be resolved within the system that defines the set itself (Russell 1902, 1903).

- A function $\mathrm{Ob}(F) : \mathrm{Ob}(\mathcal{C}) \longrightarrow \mathrm{Ob}(\mathcal{C}')$ between categories $\mathcal{C}$ and $\mathcal{C}'$, which will be sometimes denoted simply by $F : \mathrm{Ob}(\mathcal{C}) \longrightarrow \mathrm{Ob}(\mathcal{C}')$, and
- For every pairs of objects $C, D \in \mathrm{Ob}(\mathcal{C})$, a function

$$\mathrm{Hom}_F(C, D) : \mathrm{Hom}_{\mathcal{C}}(C, D) \longrightarrow \mathrm{Hom}_{\mathcal{C}'}(F(C), F(D)), \tag{8.46}$$

which sometimes will be written: $F : \mathrm{Hom}_{\mathcal{C}}(C, D) \longrightarrow \mathrm{Hom}_{\mathcal{C}'}(F(C), F(D))$.

In agreement with Eqs. (8.24) and (8.26), one asserts that the following laws hold:

- Identities are preserved by F. That is, for any object $C \in \mathrm{Ob}(\mathcal{C})$, we have

$$F(1_C) = 1_{F(C)}; \tag{8.47}$$

and
- Composition is preserved by F. That is, for any objects $B, C, D \in \mathrm{Ob}(\mathcal{C})$ and morphisms $g : B \longrightarrow C$ and $h : C \longrightarrow D$, we have

$$F(h \circ g) = F(h) \circ F(g). \tag{8.48}$$

Note also that a functor $F : \mathcal{C} \longrightarrow \mathcal{C}'$ is said *faithful* (respectively, *full*) if $\forall C, D \in \mathcal{C}$, the function

$$\mathcal{C}(C, D) \longrightarrow \mathcal{C}'(F(C), F(D)) \tag{8.49}$$

$$f \longmapsto F(f) \tag{8.50}$$

is injective (respectively, surjective).

Functors preserve isomorphism. For instance, we have a functor

$$U : \mathbf{Mon} \longrightarrow \mathbf{Set} \tag{8.51}$$

that takes every monoid to its underlying set and every monoid homomorphism to its underlying set morphism (Eq. (8.39)). Moreover, to every group, we can assign its underlying monoid. Similarly, a group homomorphism is just a monoid homomorphism of its underlying monoids. This means that there is a functor

$$V : \mathbf{Grp} \longrightarrow \mathbf{Mon} \tag{8.52}$$

that sends every group or group homomorphism to its underlying monoid or monoid homomorphism. Suppose you are working with symmetries, but suppose also that the symmetry breaks somewhere, or you add some extra observable such that there is no longer reversibility under the symmetry. These are physical contexts that we have amply explored in the last chapters. We can desire to relax the requirement that every

action be reversible without changing anything else, as it happens with decoherence (Sects. 3.1.3 and 5.1.3). You want to know where you can go, or what's allowed. The answer is to simply pass from the category of groups (or group actions) to the category of monoids (or monoid actions). As we shall see, we could also proceed the other way around and make use of monoids in the general case and pass to the category of groups in special cases.

I have mentioned the universal applicability of graphs. In fact, all categories have underlying graphs. We can make this notion precise by saying that, if $\mathcal{C} = (\mathrm{Ob}(\mathcal{C}), \mathrm{Hom}_{\mathcal{C}}, \mathrm{dom}, \mathrm{cod}, \mathrm{ids}, \circ)$ is a category with its objects, domain, codomain, identities and composition, there is a graph $\Gamma = (\mathrm{Ob}(\mathcal{C}), \mathrm{Hom}_{\mathcal{C}}, \mathrm{dom}, \mathrm{cod})$ underlying $\mathcal{C}$, so that we have the functor

$$U : \mathbf{Cat} \longrightarrow \mathbf{Grph}, \tag{8.53}$$

where **Cat** stands for whatever category.

Let **PrO** be the category of preorders and **Grph** be the category of graphs. There is unique functor $P : \mathbf{PrO} \longrightarrow \mathbf{Grph}$ such that for any preorder $\mathcal{X} = (X, \leq)$, the graph $P(\mathcal{X})$ has vertices X. Since there are also functors $W : \mathbf{PrO} \longrightarrow \mathbf{Set}$ sending $(X, \leq)$ to X, and $T : \mathbf{Grph} \longrightarrow \mathbf{Set}$ sending $(V, A, \mathrm{src}, \mathrm{tgt})$ to V, we have the following commutative triangle:

$$\begin{array}{ccc} \mathbf{PrO} & \xrightarrow{\quad P \quad} & \mathbf{Grph} \\ & {\scriptstyle W}\searrow \quad \swarrow{\scriptstyle T} & \\ & \mathbf{Set} & \end{array} \tag{8.54}$$

A database schema (or schema, for short) consists of a graph together with a certain kind of equivalence relation on its paths. We can introduce a category **Sch** that has schemas as objects and appropriately modified graph homomorphisms as morphisms. It can be proved that the category of schemas is equivalent (in a specific sense that will be defined below) to the category of categories

$$\mathbf{Sch} \sim \mathbf{CAT}. \tag{8.55}$$

The difference between schemas and categories is like the difference between monoid presentations, given by generators and relations, and the monoids themselves. The same monoid has (infinitely) many different presentations, and so it is for categories: many different schemas can present the same category. Thus, a database schema is a category presentation (a notion on which I shall come back). Computer scientists may think of the schema as syntax and the category it presents as the corresponding semantics. A schema is a compact form, and can be specified in finite space and time while generating something infinite.

8.1.4 Universal Constructions

Universality

This examination raises the problem of universality.[14] *Universality* arises when we are interested not just in the existence of a solution to problems of construction but in its *canonicity*. This canonicity should guarantee uniqueness, in the sense that a canonical solution should be unique up to (unique) isomorphism. The notion of canonicity has a simple interpretation in the case of POSets, as an extremal solution: one that is the least or the greatest among all solutions. Such an extremal solution is obviously unique. When it comes to sets that map to, e.g. elements A and B, the $(A \times B)$-grid is ideal since it projects on to both A and B as straightforwardly as possible (see Diagram (8.11)). When it comes to sets that collect the elements of both A and B, the disjoint union $A \sqcup B$ is ideal since it includes both A and B without confusion or superfluity (see Diagram (8.16)). It may be recalled that in **Set** these are spans (Diagram (8.9)) and cospans (Diagram (8.15)). We generalise these notions to many categories by introducing now the notions of *limits* and *colimits* as universal constructions. Moreover, as we shall see, limits and colimits represent only one of the possible universal constructions.

Limit

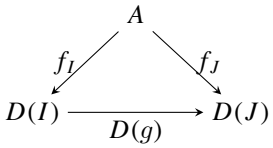
We can affirm that products (8.11), pullbacks (8.14) and equalisers (8.21) are all *limits* while coproducts (8.16), pushouts (8.18) and coequalisers (8.23) are all *colimits*. First, each of these constructions starts with some objects and some maps. Limits in a preorder are meets, colimits in a preorder are joins. Limits and colimits also exist for database instances and monoid actions, allowing us to discuss for example the product or union of different state machines (see Figs. 8.7 and 8.8). Limits and colimits exist for spaces, giving rise to products and unions, as well as quotients. However, limits and colimits do not exist in every category; but when a category $\mathcal{C}$ is complete with respect to limits (or colimits), these limits always seem to mean something valuable to human intuition. For example products, unions, equivalence relations and so on are classical ideas in set theory that are naturally captured by limits and colimits in **Set**.

We can summarise this by saying that there is a common category $\mathcal{L}$ and a functor $\mathcal{L} \longrightarrow \mathcal{C}$ to a category $\mathcal{C}$. Note that $\mathcal{L}$ is a *small* category. If $\mathcal{C}$ is a category and $\mathcal{L}$ is small, having the map $D : \mathcal{L} \longrightarrow \mathcal{C}$ as a diagram in $\mathcal{C}$, then a *cone* on D is an object $A \in \mathcal{C}$ (the vertex of the cone) together with the family of maps on $\mathcal{C}$

$$\left(A \xrightarrow{f_I} D(I)\right)_{I \in \mathcal{L}}, \tag{8.56}$$

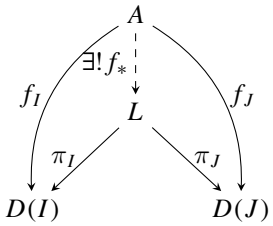
[14] Leinster (2014, Chap. 5).

such that for all maps $I \xrightarrow{g} J \in \mathcal{L}$, the triangle in $\mathcal{C}$



(8.57)

commutes. Then, collecting the results (8.56), (8.57), the limit L of D is the cone



(8.58)

where the maps π_I, π_J are the projections of the limit. In other words, $\forall A \in \mathcal{C}$, the map $A \longrightarrow L$ corresponds one-to-one with cones on D with vertex A. The cone expresses all of the three limits (products, pullbacks, and equalisers). The cone (8.58) has the property that for any form (8.56) there exists a unique map (whose uniqueness is denoted by the symbol $\exists!$) $f_* : A \longrightarrow L$ such that, $\forall I \in \mathcal{L}$, we have

$$\pi_I \circ f_* = f_I. \tag{8.59}$$

Note that the function f_* is called the *transpose* of f and transposition is self-inverse ($f_{**} = f$).

Colimit

The notion of *colimit* is defined as follows. Consider again the category $\mathcal{C}$ and the small category $\mathcal{L}$. Let $D : \mathcal{L} \longrightarrow \mathcal{C}$ be a diagram in $\mathcal{C}$, and write D^{op} for the corresponding functor between the dual or opposite categories $\mathcal{L}^{\text{op}}, \mathcal{C}^{\text{op}}$, as follows:

$$\mathcal{L}^{\text{op}} \longrightarrow \mathcal{C}^{\text{op}}. \tag{8.60}$$

We can, in fact, distinguish between two different kinds of functors between categories: the so-called covariant functors and the so-called contravariant functors. *Covariant* functors are what we have used up to now. Note that every category $\mathcal{C}$ has an *opposite* or dual category $\mathcal{C}^{\text{op}}$ with the same objects but inverse arrows. In other words, if

$$A \xrightarrow{f} B \xrightarrow{g} C, \tag{8.61}$$

are maps in $\mathcal{C}$, then

$$A \xleftarrow{f} B \xleftarrow{g} C \text{ or } A \xrightarrow{f_*} B \xrightarrow{g_*} C \tag{8.62}$$

are maps in $\mathcal{C}^{\text{op}}$. Thus, if $\mathcal{C}$ and $\mathcal{D}$ are categories, a *contravariant* functor from $\mathcal{C}$ to $\mathcal{D}$ is a functor $\mathcal{C}^{\text{op}} \longrightarrow \mathcal{D}$.

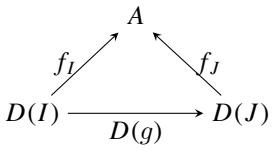
A *cocone* on D is a cone on D^{op}, and a colimit of D is a limit of D^{op}. The cocone on D is an object $A \in \mathcal{C}$ (the vertex of the cocone) together with the family of maps in $\mathcal{C}$

$$\left(D(I) \xrightarrow{f_I} A\right)_{I \in \mathcal{L}}, \tag{8.63a}$$

or in $\mathcal{C}^{\text{op}}$:

$$\left(A \xrightarrow{(f_I)_*} D(I)\right)_{I \in \mathcal{L}}, \tag{8.63b}$$

such that for all maps $I \xrightarrow{g} J \in \mathcal{L}$, the diagram in $\mathcal{C}^{\text{op}}$

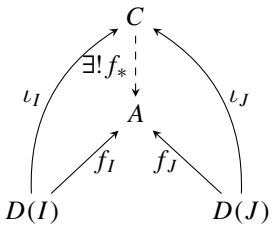


(8.64)

commutes. This cocone has the property that for any form (8.63) on D there exists a unique map $f_* : C \longrightarrow A$, where C denotes the colimit, such that, $\forall I \in \mathcal{L}$, we have

$$f_* \circ \iota_I = f_I, \tag{8.65}$$

which, collecting the results (8.63a) and (8.64), can be displayed by the diagram



(8.66)

Then, we get a mirror image of the previous diagram for cones. Thus, a cocone represents all colimits (coproducts, pushouts and coequalisers). Let $\mathcal{C}$ be a category.

An object $A \in \mathrm{Ob}(\mathcal{C})$ is called *initial* if, for all objects $C \in \mathrm{Ob}(\mathcal{C})$ there exists a unique morphism $A \longrightarrow C$. An object $Z \in \mathrm{Ob}(\mathcal{C})$ is called *terminal*, if for all objects $A \in \mathrm{Ob}(\mathcal{C})$, there is exists a unique morphism $C \longrightarrow Z$. For instance, the empty set is initial in **Set**, while the one-element set is terminal. An object in a category is called *universal* if it is either initial or terminal. Thus, colimits are terminal things of a certain sort, while limits are initial things of a certain sort.

Adjoint Functors

As mentioned, limits and colimits are only one form of universal property. There are also two other notions that are crucial to category theory: that of adjointness and that of representability. Note that everything that can be described with one of these three tools can also be described by using any of the other two. It is like using different coordinate systems. Let us begin with the notion of *adjoint functors* (also known as adjunctions), which are like dictionaries that translate back and forth between different categories.[15] As seen at the end of Sect. 8.1.3, in category theory, we often have two categories that are not on the same conceptual level, and their relations can be particularly interesting.

To be more specific, consider the category of monoids and the category of sets. A monoid $\mathcal{M} = (M, e, *)$ is a set M with an identity element e and a multiplication formula $*$ that is associative (Sect. 8.1.2). A set is just a set. A dictionary between **Mon** and **Set** is required to set up an exchange that is appropriate to the structures at hand. It will be in the form of two functors, one is denoted by $L : \mathbf{Set} \longrightarrow \mathbf{Mon}$, and the other is denoted by $R : \mathbf{Mon} \longrightarrow \mathbf{Set}$. These two adjunctions are called *left* and *right* adjoints, respectively. We may say that L is the left adjoint of R or that R is the right adjoint of L.

Of course, these relations need to be appropriate. Let us consider an analogy. A few months old child can make repeatable noises and an adult can make repeatable noises. One might say "after all, talking is nothing but making repeatable noises." But the adult's repeatable noises are called words, they form sentences, and these sentences can cause even nuclear wars. Thus, there is something more in adult language than simply repeatable sounds. So, we have something analogous to two categories here ((repeated noises) and (meaningful words)). To translate baby talk into adult language, we would make every repeated noise a kind of word, thereby granting it *meaning* (like interpreting a certain sound as meaning distress). We don't know what a given repeated noise should mean, but we give it a slot in our conceptual space. On the other hand, to translate from meaningful words to repeatable noises is easy. We just hear the word as a repeated noise, which is how the baby probably hears it.

This example introduces us to one of the most important distinctions: adjoint functors come here in the form of "free" (left adjoint) and "forgetful" (right adjoint). This is crucial for the notion of emergence (Sect. 7.3.3), since it provides us with a formal tool for dealing with the possible two operations that connect objects at different

[15] Spivak (2013, Sect. 5.1).

levels of complexity: we can either try to reduce an object to its basic components by considering some of its aspects as not relevant or we can consider a complex object as emerging out of a lower level of complexity thanks to additional constraints. What is important here is that there are 'neutral' lower level operations that have no meaning at the upper level, in agreement with our definition of emergence. Moreover, these adjunctions play also an important role in theory of knowledge. In fact, forgetful adjunctions have been traditionally called *abstractions*, while free adjunctions could be called *specifications*. Let us now come back to the previous example. Here, we freely add a child's 'word' to our conceptual space without having any idea of how it adheres to the rest of the child's noises or feelings. But it doesn't act like a sound to us, it acts like a word; even when we cannot figure out what it means, we assume that it means something. This is a specification of meaningless sounds. Conversely, the translation going the other way is "forgetful", by forgetting the meaning of our words and just hearing them as sounds. In fact, the baby hears our words and accepts them as mere sounds, not knowing that there is anything extra to get. This is abstraction since we lose here the characters that make of words a specific subset of sounds (described by linguistics). Note that formal structures have a specific character. Then, free adjunction essentially increases the number of relations among elements without necessarily reducing the number of elements.

Back to sets and monoids, the sets are like the babies from the previous example: they are simple objects full of unconnected dots. The monoids are like adults, forming words and performing actions. In the monoid, each element means something and combines with other elements in some way. There are lots of different sets and lots of different monoids, just as there are many babies and many adults, but there are patterns to the behaviour of each kind and we put them in different categories.

As mentioned, universal concepts, i.e. initial objects and terminal objects, colimits and limits, are easily phrased in the language of adjoint functors. We will say that a functor $F : \mathcal{C} \longrightarrow \mathcal{D}$ is a left adjoint if there exists a functor $G : \mathcal{D} \longrightarrow \mathcal{C}$ such that F is a left adjoint of G. We have shown that if F is a left adjoint of some functor G, then it is isomorphic to every other left adjoint of G, and G is isomorphic to every other right adjoint of F. The useful fact about adjunctions is that left adjoints preserve all colimits and right adjoints preserve all limits.

Representable Functors

Let us now discuss *representable* functors, namely the third approach to the idea of universal property.[16] We have already somehow introduced this notion when dealt with different schemas of data in Sect. 8.1.3 and when discussed different action models in Sect. 8.1.5. Let $\mathcal{C}$ be a locally small category (in which some classes are in fact sets). We define a functor (where $-$ denotes a placeholder, into which arguments can be inserted)

$$H^A = \mathcal{C}(A, -) : \mathcal{C} \longrightarrow \mathbf{Set} \tag{8.67}$$

[16] Leinster (2014, Chap. 4).

defined as follows:

- For objects $B \in \mathcal{C}$, put $H^A(B) = \mathcal{C}(A, B)$, where $\mathcal{C}(A, B)$ denotes the set (or class) of all maps $A \longrightarrow B$;
- For maps $B \xrightarrow{g} B'$ in $\mathcal{C}$, define

$$H^A(g) = \mathcal{C}(A, g) : \mathcal{C}(A, B) \longrightarrow \mathcal{C}(A, B') \tag{8.68}$$

by $p \longmapsto g \circ p, \forall p : A \longrightarrow B$.

Now, if $\mathcal{C}$ is a locally small category, a functor $F : \mathcal{C} \longrightarrow \mathbf{Set}$ is *representable* if $F \cong H^A$ for some $A \in \mathcal{C}$. A *representation* of F is a choice of an object $A \in \mathcal{C}$ and an isomorphism between H^A and F. Note that only so-called set-valued functors (i.e. functors with codomain **Set**) can be representable.

For each object A of a category $\mathcal{C}$, the functor $H^A \in [\mathcal{C}, \mathbf{Set}]$ describes how A sees the world. As A varies, this view varies as well. On the other hand, since it is always the same world being seen, different views from different objects are somehow related. Thus, the family $(H^A)_{A \in \mathcal{C}}$ of views has some consistence to it, what means that, whenever there is map between objects A, A', there is also a map between H^A and $H^{A'}$. In other words, the map $A' \xrightarrow{f} A$ induces a (natural) transformation H^f such that

$$\begin{array}{ccc} & H^A & \\ \mathcal{C} & \Downarrow H^f & \mathbf{Set} \\ & H^{A'} & \end{array} \tag{8.69}$$

whose B-component (for $B \in \mathcal{C}$) is the function

$$\begin{aligned} H^A(B) = \mathcal{C}(A, B) &\longrightarrow H^{A'}(B) = \mathcal{C}(A', B) \\ p &\longmapsto p \circ f. \end{aligned} \tag{8.70}$$

We come back here to the notion of aspect (Sect. 8.1.2): the properties that we attribute to things are in fact aspects that they present to "us" according to the different contexts and interactions. What is crucial of this view is that it is both *relational* and *objective* as far as there is a congruence among these different views (Sects. 3.2.3 and 5.2.6).

The three approaches to universality (limits and colimits, adjunctions and representations) build an important bridge between formal aspects and empirical sciences. In fact, these approaches are formal but hinting at something more:

- The first one, based on limits and colimits, is based on fundamental structural (and horizontal) equivalences among categories together with their initial and terminal objects, and can therefore be called the *logical* approach.
- The universal construction based on adjunctions deals with the (vertical) relations among different domains of knowledge (and disciplines) and can therefore be called the *ontological* approach.

- Finally, the representational construction of universality deals with the different ways (and contexts) in which we could consider the same object and their relations; therefore, it could be called the *epistemological* approach.

These three different approaches are also interrelated in the sense that there are mappings from one to another. This sets logic, ontology and epistemology on a common ground. These are the issues that will be developed further in the present chapter.

8.1.5 Natural Transformations

Definition of Natural Transformation

Category theory was originally invented to deal with natural transformations.[17] These were sufficiently conceptually challenging to require a specific formalisation and thus the invention of category theory. If we think of categories as domains (of discourse, interaction, comparability, etc.) and of functors as transformations between different domains, natural transformations compare different transformations, that is, different functors.

Let us introduce the formal definition. Let $\mathcal{C}$, $\mathcal{D}$ be categories and let $F : \mathcal{C} \longrightarrow \mathcal{D}$ and $G : \mathcal{C} \longrightarrow \mathcal{D}$ be functors. A *natural transformation* α from F to G, which requires that these functors have the same domain and the same codomain, is denoted by

$$\alpha : F \longrightarrow G \tag{8.71}$$

and is defined as follows. One announces some constituents:

- For each object $C \in \mathrm{Ob}(\mathcal{C})$, a morphism $\alpha_C : F(C) \longrightarrow G(C)$ in $\mathcal{D}$, called the C-component of α.

Moreover, one asserts that the following law holds:

- For every morphism $h : C \longmapsto C'$ in $\mathcal{C}$, the following square (called the *naturality* square for h) must commute:

$$\begin{array}{ccc} F(C) & \xrightarrow{\alpha_C} & G(C) \\ {\scriptstyle F(h)}\downarrow & & \downarrow{\scriptstyle G(h)} \\ F(C') & \xrightarrow{\alpha_{C'}} & G(C') \end{array} \tag{8.72}$$

Suppose now that there is a path $p : C_0 \xrightarrow{f_1} C_1 \xrightarrow{f_2} \ldots \xrightarrow{f_n} C_n$. Suppose further that $\forall\, 1 \leq i \leq n$ the naturality square commutes:

[17] Spivak (2013, Sect. 4.3).

Table 8.2 Original model X

ID	a	b
State 0	State 1	State 2
State 1	State 2	State 1
State 2	State 0	State 0

$$\begin{array}{ccc} F(C_{i-1}) & \xrightarrow{\alpha_{C_{i-1}}} & G(C_{i-1}) \\ {\scriptstyle F(f_i)}\downarrow & & \downarrow{\scriptstyle G(f_i)} \\ F(C_i) & \xrightarrow{\alpha_{C_i}} & G(C_i) \end{array} \tag{8.73}$$

Then, the naturality square for the composite path p also commutes:

$$\begin{array}{ccc} F(C_0) & \xrightarrow{\alpha_{C_0}} & G(C_0) \\ {\scriptstyle F(p)}\downarrow & & \downarrow{\scriptstyle G(p)} \\ F(C_n) & \xrightarrow{\alpha_{C_n}} & G(C_n) \end{array} \tag{8.74}$$

An Example

Let us give an example. Suppose a functor List $\circ$ List : **Set** $\longrightarrow$ **Set** that sends a set of lists of lists of Y to the set of lists of lists in X. Now a natural transformation is the following. Using a function $f : X \longrightarrow Y$, where X, Y are sets, (i) to convert a list of lists of X's into a list of Y's thanks to the functor List $\circ$ List (f), and then (ii) concatenating to get a simple list of Y 's thanks to the natural transformation α_Y does the same thing as (i) concatenating our list of lists of X's into a simple list of X's thanks to the natural transformation α_X and then (ii) using the functor List(f) to convert it into a list of Y's. Suppose that $X = \{a, b, c\}$ and $Y = \{1, 2\}$, and let $f : X \longrightarrow Y$ assign $f(a) = f(b) = 1$, $f(c) = 2$. Let us start with a list of X's like $[[a, b], [a, c, a, b, c], [c]]$ and introduce the natural transformation α_X : List $\circ$ List$(X) \longrightarrow$ List(X) (and similarly for Y). Then we have

$$\begin{array}{ccc} [[a,b],[a,c,a,b,c],[c]] & \xmapsto{\alpha_X} & [a,b,a,c,a,b,c,c] \\ \Big\downarrow{\scriptstyle \text{List}\circ\text{List}(f)} & & \Big\downarrow{\scriptstyle \text{List}(f)} \\ [[1,1],[1,2,1,1,2],[2]] & \xmapsto{\alpha_Y} & [1,1,1,2,1,1,2,2] \end{array} \tag{8.75}$$

Table 8.3 New model Y

ID	a	b
State 0	State 1A	State 2A
State 1A	State 2A	State 1B
State 1B	State 2B	State 1C
State 1C	State 2B	State 1B
State 2A	State 0	State 0
State 2B	State 0	State 0

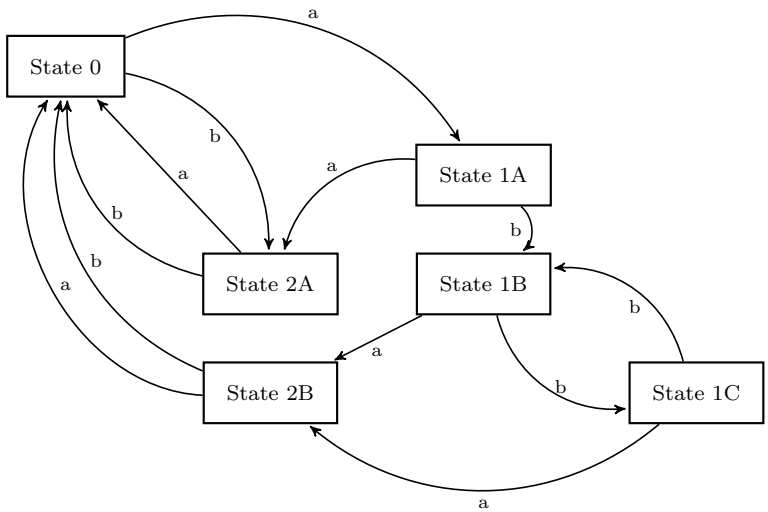


Fig. 8.8 Let us call X the original model displayed in Fig. 8.7 and Y the refined one above. The crucial question is: are these models comparable? Adapted from Spivak (2013, p. 148)

Second Example

A further example is the following. I have introduced the example displayed in Fig. 8.7 and Table 8.1. An interesting situation arises when somebody has a more refined model of the same machine. Suppose that we like to introduce this new model, i.e. Y, as in Fig. 8.8 and compare it with the previous one, namely X. Looking at Table 8.3 for Y, it is evident that States 1A, 1B, 1C of Model Y are expansions of State 1 of Model X as well as States 2A, 2B are of State 2, so that, if one removes the distinction between States 1A, 1B, 1C as well as between States 2 A and 2B, one goes back to Table 8.2 for X. The table for Y is more specific, but it is fully compatible with table X. The sense in which it is compatible is precisely the sense defined by there being a natural transformation.

Now, let us check whether the following two squares (representing homomorphisms a, b) commute:

Table 8.4 Naturality table for a : ▲ ⟶ ▲. The first column coincides with the first column of Table 8.3, the second column with the second column of Table 8.3 and is the result of the transformation effected in the left diagram (8.76) by the left vertical arrow, the third column is the result of the transformation (always on the states of the first column) effected in the same diagram by the sequence left vertical arrow–bottom horizontal arrow (note that this result coincides with Column a of the table for X), the fourth column shows the results of the transformation performed by the top horizontal arrow and the fifth and last column is the result of the sequence top horizontal arrow–right vertical arrow that brings to the same result of the other sequence

$Y(\blacktriangle)$ (ID)	Y(a)	$\alpha_\blacktriangle \circ Y$(a)	$\alpha_\blacktriangle$	X(a) $\circ \alpha_\blacktriangle$
State 0	State 1A	**State 1**	State 0	**State 1**
State 1A	State 2A	**State 2**	State 1	**State 2**
State 1B	State 2B	**State 2**	State 1	**State 2**
State 1C	State 2B	**State 2**	State 1	**State 2**
State 2A	State 0	**State 0**	State 2	**State 0**
State 2B	State 0	**State 0**	State 2	**State 0**

$$\begin{array}{ccc} Y(\blacktriangle) & \xrightarrow{\alpha_\blacktriangle} & X(\blacktriangle) \\ {\scriptstyle Y(\mathrm{a})}\downarrow & & \downarrow{\scriptstyle X(\mathrm{a})} \\ Y(\blacktriangle) & \xrightarrow{\alpha_\blacktriangle} & X(\blacktriangle) \end{array} \qquad \begin{array}{ccc} Y(\blacktriangle) & \xrightarrow{\alpha_\blacktriangle} & X(\blacktriangle) \\ {\scriptstyle Y(\mathrm{b})}\downarrow & & \downarrow{\scriptstyle X(\mathrm{b})} \\ Y(\blacktriangle) & \xrightarrow{\alpha_\blacktriangle} & X(\blacktriangle) \end{array} \tag{8.76}$$

where $\alpha_\blacktriangle$ is a natural transformation from Y to X while $X(a), X(b), Y(a), Y(b)$ are functors. Moreover, $X(\blacktriangle), Y(\blacktriangle)$ represent the possible states of the set ▲ in the models X, Y, respectively, as displayed in Tables 8.4 and 8.5: the columns that need to match under these natural transformations are in bold. We should check that for every morphism a similar diagram commutes, but this holds automatically: we can put all this stuff together in a single totally commuting diagram (where now the natural transformations $\alpha_\blacktriangle$ are represented by the vertical arrows):

$$\begin{array}{ccccc} Y(\blacktriangle) & \xrightarrow{Y(\mathrm{a})} & Y(\blacktriangle) & \xrightarrow{Y(\mathrm{b})} & Y(\blacktriangle) \\ {\scriptstyle \alpha_\blacktriangle}\downarrow & & \downarrow{\scriptstyle \alpha_\blacktriangle} & & {\scriptstyle \alpha_\blacktriangle}\downarrow \\ X(\blacktriangle) & \xrightarrow{X(\mathrm{a})} & X(\blacktriangle) & \xrightarrow{X(\mathrm{b})} & X(\blacktriangle) \end{array} \tag{8.77}$$

Some of these transformations are displayed in Table 8.6. Obviously, similar diagrams and tables can be written by considering the inverse transformation $\alpha_\blacktriangle^{-1}$ from X to Y (which in fact was implicitly taken into account).

The fundamental lesson here is that we can build different models of our physical reality (according to the general notion of representation introduced in the previous subsection, which already involved the notion of natural transformation). Some are more elementary (reductionist), others more fine-tuned for dealing with more complex (emergent) realities (Sect. 7.3).

Table 8.5 Naturality table for b : ▲ ⟶ ▲. The conventions are the same as for the previous table, but this time making reference to the right diagram (8.76)

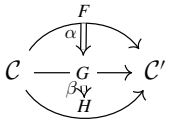
$Y(\blacktriangle)$ (ID)	Y(b)	$\alpha_{\blacktriangle} \circ Y$(b)	$\alpha_{\blacktriangle}$	X(b) $\circ \alpha_{\blacktriangle}$
State 0	State 2A	**State 2**	State 0	**State 2**
State 1A	State 1B	**State 1**	State 1	**State 1**
State 1B	State 1C	**State 1**	State 1	**State 1**
State 1C	State 1B	**State 1**	State 1	**State 1**
State 2A	State 0	**State 0**	State 2	**State 0**
State 2B	State 0	**State 0**	State 2	**State 0**

Table 8.6 Naturality table for some transformations of diagram (8.77). The three columns on the left after the first (ID) column and the three columns on the right bring to the same result through different paths

$Y(\blacktriangle)$ (ID)	Y(a)	Y(b) $\circ Y$(a)	$\alpha_{\blacktriangle} \circ Y$(b) $\circ$ Y(a)	$\alpha_{\blacktriangle}$	X(a) $\circ \alpha_{\blacktriangle}$	X(b) $\circ$ X(a) $\circ \alpha_{\blacktriangle}$
State 0	State 1A	State 1B	**State 1**	State 0	State 1	**State 1**
State 1A	State 2A	State 0	**State 0**	State 1	State 2	**State 0**
State 1B	State 2B	State 0	**State 0**	State 1	State 2	**State 0**
State 1C	State 2B	State 0	**State 0**	State 1	State 2	**State 0**
State 2A	State 0	State 2A	**State 2**	State 2	State 0	**State 2**
State 2B	State 0	State 2A	**State 2**	State 2	State 0	**State 2**

Vertical and Horizontal Composition

Now, we discuss two types of compositions for natural transformations. The terms vertical and horizontal are used to describe them. First, let us consider *vertical composition*:

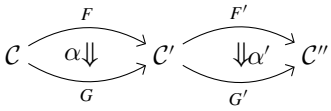


(8.78a)

This can be also expressed by the commutative diagram:

$$\begin{array}{ccc} F & \xrightarrow{\alpha} & G \\ {\scriptstyle\alpha}\downarrow & & \downarrow{\scriptstyle\beta} \\ G & \xrightarrow{\beta} & H \end{array} \tag{8.78b}$$

Quite different is the case of *horizontal composition*:



(8.79a)

We can reformulate the above diagram by means of the following commuting square:

$$\begin{array}{ccc} F' \circ F & \xrightarrow{\alpha} & F' \circ G \\ {\scriptstyle \alpha'}\downarrow & & \downarrow{\scriptstyle \alpha'} \\ G' \circ F & \xrightarrow{\alpha} & G' \circ G \end{array} \tag{8.79b}$$

We can summarise these two kinds of composition by writing

$$\beta \circ \alpha : F \longrightarrow H \text{ and } \alpha' \circ \alpha = \alpha \circ \alpha' : F' \circ F \longrightarrow G' \circ G \tag{8.80}$$

for vertical and horizontal composition, respectively.

Let $\mathcal{C}$ and $\mathcal{D}$ be categories. There exists a category, called the *category of functors* from $\mathcal{C}$ to $\mathcal{D}$ and denoted $\mathrm{Fun}(\mathcal{C}, \mathcal{D})$ or in short $[\mathcal{C}, \mathcal{D}]$, whose objects are the functors $\mathcal{C} \longrightarrow \mathcal{D}$ and whose morphisms are the natural transformations

$$\mathrm{Hom}_{\mathbf{Fun}}(\mathcal{C}, \mathcal{D})(F, G) = \{\alpha : F \longrightarrow G | \alpha = \mathrm{NT}\}\,, \tag{8.81}$$

where NT denotes "natural transformation". In other words, being $[\mathcal{C}, \mathcal{D}]$ a category, in agreement with definitions (8.24)–(8.26), there are identity natural transformations and natural transformations can be composed, so that the identity and associativity laws hold also for natural transformations.

Equivalence of Categories

Our purpose now is to explain the "good notion" of sameness for categories, which appropriately take natural transformations into account. Instead of "functors going both ways with round trips equal to identity", which is required in order to be an isomorphism of categories, we adopt equivalence of categories which demands "functors going both ways with round trips isomorphic to identity". In fact, as mentioned, isomorphism is stronger than equivalence as far as it supposes not only equivalence among counterparts but also several relations among the elements of the two classes. Let $\mathcal{C}$ and $\mathcal{C}'$ be categories. A functor $F : \mathcal{C} \longrightarrow \mathcal{C}'$ is called an *equivalence of categories*, if there exists a functor $F' : \mathcal{C}' \longrightarrow \mathcal{C}$ and natural isomorphisms

$$\alpha : 1_{\mathcal{C}} \xrightarrow{\cong} F' \circ F \text{ and } \alpha' : 1_{\mathcal{C}'} \xrightarrow{\cong} F \circ F', \tag{8.82}$$

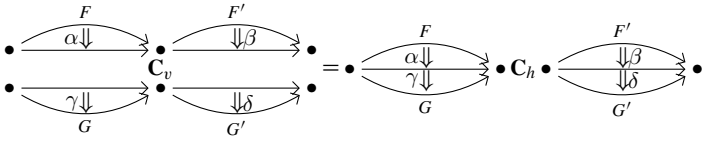
which ensure us of the correct "round trip".

2-Categories

Thus, categories are often similar enough to be considered equivalent without being isomorphic. The reason is that, *in* the framework of a category, there are (classes of) *objects* and (classes of) *morphisms*, whereas when we establish relations *among* categories, we have *categories* and *functors*, plus *natural transformations*. The latter serve as mappings between mappings, and this is not part of the structure of an ordinary category. This whole subject leads us to the study of 2-categories (or n-categories, or ∞-categories), which goes further than the scope of this sketchy presentation. I limit myself to remark that, while an ordinary (1-)category has objects and morphisms, a 2-category generalises this by also including 2-morphisms between the 1-morphisms. An example has already been provided. In fact, the category of small categories and functors **Cat** is actually a 2-category with natural transformations as its 2-morphisms. Also with vertical–horizontal compositions (8.78), (8.79) we have already shifted towards 2-categories. In fact, if we denote by $\mathbf{C}_h$, $\mathbf{C}_v$ horizontal and vertical compositions, respectively of natural transformations $\alpha, \beta, \gamma, \delta$, we have a kind of De Morgan law (Eq. (3.130))

$$(\alpha \mathbf{C}_h \beta)\mathbf{C}_v(\gamma \mathbf{C}_h \delta) = (\alpha \mathbf{C}_v \gamma)\mathbf{C}_h(\beta \mathbf{C}_v \delta), \tag{8.83a}$$

which is graphically represented by



(8.83b)

where the bullets are 1-categories (while objects are called 0-cells), functors F, F', G, G' that are mappings between 0-cells are called 1-cells, and natural transformations that are mappings between 1-cells are called 2-cells.

8.1.6 Category Theory and Quantum Information

A category has objects and morphisms, which represent things and ways to go between things. In physics, the objects are physical systems, and the morphisms are processes turning a state of one physical system into a state of another or the same system. This establishes a fruitful analogy.

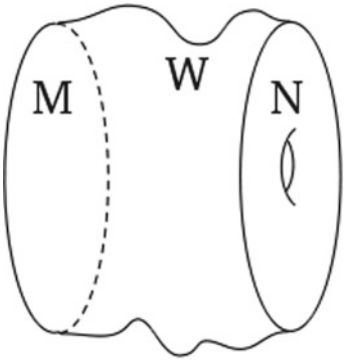


Fig. 8.9 Example of cobordism: a (squeezed) can. Cobordism between manifolds M and N is a compact manifold W whose boundary is the disjoint union of M and N. Adapted from https://en.wikipedia.org/wiki/Cobordism

Cobordism

In fact, in the 1980s, it became clear that underlying Feynman's diagrams (Sect. 7.1.4), there is a powerful analogy between quantum physics and topology![18] (Do not forget that topology played a major role in the building of category theory.) Namely, a linear operator (involved in Feynman's diagrams) behaves very much like a *cobordism*: two manifolds (on this concept see Sect. 2.3.2) of the same dimension are cobordant if their disjoint union is the boundary of a compact manifold one dimension higher, as displayed in Fig. 8.9. The category that has $(n-1)$-dimensional manifolds as objects is called n**Cob**. In fact, diagram (7.144), or its inverse (in our space-time), can be reproduced in topological terms as in Fig. 8.10.

Adjoint

In order to deal with QM, we need to reflect in the categorial structure the important notion of Hilbert space and of adjoint in Hilbert spaces. First, we need to introduce the cross product $X \otimes Y$ that express either concomitant or independent action, depending on whether the objects X and Y are correlated or not. Moreover, we need to satisfy the assignment $\mathcal{H} \longmapsto \mathcal{H}^*$ on objects from a Hilbert space to its dual (Sect. 1.2.1), which 'hosts' both the complex conjugate and the transposed of the vectors $\in \mathcal{H}$. It is not my aim to enter in a complex matter but only to give a sketch of this formalism.[19]

Thus, we can define a *dagger category*, which is a category $\mathcal{C}$ equipped with an identity-on-objects, contravariant, strictly involutive functor $f^\dagger$ that units complex

[18] Baez and Stay (2011).

[19] For what follows see Abramsky and Coecke (2009).

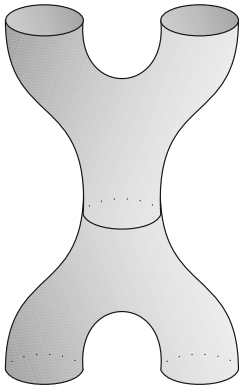


Fig. 8.10 Cobordism representing a Feynman diagram. Adapted from Baez and Stay (2011)

conjugation and transposition

$$f \longmapsto f^{\dagger} = (f^{*})_{*} = (f_{*})^{*}, \tag{8.84}$$

such that we have

$$1^{\dagger} = 1,\ (g \circ f)^{\dagger} = f^{\dagger} \circ g^{\dagger},\ f^{\dagger\dagger} = f. \tag{8.85}$$

Note that these definitions correspond in fact to the notion of adjointness for operators in Hilbert space (see Eq. (1.82)). Moreover, we must have

$$(f \otimes g)^{\dagger} = f^{\dagger} \otimes g^{\dagger}, \tag{8.86}$$

where I recall that the two maps do not express subsequent operations as in the second Eq. (8.85).

Moreover, we define an arrow $f : X \longrightarrow Y$ to be unitary if it is an isomorphism such that—see comments on Eqs. (1.14) and (1.308)—

$$f^{-1} = f^{\dagger}. \tag{8.87}$$

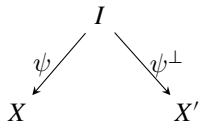
Kets and Bras

Since we are here interested especially in the applications of category theory to quantum information, in the following we shall deal with *finite-dimensional* Hilbert

spaces and their relative category: **FdHilb**.[20] With the tools so far acquired, it is straightforward to interpret Dirac notation. A ket $|\psi\rangle$ is simply an arrow

$$\psi : I \longrightarrow X. \tag{8.88}$$

We think of kets as *states* of a given type of system X. Morphism (8.88) corresponds to a preparation of a system in the specific state $|\psi\rangle$ if, according to Sect. 5.2.2, we write

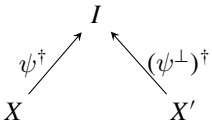


(8.89)

The corresponding bra, which, e.g. represents a detection output (Sect. 5.2.4) will then be

$$\psi^{\dagger} : X \longrightarrow I, \tag{8.90}$$

which gives the diagram



(8.91)

We can think of these objects as a *costates*. This shows that the notion of state is rather a mapping than an object or a character of an object in the classical sense of the word. This explains why, in Sect. 5.2.2, states have been associated with preparations as specific and controlled procedures (mappings) and be said to be equivalence classes of such mappings. Similarly, detections are equivalence classes of interactions (Sect. 5.2.4). Note that the I in the previous equations stands for whatever scalar and, in Eq. (8.89), corresponds to the generation of qubit out of an indeterminate background. I recall in fact that the state previous to preparation is trivial as well as it is the state after detection (Sect. 6.3.1). In other words, we have an object (a physical system), i.e. a type–token, when its dimension, basis vectors spanning its Hilbert space and expansion coefficients are specified. Then, this gives to preparation the meaning of *extracting* a specific qubit from the universal superposition. When we detect, we *send* again the qubit *back* to the universal superposition (thanks to the locally inaccessibility to most of the information) by *retaining a bit locally* thanks to some interaction.

[20] For what follows see Abramsky and Coecke (2009).

Scalar Product

Now, given $\psi, \varphi : I \longrightarrow X$, we define their scalar product $\langle \psi \mid \varphi \rangle$ (which combines Eqs. (8.88) and (8.90)) as a morphism from scalars to scalars

$$\psi^\dagger \circ \varphi : I \longrightarrow I. \tag{8.92}$$

Note that, in this way, the composition $\circ$ expresses the scalar product in this context. This is justified by the fact that we can consider φ as input and ψ as output. Note also that, in analogy with Eq. (1.83), for

$$\psi : I \longrightarrow X,\ \varphi : I \longrightarrow Y,\ f : X \longrightarrow Y, \tag{8.93}$$

we have

$$\langle f \circ \psi \mid \varphi \rangle = \left\langle \psi \mid f^\dagger \circ \varphi \right\rangle, \tag{8.94}$$

since, thanks to Eqs. (8.84) and (8.85), we get

$$\langle f \circ \psi \mid \varphi \rangle = (f \circ \psi)^\dagger \circ \varphi = \psi^\dagger \circ f^\dagger \circ \varphi = \left\langle \psi \mid f^\dagger \circ \varphi \right\rangle, \tag{8.95}$$

where given the associativity of the composition, parentheses have been dropped in the second-last step. It is also obvious that unitary functor $U : X \longrightarrow Y$ preserve the scalar product:

$$\langle U \circ \psi \mid U \circ \varphi \rangle_Y = \psi^\dagger \circ U^\dagger \circ U \circ \varphi = \langle \psi \mid \varphi \rangle_X, \tag{8.96}$$

due to the fact that $U^\dagger \circ U = 1_X : X \longrightarrow Y \longrightarrow X$.

Projectors

It is noticeable of this approach that one is able to delineate a fine structure of bipartite projectors.[21] A *projector* on an object X is an arrow $P : X \longrightarrow X$ which needs to be both idempotent and self-adjoint (see Sect. 1.2.2):

$$P^2 = P,\ \ P = P^\dagger. \tag{8.97}$$

Suppose that we have a state $\psi : I \longrightarrow X$ which is normalised, i.e.

$$\langle \psi \mid \psi \rangle = 1_I. \tag{8.98}$$

Then, the composition

$$\psi \circ \psi^\dagger : X \longrightarrow I \longrightarrow X \tag{8.99}$$

[21] I make here a simplification avoiding the complexities arising from compact closed categories.

represents a ket–bra product and is therefore a projector. Let us now apply this idea to the state

$$\Psi : I \longrightarrow X \otimes Y \tag{8.100}$$

of a compound system. Taking into account that the correspondent bra must be the morphism

$$\Psi^{\dagger} : X \otimes Y \longrightarrow I, \tag{8.101}$$

we can then define a projector for a compound system as

$$P_f := \Psi \circ \Psi^{\dagger} : X \otimes Y \longrightarrow X \otimes Y. \tag{8.102}$$

Note that the state (8.100) must represent an entangled state. Indeed, a product state (like, e.g. a Lüders mixture) needs to have the form

$$\psi \otimes \psi : I \otimes I \longrightarrow X \otimes X. \tag{8.103}$$

Teleportation in Categorial Terms

As an example of how the previous formalism applies to informational protocols, I deal here with teleportation by generalising the results of Sect. 6.3.3. We can logically analyse teleportation as a transmission of a qubit Q from a 'location' 1 to a 'location' 3 thanks to a 'location' 2. Such a location will be explicitly indicated. Let us now introduce a permutation operation Π that act as

$$(X)_1 \otimes (Y)_2 \otimes (Z)_3 \xrightarrow{\Pi_{123}} (Y)_1 \otimes (Z)_2 \otimes (X)_3. \tag{8.104}$$

Then, teleportation can be understood as a swap operation between 'locations' 1 and 3 such that it 'moves' a qubit by leaving a blank behind:

$$(Q)_1 \otimes I_3 \xrightarrow{\sigma_{13}} I_1 \otimes (Q)_3. \tag{8.105}$$

Such an operation can be easily analysed as

$$\begin{array}{ccc} (Q)_1 \otimes I_3 & \xrightarrow{\quad\sigma_{13}\quad} & I_1 \otimes (Q)_3 \\ {\scriptstyle 1_Q \otimes \Psi}\downarrow & & \uparrow{\scriptstyle \Psi^{\dagger} \otimes 1_Q} \\ (Q)_1 \otimes (X)_2 \otimes (Y)_3 & \xrightarrow[\Pi_{123}]{} & (X)_1 \otimes (Y)_2 \otimes (Q)_3 \end{array} \tag{8.106}$$

Note that the permutation operation covers therefore both the change of basis for the three systems and the appropriate unitary operation on qubit 3. Of course, the input entangled state and the measurement outcome can be different, what does not reduce

the validity of this scheme, as shown by the transformations introduced in Sect. 4.4. Thus, we have the sequence: input qubit, addition of an entangled state, change of basis and unitary operation, measurement of systems 1–2 and transfer of qubit Q to location 3.

8.2 Logic Assessment: Boolean Algebra and Quantum Mechanics

If QM represents a general theory of information, this must be cast in logical terms, according to the analysis developed in the two previous sections. Logic is the most general and fundamental formal tool, being the basic category theory. This is so true that, although logic is used everywhere, from mathematics to physics or even practical matters, as a discipline it relies only on its own methods. In other words, it is the only self-sufficient discipline that exists. Thus, we can say that even category theory is a branch of logic. As I shall show, logic has deep connection with information theory, and this connection is rooted in the notion of combinatorics. Logic presents a binary choice (true or false) that perfectly corresponds to the binary codification of information (0 and 1 representing falsity and truth, respectively), and as the latter can be treated as a pure combinatorial calculus.[22] In fact, combinatorics is inscribed in the logical relations themselves. To this purpose, I shall introduce the notion of logical space, which can be considered an extension of the notion of Boolean algebra,[23] as far as it can be treated as a vectorial space and the basic variables that span it as a kind of logical basis. This is particularly relevant for dealing with quantum computation.[24]

8.2.1 Basic Notions

Basic Operations

Let us start with the usual definition of Boolean algebra. Fix, once and for all, a set S, and denote by $\mathcal{B}(S)$ the collection of all subsets of S. Next, let f be a Boolean function, i.e. mapping from $\mathcal{B}(S)$ to the set $\{0, 1\}$ of truth values, satisfying $\forall X, Y \in \mathcal{B}(S)$[25]:

- *Intersection*, denoted by $X \times Y$ or XY, which, $\forall X, Y \in \mathcal{B}$, satisfies

$$f(XY) = 1 \text{ iff } f(X) = f(Y) = 1; \text{ otherwise } f(XY) = 0. \quad (8.107)$$

It is an instance of the product (8.11).

[22] Auletta (2013b).
[23] Boole (1854).
[24] Auletta (2013c).
[25] For a canonical introduction to Boolean algebra see Givant and Halmos (2009).

- *Union*, denoted by $X + Y$, which, $\forall X, Y \in \mathcal{B}$, satisfies

$$f(X + Y) = 0 \text{ iff } f(X) = f(Y) = 0; \text{ otherwise } f(X + Y) = 1. \quad (8.108)$$

  It is an instance of the coproduct (8.16).
- *Complementation*, denoted by X', which, $\forall X \in \mathcal{B}$, satisfies

$$f(X') = 1 \text{ iff } f(X) = 0 \text{ and } f(X') = 0 \text{ iff } f(X) = 1. \quad (8.109)$$

The map f comes therefore to represent truth-value assignment and the numbers 0 (false) and 1 (true) to represent truth values, which allows us to conceive sets and subsets as two-valued variables. Note that I am faithful to Boole's original symbols, for the reason that these operations can also be understood in mathematical terms, e.g. as usual set operations.[26]

In the Boolean algebra $\mathcal{B}(S)$, the binary operations $\times$, $+$ (also called logical sum and product or inclusive disjunction and conjunction) satisfy the following properties:

1. *Idempotency*: $\forall X \in \mathcal{B}$, $X = XX = X + X$. From this it follows that, $\forall X \in \mathcal{B}$, either $f(X) = 1$ or $f(X) = 0$. In fact, suppose that $f(X)$ is not 1. Then, according to Eq. (8.107), $f(XX) = 0$. However, from idempotency, it follows that $f(X) = 0$. Similarly, suppose that, $f(X)$ is not 0. Then, from Eq. (8.108), we have $f(X + X) = 1$. However, according to idempotency, it follows that $f(X) = 1$.
2. *Commutative law*: $\forall X, Y \in \mathcal{B}$, $XY = YX$ and $X + Y = Y + X$.
3. *Associative law*: $\forall X, Y, Z \in \mathcal{B}$, $X(YZ) = (XY)Z$ and $X + (Y + Z) = (X + Y) + Z$ (Eq. (8.25)).

Moreover, the two binary operations are *mutually distributive*, i.e. $\forall X, Y, Z \in \mathcal{B}$,

$$X(Y + Z) = XY + XZ \text{ and } X + YZ = (X + Y)(X + Z). \quad (8.110)$$

Traditional Definition of Boolean Algebra

In order that this collection $\mathcal{B}(S)$ of all subsets of S, together with the operations of union, intersection, and complementation, be a Boolean algebra, there are some requirements. In the literature, seems that there is no general agreement about the number of those requirements or about what are the primitive ones. Here, I shall introduce them in a rather loose way since in the following I shall present another approach that can bypass this problem by simply looking at the *structural* properties of Boolean algebras. First, such a collection must be a POSet (Sect. 8.1.3) by inclusion (*implication*), symbolised as $\rightarrow$, which displays:

- *Reflexivity*: $\forall X \in \mathcal{B}$, $X \rightarrow X$;

[26] See Boole (1854, p. 33).

- *Transitivity*: $\forall X, Y \in \mathcal{B}$, if $XY \to X$ and $X \to X + Y$, then $XY \to X + Y$;
- *Antisymmetry*: $\forall X, Y \in \mathcal{B}$, the two expressions $XY \to Y$ and $Y \to XY$ are logically different and neither implies or is equivalent to the other.

Note that inclusion or implication is a *logical relation* and not the so-called material implication, unfortunately very much used today in logic, which has in fact no logical ground. Indeed, if expressions like $X \to Y$ were admitted (for arbitrary sets X, Y), we would count as a logical proposition a connection between two logically unrelated variables. It is indeed weird to assume that material implication expresses the notion of sufficient condition when the antecedent (here X) is no condition whatsoever of the consequent (here Y). The consequences of this standpoint are well-known: we would run into the so-called logical paradoxes like "If Napoleon won at Waterloo, I was the emperor of China", and similar absurd expressions. I think that to affirm that there are logical paradoxes is the only true paradox here. At the opposite, I am convinced that paradoxes simply do not exist neither in logic nor in any other scientific discipline. "Paradox" may be only a useful term for denoting difficult problems that have not been completely solved at a certain stage of our knowledge, like the 'Schrödinger's cat paradox' (Sects. 2.2.5 and 3.1.4) and similar ones. Everything is fine if we use the notion of *sufficient* (or *necessary*) condition for expressing what the material implication would represent but with the constraint that it needs to be a *condition*, in the same sense in which, e.g. we say that degeneracy of eigenvalues is sufficient condition of commutativity between the involved observables.

On the other hand, the definition of logical implication is rigorous. As we shall see, we shall assign to each set or class of the algebra a logical ID that comprehends all its truth values. Then, the logical ID of the consequent of any logical implication needs to preserve *all* 1s (truths) of the antecedent in the *same* position and in most cases increase their number. In the case in which the number does not increase, antecedent and consequent coincide according to the reflexive property of the logical implication. In other words, implication is a *truth propagator* as functors are conductors of mathematical truth in category theory (Sect. 8.1.2). Thus, implication needs to be a tautology (always true), like the form $XY \to X$. Logical implications build (logical) *propositions*, while the relations among classes (like $X + Y$ or XY) are contingent (can be true but also false) and cannot be considered logical propositions. Thus, we do not admit as logical proposition statements like "Today is a nice day", as it is again the case in current propositional calculus where arbitrary propositions $p, q, r \ldots$ are used (in fact, in logic there is nothing that could be considered arbitrary). Often, these two concepts are confused but need to be kept well distinct.

With the symbol $\sim$ I shall represent here the so-called material equivalence (as in ordinary mathematics and category theory) to be kept distinct from logical equivalence (denoted by $=$) which is also a tautology (expressing again logical relations among classes). I shall maintain material equivalence for its relevance in mathematics. Since logical equivalence individuates the same node (as it happens e.g. for $X(X + Y) = X$), it can be also called logical identity (and this justifies the symbol "$=$") and is a stronger notion than isomorphism. Thus, isomorphism is middle way between logical and material equivalence.

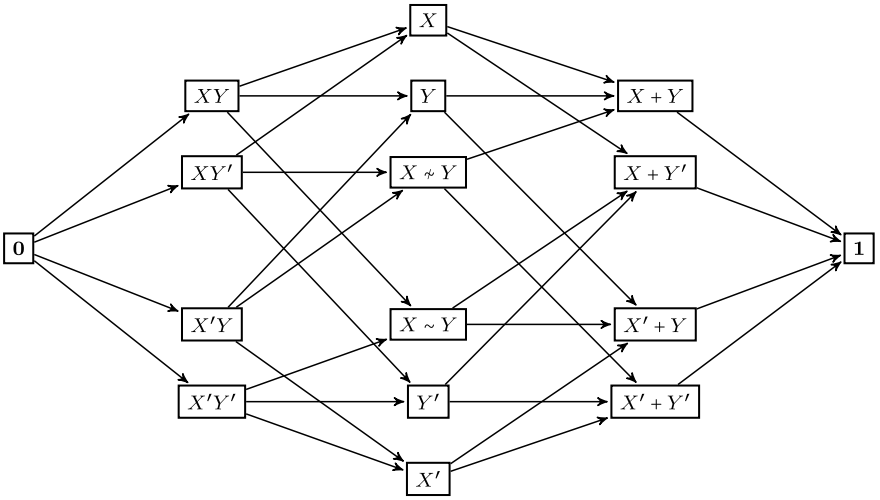


Fig. 8.11 The relations of inclusion with sets X, Y and their complements. The symbol $\nsim$ denotes countervalence, defined $\forall X, Y \in \mathcal{B}$ as $f(X \nsim Y) = 1$ iff either $f(X) = 1$ and $f(Y) = 0$ or $f(X) = 0$ and $f(Y) = 1$. Note that each implication shown here is in fact a tautology, i.e. a constant with truth value 1: e.g. $f(XY \to X) = \mathbf{1}$. In other words, any of these order relations map to the universal set $\mathbf{1}$. Arrows departing from a node are examples of span (8.9), while arrow converging on a node are examples of cospans (8.15)

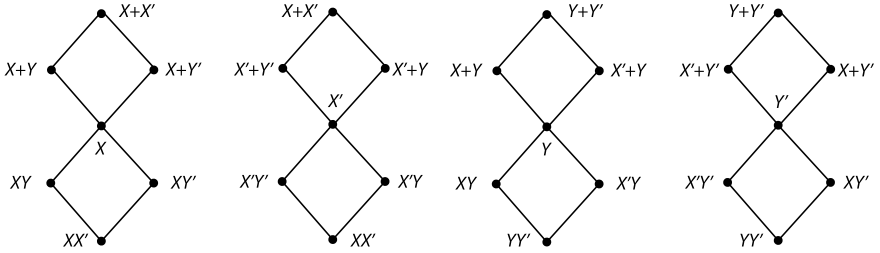


Fig. 8.12 In the 2D algebra, we can select four linear subalgebras individuated by the variables X, X', Y, Y' (or X, Y and their negations). Note in fact that every pair of nodes being on one of the paths satisfies the comparability requirement. Note that the whole structure has the form of a double diamond with the variable representing the joining point between the two

While a Boolean algebra satisfies the three requirements for a POSet, i.e. reflexivity, transitivity and antisymmetry, it is also well-known that it is not a linear POSet,[27] i.e. it does not satisfy

- Comparability: it is not true that $\forall X, Y \in \mathcal{B}$ we have, e.g. either $XY \to Y'$ or $Y' \to XY$.

[27] Spivak (2013, Sect. 3.4).

The nonlinearity of $\mathcal{B}$ can be easily ascertained in the case of two variables by following the implication arrows displayed in Fig. 8.11. The reason for that is represented by the existence of a set of irreducible atomic variables and of their relations in terms of resolution of identity, and issue to be dealt with below. Now, each of such variables selects a subspace in every $n \geq 2$ logical space that is linear if we consider paths that follow either meet (conjunctions) or join (disjunctions): see Fig. 8.12 for an example.[28] (Note that such subalgebras do not represent the $n-1, n-2, \ldots$ proper subalgebras of each n-dimensional algebra: for instance, there are 12 2D and 8 1D subalgebras embedded in the 3D algebra.)[29]

Finally, the Boolean algebra comprehends universal bounds: **1** (the universal set or *supremum*, given by the union of all elements) and **0** (the empty set or *infimum*, given by the intersection of all elements), which satisfy following properties:

$$\mathbf{0}X = \mathbf{0},\ \mathbf{1}X = X,\ \mathbf{0} + X = X,\ \mathbf{1} + X = \mathbf{1}, \tag{8.111}$$

so that they are the identity elements for intersection and union, respectively (Eq. (8.26)). In the language of category theory, **0** is the initial object (limit) while **1** is the terminal object (colimit) (Sect. 8.1.4) A Boolean algebra with the property that every subset of it has both a *supremum* and an *infimum* is called *complete*.[30] The universal set logically represents a tautology (it is always true, i.e. **1** is a constant with value 1), while the empty set represents a contradiction (it is always false, i.e. **0** is a constant with value 0). Being they limit and colimit, respectively, we have that $\forall X \in \mathcal{B}$, $\mathbf{0} \to X$ and $X \to \mathbf{1}$, which allows us to write $\mathbf{0} \to X \to \mathbf{1}$, $\forall X \in \mathcal{B}$, so that, by transitivity, the whole preorder can be synthetically expressed as $\mathbf{0} \to \mathbf{1}$: falsity is sufficient condition of truth or also truth is a necessary condition of falsity. If we have a complete Boolean algebra, the set S of the Boolean algebra turns in fact to be **1**, so that we can write $\mathcal{B}(\mathbf{1})$ or simply $\mathcal{B}$. Here and in the following, I shall deal only with such algebras.

Now, it turns out that such a Boolean logical system is complete in the stronger sense that (in appropriate dimension) contains in itself all logical laws inscribed in its own relations. This is evident since all logical laws are tautologies. Thus, no assumption of logical laws or principles is necessary since, as anticipated, any logical statement is embedded in the structure of the algebra. This is shown by the combinatorial approach.

[28] Auletta (2013b, Chap. 1).

[29] Auletta (2013b, Chap. 8).

[30] Givant and Halmos (2009, p. 45).

Table 8.7 Sum and product of IDs. Note that the value $1 + 1 = 1$ is not determined by a special property of the logical sum of IDs but from the fact that the set of the allowed number is $\{0, 1\}$. Also division (OR NOT) and subtraction (AND NOT) could be defined in similar ways

Sum	$0+0=0$	$0+1=1$	$1+0=1$	$1+1=1$
Product	$0\times 0=0$	$0\times 1=0$	$1\times 0=0$	$1\times 1=1$

8.2.2 Combinatorial Approach

Logical IDs

I shall present now an approach highlighting the structural properties of Boolean algebra.[31] This is based on combinatorics, or better, the combinatorial calculus is inscribed in the fundamental structures of the Boolean algebra. The first scholar to have thought about logic as a combinatorial calculus was G. Leibniz, and this in terms of combinatorics of possibles.[32] The relevance of this approach can be appreciated if we think that everywhere (from QM to LQG, from statistical mechanics to molecular physics) combinatorics is the dominant formal structure.

In fact, each element of the Boolean algebra with n variables, denoted by $\mathcal{B}_n$, can be picked up through pure combinatorial calculus and the structural relations between the elements (nodes) generated in this way suffice to fully characterise the complete Boolean algebra:

Theorem 8.1 $\forall n \in \mathbb{N}$, *all elements of* $\mathcal{B}_n$ *are given by*

$$\sum_{k=0}^{2^n} \binom{2^n}{k} = 2^{\sum_{j=0}^{n} \binom{n}{j}}. \tag{8.112}$$

In fact, $\forall \mathcal{B}_n$ we have 1 element with zero 1s and 2^n 0s (**0**), a number 2^n of all possible permutations with a single 1, a number $\binom{2^n}{2}$ of all possible permutations with two 1s and so on up to a single element with 2^n 1s (**1**). This allows us to identify each element of $\mathcal{B}_n$ by a string of 0s and 1s with length 2^n. This is the *logical ID* (in short, ID) of the element, and each of the values $0 \leq k \leq 2^n$ determines the number of 1s present in the ID.

In such a way, the combinations of the different elements can be numerically computed (column by column) on the IDs thanks to sum and product (Table 8.7). For instance (for $\mathcal{B}_3$, with ID length $2^3 = 8$),

[31] Auletta (2013b, 2015c).

[32] Leibniz (1666).

$$\begin{array}{llc} 0011 & 0101 & + \\ 1000 & 0001 & = \\ 1011 & 0101 & \end{array}$$

and

$$\begin{array}{llc} 0011 & 0101 & \times \\ 1000 & 0001 & = \\ 0000 & 0001 & \end{array}$$

Note that often I partition the logical ID in segments (chunks) composed of four numbers for making the visualisation easier. Often, elements with a single 1 in finite Boolean algebras are considered as atoms.[33] However, this is little bit misleading since they could be better called the smallest classes of any nD algebra (with n finite) after the *infimum*. In the infinite case, they cannot be picked up with any finite means since the ID of any element of the algebra is infinitely long.

Logical Operations Again

Dealing with IDs has many advantages. First, idempotency, commutation, association and distribution can be proved by simply summing and multiplying IDs. In fact, if XY is true, also X and Y are true, thus are logical consequences of that expression. This is due to the fact that the IDs of both X and Y contain the ID of XY (where in the latter are 1s, also in the former two there are 1s). The rule of *disjunctive syllogism* is proved in this way. If $X' + Y$, whose ID is 1111 0011 is true, we have that $X(X' + Y)$ has ID 0000 0011, which we know to be the ID of XY. But, as said, Y is a logical consequence of XY.

Moreover, IDs allow us to define two new operations (morphisms) that will be crucial for the following:

- The *reversal* operation, which reverse the ID; e.g (for $\mathcal{B}_3$) $(10001101)^{-1} = 10110001$. Such an operation transforms, e.g. XY into $X'Y'$ as well as $X + Y$ into $X' + Y'$. In other words, reversing a statement abrogates the latter. These expressions are also called controdual of each other.[34] Note that

$$(XY)(XY)^{-1} = \mathbf{0} \text{ and } (X + Y) + (X + Y)^{-1} = \mathbf{1}. \tag{8.113}$$

- The *neg-reversal* operation (it does not matter the order of negation and reversal): $(10001100)^{\dagger} = 11001110$. In other words, $\forall X$, we have

$$X^{\dagger} = (X')^{-1} = (X^{-1})', \text{ which implies } (X^{\dagger})^{-1} = (X^{-1})^{\dagger} = X'. \tag{8.114}$$

[33] Givant and Halmos (2009, pp. 117 and 127).

[34] Givant and Halmos (2009, Chap. 4).

Such an operation transforms, e.g. XY into $X + Y$ and vice versa. These expressions are also called the dual of each other. In other words, as complementation, both two applications of reversal and neg-reversal operations bring back to the initial element (and so their compositions). They correspond therefore to the inversion in group theory. Note that we have

$$(XY)(XY)^{\dagger} = XY \text{ and } (XY) + (XY)^{\dagger} = X + Y. \tag{8.115}$$

Self-dual Variables

I have spoken of basic variables. It is time now to introduce them formally. In fact, $\forall n \in \mathbb{N}$, there is a set of variables $\in \mathcal{B}_n$ whose ID is divided in two equal halves whose the first is the neg-reversal of the second. In other words, they are the neg-reversal of themselves ($X = X^{\dagger}$). Examples (for $\mathcal{B}_3$) are: 0000 1111, 0011 0011, 1001 0110. Let us call these *self-dual* variables, and denote their set as $\mathrm{Self}(n)$. Since such variables are the neg-reversal of themselves, this means that, $\forall X \in \mathrm{Self}(n)$, we have $X' = X^{-1}$, as it is evident by looking at the IDs listed previously.

It is also easy to see that, $\forall n \in \mathbb{N}$, all elements of $\mathcal{B}_n$ can be derived from some combination of the self-dual variables. Thus, of all possible collections of objects $\subset \mathcal{B}_n$, we consider self-dual variables (like X or Y) as representing sets denoting general properties, while their combinations (expressing different kinds of relations among sets, like XY or $X + Y$) as classes. Thus, if mathematics studies the possible relations among objects,[35] logic studies the relations among *classes* of objects.[36] Since, however, sets can be considered zero-level classes, in the following I shall often use the term *class* for denoting both sets and classes in strict sense. It can be proved that, for each $\mathcal{B}_n$, the number of those sets (and their complements) is equal to the number of all collections of objects (both sets and classes) of $\mathcal{B}_{n-1}$. In fact, self-dual variables are identified by half the sequence of their binary ID. An immediate consequence is that, for each $\mathcal{B}_n$ ($n \in \mathbb{N}$), with $k(n) = 2^{m(n)}$ (where $m(n) = 2^n$) collections of objects, the number of these variables is $k(n-1)$, with $m(n-1) = 2^{n-1}$.

As mentioned previously, each $\mathcal{B}_n$ contains several complete subalgebras $\mathcal{B}_m$, with $m < n$. For instance, $\mathcal{B}_2$ contains two $\mathcal{B}_1$ complete subalgebras; $\mathcal{B}_3$ contains eight $\mathcal{B}_1$ complete subalgebras and twelve $\mathcal{B}_2$ complete subalgebras (six couples for each spanning set displayed in Table 8.11); and so on. These subalgebras are complete because each of them contains the *supremum* and the *infimum*. For any fixed m, these subalgebras areisomorphic to each other across any $\mathcal{B}_n$ algebra, with $m < n$.

[35] Poincaré (1902, p. 49).

[36] Poincaré tells us that formal logic is nothing else than the study of the properties that are common to any classification (Poincaré 1909, p. 9).

Moreover, all of the $\mathcal{B}_m$ so defined are ideal subalgebras, where with *ideal* is meant a subset of any $\mathcal{B}_n$ algebra such that[37]

- $\mathbf{0} \subset \mathcal{B}_m$,
- If $X \subset \mathcal{B}_m$ and $Y \subset \mathcal{B}_m$, then $X + Y \subset \mathcal{B}_m$,
- If $X \subset \mathcal{B}_m$ and $Y \subset \mathcal{B}_m$, then $XY \subset \mathcal{B}_m$.

Numbers Determining the Algebra

I have already introduced some fundamental numbers that characterise any Boolean algebra. I summarise them and add some other ones:

- n is the number of basic (self-dual) variables that determine the dimension of the algebra. As we shall see, there can be different alternative sets of such variables.
- $m = 2^n$ is the number of truth-value assignments that determine the truth table (as well as the ID length).
- $k = 2^m$ is the overall number of collections of objects that can be generated in such an algebra through relations among the n variables. This number is given by the laws of combinatorics, i.e. analogously to Eq. (8.112), we have

$$k = \sum_{x=1}^{m} \binom{m}{x} = 2^{\sum_{y=0}^{n} \binom{n}{y}}, \tag{8.116}$$

  where

$$m = \sum_{y=0}^{n} \binom{n}{y}. \tag{8.117}$$

  It is evident that for $n = 2$, we have $k = 16$, for $n = 3$, we have $k = 256$ and so on.
- Another important number is that of the levels of the algebra. All the nodes of the same level have the same number of 0s and 1s (differently permuted). The number l of levels is given by $l = m + 1$.
- Since the Boolean algebra is in fact a POSet, it builds a network where several nodes are connected through logical implication, which turns out to be a graph. Thus, the number r of logical relations (logical implications) that can be constituted among all nodes of the network is given by

$$r(n) = \sum_{t=0}^{m-1} (m - t) \binom{m}{m - t}, \tag{8.118}$$

[37] See Givant and Halmos (2009, pp. 149–150).

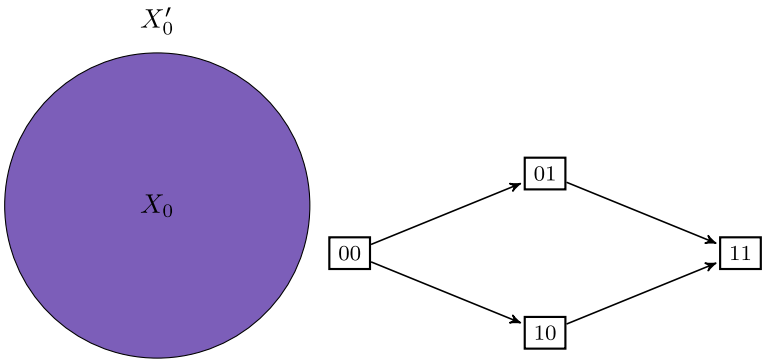


Fig. 8.13 Left panel: the set X_0 and its complement X_0'. Clearly, the sum of the two covers the whole universe of all possible collections of objects, i.e. $X_0 + X_0' = \mathbf{1}$. On the other hand, their product is $X_0 X_0' = \mathbf{0}$. Right panel: then resulting $\mathcal{B}_1$ network. We have two pathways for going from 00 to 11. Note that we also have arrows going from each node back to itself (not shown in the diagram). Clearly, such 1D algebra is linear

where t is the number of 1s (truths) present in the $m = 2^n$ ID. By defining $x := m - t$, we have

$$r(n) = \sum_{x=1}^{m} x \binom{m}{x}, \tag{8.119}$$

Since we have the mathematical equality

$$\sum_{k=0}^{n} k \binom{n}{k} = n2^{n-1}, \tag{8.120}$$

we also have

$$r(n) = m2^{m-1} = 2^n 2^{m-1} = 2^{m+n-1}. \tag{8.121}$$

- The product of all relations gives the number

$$p(n) = m(n)! \tag{8.122}$$

of possible paths going from the lowest to the highest node.

Thus, the Boolean algebra is a category because it satisfies all basic requirements (8.24)–(8.26). Composition is ensured by the transitivity of implication. Moreover, it is a (directed) graph with its paths, it is a POSet, it is the monoid $(\mathbf{0}, \mathbf{1}, +, \times)$, it is also a group with the reversal operation (and $\mathbf{0}, \mathbf{1}$ as the identity elements for product and sum, respectively), and it is the category of all categories (**CAT**) since any category can be found in it (Sect. 8.1.3). Evidence for the latter point will be shown below.

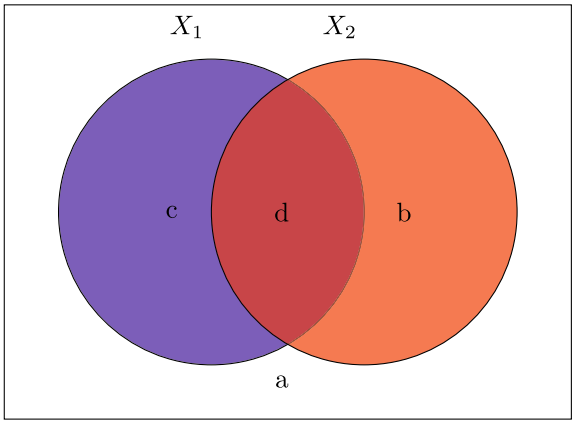


Fig. 8.14 The sets X_1 and X_2. X_1 is represented as the sum of areas c e d while X_2 as the union of areas b and d: see Table 8.8

Table 8.8 The IDs of the self-dual variables X_1 and X_2 for $\mathcal{B}_2$

	a	bc	d		a	bc	d
X_1	0	01	1	X_2	0	10	1
X_1'	1	10	0	X_2'	1	01	0

The First Algebras

Let us now present the first n-level algebras in a systematic way. For $\mathcal{B}_1$, we have just one variable. Let us introduce a convention[38] that is very helpful for the following and denote such variable X_0 and its complement X_0', so that we have four classes as a whole, as displayed in Fig. 8.13.

For $\mathcal{B}_2$ (with $m(2) = 4$, $k(2) = 16$, $l(2) = 5$, $r(2) = 32$, $p(2) = 4! = 24$), we get the situation represented in Fig. 8.14, which gives rise to the IDs for the basic sets X_1, X_2 displayed in Table 8.8. Note that we are relatively free in denoting the different areas displayed in Fig. 8.14. For instance, we could have used Greek letter instead of Latin ones. We could even reversed the order putting what here is d at the place of a. Nevertheless, the relations that are established between X_1 and X_2 are invariant across all these possible choices.

To obtain the results shown in Table 8.8, we use now a generative (iterative) procedure. These IDs are generated from those of $\mathcal{B}_1$ as follows: we split the original IDs of X_0 and X_0' in two parts and we put in between again the IDs of X_0 and X_0'. Note that either the first two numbers or the latter two of the two IDs of the variables and their complements cover all 4 classes of $\mathcal{B}_1$. All 16 classes can be summarised as in Table 8.9. Each of the IDs generated by combining (either through sum or product) the basic sets (self-dual variables) is computed by using the IDs of the sets themselves. The corresponding Boolean network is shown in Fig. 8.15.

[38] Presented for the first time in Auletta (2015c).

Table 8.9 The different classes generated in $\mathcal{B}_2$

Level	ID	Classes	Areas
4/4	1111	$X_1 + X_1' + X_2 + X_2'$	$a + b + c + d = (a'b'c'd')'$
3/4	1110	$X_1' + X_2'$	$a + b + c = d'$
	1101	$X_1' + X_2$	$a + b + d = c'$
	1011	$X_1 + X_2'$	$a + c + d = b'$
	0111	$X_1 + X_2$	$b + c + d = a'$
2/4	1100	X_1'	$a + b = c'd'$
	1010	X_2'	$a + c = b'd'$
	1001	$X_1 \sim X_2$	$a + d = b'c'$
	0110	$X_1 \nsim X_2$	$b + c = a'd'$
	0101	X_2	$b + d = a'c'$
	0011	X_1	$c + d = a'b'$
1/4	1000	$X_1'X_2'$	$a = b'c'd'$
	0100	$X_1'X_2$	$b = a'c'd'$
	0010	X_1X_2'	$c = a'b'd'$
	0001	X_1X_2	$d = a'b'c'$
0/4	0000	$X_1X_1'X_2X_2'$	$a'b'c'd' = (a + b + c + d)'$

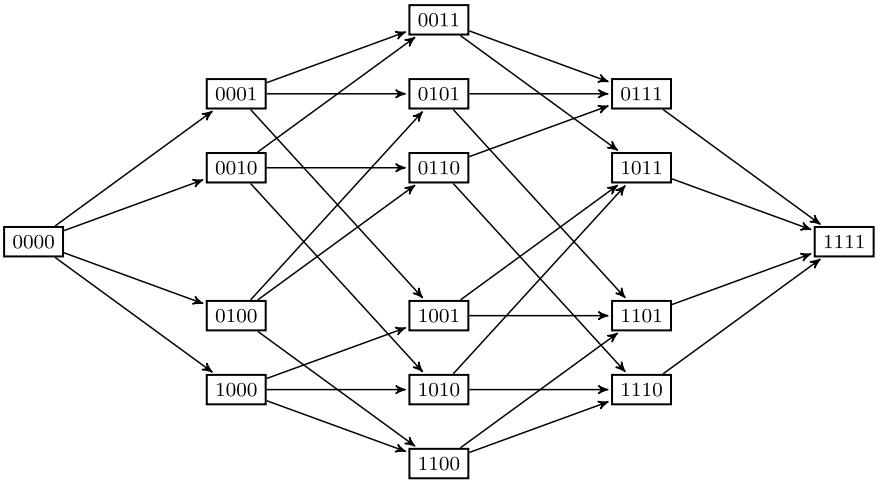


Fig. 8.15 The $\mathcal{B}_2$ network. We have 24 pathways bringing us from 0000 to 1111

Table 8.10 The generation of the basic sets (self-dual variables) of $\mathcal{B}_3$ and their complements. Note that either the columns represented by the first four numbers or those represented by the latter 4 ones cover all 16 classes of $\mathcal{B}_2$. Parenthetically, this explains why $\forall n$ the number of the basic sets and their complements of $\mathcal{B}_n$ have the size of $k(n-1)$

$X_{1.1}$	00	0011	11	$X_{2.1}$	01	0011	01
$X_{1.2}$	00	0101	11	$X_{2.2}$	01	0101	01
$X_{1.3}$	00	1010	11	$X_{2.3}$	01	1010	01
$X_{1.4}$	00	1100	11	$X_{2.4}$	01	1100	01
$X'_{1.1}$	11	1100	00	$X'_{2.1}$	10	1100	10
$X'_{1.2}$	11	1010	00	$X'_{2.2}$	10	1010	10
$X'_{1.3}$	11	0101	00	$X'_{2.3}$	10	0101	10
$X_{1.4}$	11	0011	00	$X'_{2.4}$	10	0011	10

Let us now consider $\mathcal{B}_3$. For $n = 3$, we have $m(3) = 8, k(3) = 256, l(3) = 9$, $r(3) = 1,024, p(3) = 40,320$. The previous examples can be considered introductory ones since, for dimensions $n \leq 2$, the Boolean algebra does not displays some of its basic features. The most distinguished feature for $\mathcal{B}_n$ with $n \geq 3$ is that the number of the basic set is larger than the dimension of the algebra. This is due to the fact that only for $\mathcal{B}_1$ and $\mathcal{B}_2$, the number of basic sets (self-dual variables) together with their complements is equal to m. For the case under consideration, we have 8 basic sets (plus their 8 complements). A choice of the basic sets is displayed in Fig. 8.16. Note that we generate the IDs of all those variables by reiterating the same procedure used for $\mathcal{B}_2$, as displayed in Table 8.10. Note again that we are relative free in choosing either this order or to reverse it and, e.g. denoting with $X_{1.1}$ what is here $X_{1.4}$. Nevertheless, the possible combinations are these ones and need to be put in such a sequence (or its inverse).

This means that we can have different choices of triplets of basic sets for giving rise to all the 256 classes. In theory, we could have

$$\binom{8}{3} = 56 \tag{8.123}$$

different choices. In practice, many combinations do not work and finally, as displayed in Table 8.11, we have two main codes of 3D self-dual variables (Code I and

Table 8.11 The two main codes of basic sets of $\mathcal{B}_3$

	abcd	efgh	Sets		abcd	efgh	Sets
Code I	0000	1111	$X_{1.1}$	Code II	0001	0111	$X_{1.2}$
	0011	0011	$X_{1.4}$		0010	1011	$X_{1.3}$
	0101	0101	$X_{2.2}$		0100	1101	$X_{2.1}$
	0110	1001	$X_{2.3}$		0111	0001	$X_{2.4}$

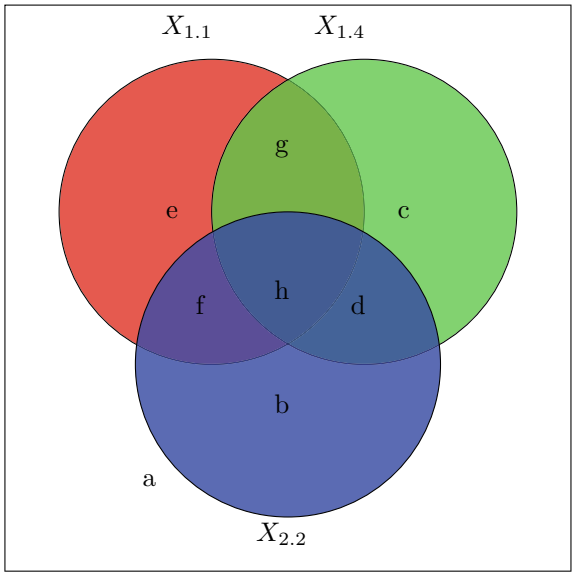


Fig. 8.16 Three basic sets of $\mathcal{B}_3$

Code II)

$$\{X_{1.1}, X_{1.4}, X_{2.2}, X_{2.3}\} \text{ and } \{X_{1.2}, X_{1.3}, X_{2.1}, X_{2.4}\}, \tag{8.124}$$

and four subcodes for each code:

$$\{\{X_{1.1}, X_{1.4}, X_{2.2}\}, \{X_{1.4}, X_{1.1}, X_{2.3}\}, \{X_{2.2}, X_{2.3}, X_{1.1}\}, \{X_{2.3}, X_{2.2}, X_{1.4}\}\},$$
$$\{\{X_{1.2}, X_{1.3}, X_{2.1}\}, \{X_{1.3}, X_{1.2}, X_{2.4}\}, \{X_{2.1}, X_{2.4}, X_{1.2}\}, \{X_{2.4}, X_{2.1}, X_{1.3}\}\}. \tag{8.125}$$

This clearly shows that a single subcode out of the previous one is sufficient to generate any class of $\mathcal{B}_3$. The order of the elements is here important since all relation among them are ruled by cyclic transpositions of elements. Although a single subcode is sufficient to generate all classes of the algebra, the use of a main code has also its advantages.

It is important to stress the general rule that governs the generation of Codes I and II and similar ones for higher dimensional algebras. Apart from the first and last columns (that have all 0s and 1s, respectively), all columns of the listed items need to have equal proportion of 1s and 0s. Moreover, apart from the second and the penultimate columns (which have all 1s under and above the horizontal central line, respectively), one-fourth of the 1s (and of the 0s) need to be above the central horizontal line and one-fourth below. Furthermore, we cannot have two or more replicates of the same column. Finally, we need to cross the two main lists of variables. It is easy to verify that the only allowed combinations are precisely Codes I and II (at most we can get some permutation of the same rows).

To have provided a complete analysis of $\mathcal{B}_3$ is crucial for understanding the fundamental characters of Boolean algebra. Thus, we do not need to reiterate such a complete analysis for dimensions >3. This would also largely exceed the scope not only of a chapter but even of a whole book! $\mathcal{B}_4$ is characterised by the following numbers: $m(4) = 16, l(4) = 17$,

$$\begin{aligned} k(4) &= 1 + 16 + 120 + 560 + 1820 + 4368 + 8008 + 11440 + 12870 \\ &\quad +11440 + 8008 + 4368 + 1820 + 560 + 120 + 16 + 1 \\ &= 65,536, \end{aligned} \tag{8.126}$$

and

$$\begin{aligned} r(4) &= 2(16 + 240 + 1680 + 7280 + 21840 + 48048 + 80080 + 102960) \\ &= 524,288 = 2^{19} = 2^{m(4)+3}. \end{aligned} \tag{8.127}$$

Finally, the number of paths is already exceedingly high:

$$p(4) = 16! = 20,922,789,888,000. \tag{8.128}$$

By iterating the same procedure employed previously, we build all the 128 basic sets, as displayed in Table 8.12.

8.2.3 Building Logical Spaces

Traditionally, Boolean algebra is understood as connecting logic and set theory. I have already shown that it, in fact, connects logic with classes (and therefore category theory). Moreover, it can also be connected with vector spaces (which is another category). This is extremely helpful when dealing with physical theories like QM and in particular with quantum information. In fact, also the connection between logic and information is a distinctive character of category theory.

Logical Spaces

The logical ID can be put in vectorial form as for instance a column or a row of numbers. Since we do not deal with complex numbers (that have no significance in logic), I use the row convention. This will give the *direction* of the logical vector (LV) representing the element of the algebra. By attributing also *length* to all LVs, we can build a logical vector space (LVS).[39] A nD LVS (with $n \in \mathbb{N}$), denoted $\mathcal{V}^n$, consists of LVs $x, y, z, \ldots$ such that, $\forall X \in \mathcal{B}_n, \exists x \in \mathcal{V}^n$ such that

[39] For the notion of vector space see Byron and Fuller (1969, I, Sect. 3.1).

Table 8.12 The 16 main codes for $\mathcal{B}_4$ (the left column is an expansion of Code I of $\mathcal{B}_3$ as well as the right column is an expansion of Code II of $\mathcal{B}_3$). Note that all ID columns in any set have four 1s and four 0s apart from the first (eight 0s) and the last (eight 1s). The 16 codes can be easily generated by focusing on the last eight values and first considering the last four values of each row. All the possible combinations for four truth values are 16: 1 for four 1s, 1 for four 0s, 4 for three 1s and one 1 and vice versa, 6 for two 1s and two 0s. Note that columns a and p have all 0s and 1s, respectively; columns b and o have the four 1s below and above the dashed line (in accordance with $\mathcal{B}_3$); columns c–d and m–n have two 1s above and below the dashed line, respectively (here the elements come in duplets)

Code	abcd	efgh	ijkl	mnop		Code	abcd	efgh	ijkl	mnop	
I.1	0000	0000	1111	1111	$X_{1.1.1}$	II.1	0001	0000	1111	0111	$X_{1.2.1}$
	0000	1111	0000	1111	$X_{1.1.16}$		0001	1111	0000	0111	$X_{1.2.16}$
	0011	0011	0011	0011	$X_{1.4.4}$		0010	0011	0011	1011	$X_{1.3.4}$
	0011	1100	1100	0011	$X_{1.4.13}$		0010	1100	1100	1011	$X_{1.3.13}$
	0101	0101	0101	0101	$X_{2.2.6}$		0100	0101	0101	1101	$X_{2.1.6}$
	0101	1010	1010	0101	$X_{2.2.11}$		0100	1010	1010	1101	$X_{2.1.11}$
	0110	0110	1001	1001	$X_{2.3.7}$		0111	0110	1001	0001	$X_{2.4.7}$
	0110	1001	0110	1001	$X_{2.3.10}$		0111	1001	0110	0001	$X_{2.4.10}$
I.2	0000	0001	0111	1111	$X_{1.1.2}$	II.2	0001	0001	0111	0111	$X_{1.2.2}$
	0000	1110	1000	1111	$X_{1.1.15}$		0001	1110	1000	0111	$X_{1.2.15}$
	0011	0010	1011	0011	$X_{1.4.3}$		0010	0010	1011	1011	$X_{1.3.3}$
	0011	1101	0100	0011	$X_{1.4.14}$		0010	1101	0100	1011	$X_{1.3.14}$
	0101	0100	1101	0101	$X_{2.2.5}$		0100	0100	1101	1101	$X_{2.1.5}$
	0101	1011	0010	0101	$X_{2.2.12}$		0100	1011	0010	1101	$X_{2.1.12}$
	0110	0111	0001	1001	$X_{2.3.8}$		0111	0111	0001	0001	$X_{2.4.8}$
	0110	1000	1110	1001	$X_{2.3.9}$		0111	1000	1110	0001	$X_{2.4.9}$
I.3	0000	0010	1011	1111	$X_{1.1.3}$	II.3	0001	0010	1011	0111	$X_{1.2.3}$
	0000	1101	0100	1111	$X_{1.1.14}$		0001	1101	0100	0111	$X_{1.2.14}$
	0011	0001	0111	0011	$X_{1.4.2}$		0010	0001	0111	1011	$X_{1.3.2}$
	0011	1110	1000	0011	$X_{1.4.15}$		0010	1110	1000	1011	$X_{1.3.15}$
	0101	0111	0001	0101	$X_{2.2.8}$		0100	0111	0001	1101	$X_{2.1.8}$
	0101	1000	1110	0101	$X_{2.2.9}$		0100	1000	1110	1101	$X_{2.1.9}$
	0110	0100	1101	1001	$X_{2.3.5}$		0111	0100	1101	0001	$X_{2.4.5}$
	0110	1011	0010	1001	$X_{2.3.12}$		0111	1011	0010	0001	$X_{2.4.12}$
I.4	0000	0011	0011	1111	$X_{1.1.4}$	II.4	0001	0011	0011	0111	$X_{1.2.4}$
	0000	1100	1100	1111	$X_{1.1.13}$		0001	1100	1100	0111	$X_{1.2.13}$
	0011	0000	1111	0011	$X_{1.4.1}$		0010	0000	1111	1011	$X_{1.3.1}$
	0011	1111	0000	0011	$X_{1.4.16}$		0010	1111	0000	1011	$X_{1.3.16}$
	0101	0110	1001	0101	$X_{2.2.7}$		0100	0110	1001	1101	$X_{2.1.7}$
	0101	1001	0110	0101	$X_{2.2.10}$		0100	1001	0110	1101	$X_{2.1.10}$
	0110	0101	0101	1001	$X_{2.3.6}$		0111	0101	0101	0001	$X_{2.4.6}$
	0110	1010	1010	1001	$X_{2.3.11}$		0111	1010	1010	0001	$X_{2.4.11}$

(continued)

Table 8.12 (continued)

Code	abcd	efgh	ijkl	mnop		Code	abcd	efgh	ijkl	mnop	
I.5	0000	0100	1101	1111	$X_{1.1.5}$	II.5	0001	0100	1101	0111	$X_{1.2.5}$
	0000	1011	0010	1111	$X_{1.1.12}$		0001	1011	0010	0111	$X_{1.2.12}$
	0011	0111	0001	0011	$X_{1.4.8}$		0010	0111	0001	1011	$X_{1.3.8}$
	0011	1000	1110	0011	$X_{1.4.9}$		0010	1000	1110	1011	$X_{1.3.9}$
	0101	0001	0111	0101	$X_{2.2.2}$		0100	0001	0111	1101	$X_{2.1.2}$
	0101	1110	1000	0101	$X_{2.2.15}$		0100	1110	1000	1101	$X_{2.1.15}$
	0110	0010	1011	1001	$X_{2.3.3}$		0111	0010	1011	0001	$X_{2.4.3}$
	0110	1101	0100	1001	$X_{2.3.14}$		0111	1101	0100	0001	$X_{2.4.14}$
I.6	0000	0101	0101	1111	$X_{1.1.6}$	II.6	0001	0101	0101	0111	$X_{1.2.6}$
	0000	1010	1010	1111	$X_{1.1.11}$		0001	1010	1010	0111	$X_{1.2.11}$
	0011	0110	1001	0011	$X_{1.4.7}$		0010	0110	1001	1011	$X_{1.3.7}$
	0011	1001	0110	0011	$X_{1.4.10}$		0010	1001	0110	1011	$X_{1.3.10}$
	0101	0000	1111	0101	$X_{2.2.1}$		0100	0000	1111	1101	$X_{2.1.1}$
	0101	1111	0000	0101	$X_{2.2.16}$		0100	1111	0000	1101	$X_{2.1.16}$
	0110	0011	0011	1001	$X_{2.3.4}$		0111	0011	0011	0001	$X_{2.4.4}$
	0110	1100	1100	1001	$X_{2.3.13}$		0111	1100	1100	0001	$X_{2.4.13}$
I.7	0000	0110	1001	1111	$X_{1.1.7}$	II.7	0001	0110	1001	0111	$X_{1.2.7}$
	0000	1001	0110	1111	$X_{1.1.10}$		0001	1001	0110	0111	$X_{1.2.10}$
	0011	0101	0101	0011	$X_{1.4.6}$		0010	0101	0101	1011	$X_{1.3.6}$
	0011	1010	1010	0011	$X_{1.4.11}$		0010	1010	1010	1011	$X_{1.3.11}$
	0101	0011	0011	0101	$X_{2.2.4}$		0100	0011	0011	1101	$X_{2.1.4}$
	0101	1100	1100	0101	$X_{2.2.13}$		0100	1100	1100	1101	$X_{2.1.13}$
	0110	0000	1111	1001	$X_{2.3.1}$		0111	0000	1111	0001	$X_{2.4.1}$
	0110	1111	0000	1001	$X_{2.3.16}$		0111	1111	0000	0001	$X_{2.4.16}$
I.8	0000	0111	0001	1111	$X_{1.1.8}$	II.8	0001	0111	0001	0111	$X_{1.2.8}$
	0000	1000	1110	1111	$X_{1.1.9}$		0001	1000	1110	0111	$X_{1.2.9}$
	0011	0100	1101	0011	$X_{1.4.5}$		0010	0100	1101	1011	$X_{1.3.5}$
	0011	1011	0010	0011	$X_{1.4.12}$		0010	1011	0010	1011	$X_{1.3.12}$
	0101	0010	1011	0101	$X_{2.2.3}$		0100	0010	1011	1101	$X_{2.1.3}$
	0101	1101	0100	0101	$X_{2.2.14}$		0100	1101	0100	1101	$X_{2.1.14}$
	0110	0001	0111	1001	$X_{2.3.2}$		0111	0001	0111	0001	$X_{2.4.2}$
	0110	1110	1000	1001	$X_{2.3.15}$		0111	1110	1000	0001	$X_{2.4.15}$

$$x = \frac{j}{2^n}(X), \tag{8.129}$$

where $j \in \mathbb{N}$ is determined by the sum of 1s present in the ID of X, and thus $j/2^n$ gives a logical level. In other words, the length of a LV depends on the ID of the corresponding logical expression. Thus, the mapping $X \mapsto x$ is injective. Note that the LV corresponding to **0** has length 0 while the LV corresponding to **1** has length 1. I

remark, however, that, although having a relation with the level of the corresponding algebra, the length of the vectors has a *geometric* significance (it determines the structure of the LVS). It should be further noted that the length of any LV is *invariant* across LVSs of different dimensions: e.g. all LVs representing self-dual variables have length of 1/2 in any nD LVS.

Coming to the geometry, the best way to represent a LVS is by means of concentric hyperspheres centred on $\mathbf{0}$ whose rays $\forall\mathcal{B}_n$, paralleling the labels for logical levels, are given by

$$\frac{0}{2^n}, \frac{1}{2^n}, \frac{2}{2^n}, \ldots, \frac{2^n}{2^n},$$

and with external hypersurface corresponding to $\mathbf{1}$, as displayed in Fig. 8.17. In other words, any LV, when prolonged to length 1, is in fact $\mathbf{1}$. For this reason, the latter can be called the universal LV (as well as the LV corresponding to $\mathbf{0}$ as the null LV). Another difference is that material equivalences and countervalences are represented by 2-directional LVs.

Moreover, the LVS is endowed with two operations: vector sum and scalar product. The *vector sum* is defined as follows: $\forall n \in \mathbb{N}$ and $\forall x_j, x_k \in \mathcal{V}^n$ it is given by the logical sum of the IDs of the corresponding $X_j, X_k \in \mathcal{B}_n$ times the number $j/2^n \in \mathbb{Q}$ given by Eq. (8.129). Thus, if (for $\mathcal{V}^3$)

$$x_{1.1} = \frac{4}{8}(X_{1.1}) = \frac{4}{8}(00001111) \text{ and } x_{1.4} = \frac{4}{8}(X_{1.4}) = \frac{4}{8}(00110011),$$

we have

$$x_{1.1} + x_{1.4} = \frac{6}{8}(X_{1.1} + X_{1.4}) = \frac{6}{8}(00111111). \tag{8.130}$$

We say that a set collects *linear independent* vectors when no vector of the set can be written as a linear combination (sum) of other vectors of the set.[40] Since any set of normal linearly independent vectors in the geometric (ordinary) vector space can be written as having a 1 in only one slot, linear independence can be expressed in this case as affirming that no such vectors can have a 1 in the same slot (they have orthogonal directions). However, in the logical space, the 0s do not denote "absence" (no component in that direction) but express truth values (as well as the 1s). Thus, the crucial notion for judging whether two or more LVs are linearly independent is how many truth values are shared (how many 0s and 1s are in the same slot). In particular, linear independent LVs need to share the minimal possible number of truth values. The minimal number is clearly one-half of the possible assignation. So, linearly independent LVs need to share exactly one-half of the truth values, no more no less.

For this reason, when dealing with geometrical properties of the LVS, it would be better to conceive the two numbers occurring here as -1 (for 0) and $+1$ (for 1) or simply as $-$, $+$. This allows us to define the *scalar product*: $\forall x, y$ and $\forall\mathcal{B}_n$ with

[40] Byron and Fuller (1969, I, Sect. 3.2).

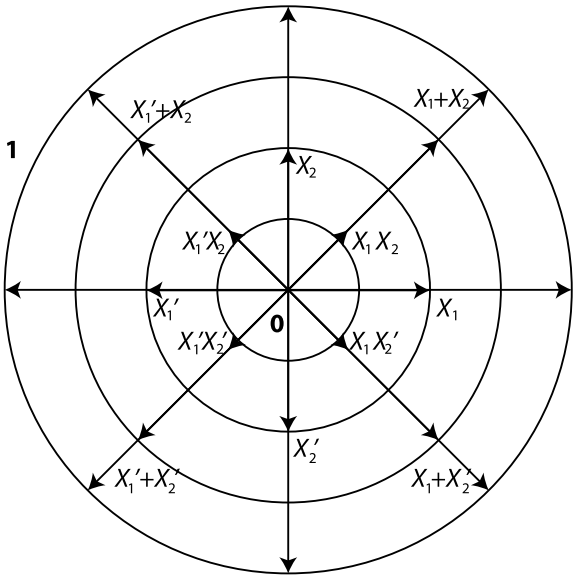


Fig. 8.17 The way in which we can represent the 2D LVS (denoted by $\mathcal{V}^2$). Note that we have here five circles (one of them represented by the **0** point) corresponding to five levels (determined by the number of 1s in the ID). Note the vectors of the kind X_1X_2 have length $1/4$, vectors of the kind X_1, X_2 have length $1/2$, and vectors of the kind $X_1 + X_2$ have length $3/4$ across any $\mathcal{B}_n$

$$x_j = \frac{j}{2^n}(X_j) \text{ and } x_k = \frac{k}{2^n}(X_k), \tag{8.131}$$

the scalar product is given by

$$(x_j, x_k) = l \left| \sum_i X_j^i Y_k^i \right|, \tag{8.132}$$

where the X_j^i, X_k^i are the ith truth value occurring in the logical part of the LVs x_j, x_k, respectively, and the number $l \in \mathbb{Q}$ given by Eq. (8.129) for the ID of XY. The products of form $X_j^i Y_k^i$ follow the rule of the logical product. It is easy to verify that the scalar product between linearly independent vectors is in fact zero: consider the LVs

$$x_{1.2} = \frac{1}{2}(----++++),\ x_{1.4} = \tfrac{1}{2}(--++--++),\ x_{2.2} = \frac{1}{2}(-+-+-+-+).$$

It is clear that the sum of all products column by column of any pair of these LV will yield 0. One can also verify that the product of any LV $x = j/2^n(X)$ by itself gives $j/2^n$. This is also true for LVs $x = j/2^n(X)$ and $x' = j/2^n(X')$ (where x and x' have opposite orientations), what shows that the scalar product and the logical product (see Eq. (8.111)) are not the same.

Bases

A subset S of $\mathcal{V}$ is said to *span* the latter if any element of $\mathcal{V}$ can be written as a linear combination of elements of S. A subset S of $\mathcal{V}$ is said to be a *basis* for $\mathcal{V}$ if S spans $\mathcal{V}$ and is a collection of linearly independent vectors.[41] Note that only LVs corresponding to self-dual variables (in short, self-dual LVs) can represent a logical basis (they are the only sets of LVs with n elements for any $\mathcal{B}_n$ to share the half of the truth values).

Having solved the foundational issues, for commodity, in the following, I often identify, e.g. the LV x with the ID of X, provided that no ambiguity arises. For LVS of dimension $n \geq 3$ ($n \in \mathbb{N}$), we have several alternative bases. For instance, for $\mathcal{B}_3$, the latter can be subdivided into subsets that are linearly independent (called spanning sets), and correspond to the two main codes displayed in Table 8.11. Any choice of three LVs in any subsets is a sufficient basis of the space. In such a case, we get the 8 bases represented in (8.125).

We can follow the same general procedure for $\mathcal{B}_4$. In this case, we obtain 56 subcodes for each of the main 16 codes. This, however, goes further than the scope of the present book.

Change of Basis

Vectors in one basis are kinds of vectorial sum of vectors in another basis. However, the rules here are a little bit more complicated than for an "ordinary" vector space. The best option is to make the problem "symmetric".[42] Indeed, with three dimensions, we have four combinations of bases (triplets) for each main code. Of course, we need to represent the change of basis as a rotation of the axes of a given basis. However, the involved vectors are here four, what means that we need a 4D LVS for performing such a rotation. Similarly, for four dimensions, we need in fact a 8D LVS. Considering Codes 1 and 2 in Table 8.11, we have the following transformations for Code 1 (considering here only the logical part of the change of basis):

$$
\begin{aligned}
X_{1.1} &= X_{1.4} \sim (X_{2.2} \sim X_{2.3}) = X_{2.2} \sim (X_{1.4} \sim X_{2.3}) = X_{2.3} \sim (X_{1.4} \sim X_{2.2}),\\
X_{1.4} &= X_{1.1} \sim (X_{2.2} \sim X_{2.3}) = X_{2.2} \sim (X_{1.1} \sim X_{2.3}) = X_{2.3} \sim (X_{1.1} \sim X_{2.2}),\\
X_{2.2} &= X_{1.1} \sim (X_{1.4} \sim X_{2.3}) = X_{1.4} \sim (X_{1.1} \sim X_{2.3}) = X_{2.3} \sim (X_{1.1} \sim X_{1.4}),\\
X_{2.3} &= X_{1.1} \sim (X_{1.4} \sim X_{2.2}) = X_{1.4} \sim (X_{1.1} \sim X_{2.2}) = X_{2.2} \sim (X_{1.1} \sim X_{1.4}),
\end{aligned}
$$

where, e.g. $X_{1.4} \sim (X_{2.2} \sim X_{2.3})$ is not $X_{1.4} \sim X_{2.2} \sim X_{2.3}$ but $X_{1.4}(X_{2.2} \sim X_{2.3}) + X'_{1.4}(X_{2.2} \sim X_{2.3})'$. The same for Code 2. Of course, since the main codes represent each a complete set of vectors for spanning such a space, we also have

[41] Byron and Fuller (1969, I, Sect. 3.3).

[42] Auletta (2015c).

$$
\begin{aligned}
X_{1.1} &= (X_{1.2} + X_{1.3})(X_{2.1} + X'_{2.4}) = (X_{1.2} + X_{2.1})(X_{1.3} + X'_{2.4}) = (X_{1.2} + X'_{2.4})(X_{1.3} + X_{2.1}),\\
X_{1.4} &= (X_{1.2} + X_{1.3})(X'_{2.1} + X_{2.4}) = (X_{1.2} + X'_{2.1})(X_{1.3} + X_{2.4}) = (X_{1.2} + X_{2.4})(X_{1.3} + X'_{2.1}),\\
X_{2.2} &= (X_{1.2} + X'_{1.3})(X_{2.1} + X_{2.4}) = (X_{1.2} + X_{2.1})(X'_{1.3} + X_{2.4}) = (X_{1.2} + X_{2.4})(X'_{1.3} + X_{2.1}),\\
X_{2.3} &= (X'_{1.2} + X_{1.3})(X_{2.1} + X_{2.4}) = (X'_{1.2} + X_{2.1})(X_{1.3} + X_{2.4}) = (X'_{1.2} + X_{2.4})(X_{1.3} + X_{2.1}),\\
X_{1.2} &= (X_{1.1} + X_{1.4})(X_{2.2} + X'_{2.3}) = (X_{1.1} + X_{2.2})(X_{1.4} + X'_{2.3}) = (X_{1.1} + X'_{2.3})(X_{1.4} + X_{2.2}),\\
X_{1.3} &= (X_{1.1} + X_{1.4})(X'_{2.2} + X_{2.3}) = (X_{1.1} + X'_{2.2})(X_{1.4} + X_{2.3}) = (X_{1.1} + X_{2.3})(X_{1.4} + X'_{2.2}),\\
X_{2.1} &= (X_{1.1} + X'_{1.4})(X_{2.2} + X_{2.3}) = (X_{1.1} + X_{2.2})(X'_{1.4} + X_{2.3}) = (X_{1.1} + X_{2.3})(X'_{1.4} + X_{2.2}),\\
X_{2.4} &= (X'_{1.1} + X_{1.4})(X_{2.2} + X_{2.3}) = (X'_{1.1} + X_{2.2})(X_{1.4} + X_{2.3}) = (X'_{1.1} + X_{2.3})(X_{1.4} + X_{2.2}).
\end{aligned}
$$

The former four rows refer to the expressions of the elements of Code 1 by means of those of Code 2, the latter four rows display the inverse transformations. As anticipated, this shows that all formulations of a self-dual variable in the terms of the other ones are ultimately a resolution of the identity. Moreover, the previous set of equations shows that what does matter is to establish which is the complement set (e.g. $X'_{2.4}$) giving rise to a certain LV through any combination with other LVs having the character of a sum times another sum. In other words, we have the general structure $(A + B)(C + D') = X$ with all pairwise permutations of the four terms on left giving the same result (same row). By displacing the prime from one variable to another, we get different sets (different rows). Note that such invariances are a logical character of self-dual variables although it comes out when we partition those variables in sets and subsets (8.124), (8.125). Moreover, the two sets of transformation rules are the same across any nD algebra (with $n \geq 3$). Let us call this important logical property the *permutation invariance* of self-dual variables of a code (or subcode) or set in those of another code (or subcode).

QM and Logical Spaces

Up to now, I have dealt with vectorial representations of the logical variables. In fact, such a logical space can also be made isomorphic to the Bloch sphere (Fig. 5.16, Sect. 5.2.4), at least for the 3D case, or, more generally, to the hypersphere of density matrices (Fig. 1.28, Sect. 1.4.1). In particular, we can map self-dual variables like $X, X', Y, \ldots$ to projectors $\hat{P}_x, \hat{P}_x^{\perp}, \hat{P}_y, \ldots$, i.e. into the *components* of a state $\hat{P}_\psi$ in that space, while statements of the form $X + Y$ to mixtures like $\hat{P}_x + \hat{P}_y$, where the weights have no logical significance (also the phase differences are logically irrelevant). In general, amplitude and phase cannot be represented in the logical space, and in fact depend on specific physical conditions and not on logical structures. Note that in the logical space projectors do not stand for individual events but rather for classes of events, which in general are equivalence classes of events, i.e. properties. Events in their specific happening require to specify particular physical conditions. Expressions like XY represent coincident (classes of) events ($\hat{P}_x \hat{P}_y$) while material equivalence ($X \sim Y$) and countervalence ($X \nsim Y$) to (parallel and antiparallel, respectively) entangled states. Obviously, we deal each time only with binary pro-

jections. Note that tautology **1** represents a pure state and is in fact a scalar covering the whole surface of the unitary sphere, in agreement with Eqs. (8.88) and (8.90). This corresponds to the fundamental property of all quantum systems to occupy as much "space" as possible and thus, when free, to cover all possibilities (Sects. 3.1.2 and 3.2.3). In such a way, we establish a biunivocal map among the notions of (i) tautology, (ii) to be in a zero-entropy state, and (iii) to display an infinite amount of possible information.

Infinity

When the number n of the dimensions of the Boolean space grows tending to infinity (and so the length of the ID), the number $2^{\frac{m}{2}}$ (with $m = 2^n$) of self-dual vectors (representing sets) relatively shrinks tending to a set of zero measure, according to

$$\lim_{n\to\infty} \frac{2^{2^{n-1}}}{2^{2^n}} = 0. \tag{8.133}$$

For instance, for a 3D LVS, the self-dual LVs are $1/16$ of all k collections of objects; for a 4D LVS, these LVs are $1/256$ of all k collections of objects; for a 5D LVS, the self-dual LVs are $1/65,536$ of all k collections of objects, and so on. On the other hand, in an infinite LVS, the different possible classes have at least the cardinality of real numbers, since they are 2^m. This means that, while the space is spanned by a number of vectors of the cardinality of natural (or rational) numbers (since they are by definition discrete), the different possible classes of objects establish a continuum. Quantum free systems reproduce this basic character.

Such an infinite Boolean algebra (denoted by $\mathcal{B}_\infty$) or LVS (denoted by $\mathcal{V}^\infty$) contains every possible class of objects or statement that could be uttered, and thus also any category, as previously anticipated. This means that whatever finite Boolean algebra (whether complete or not) can and need to be understood as an *extrapolation* out of $\mathcal{B}_\infty$ or as a subalgebra of $\mathcal{B}_\infty$. This system satisfies therefore mathematical *completeness* since any truth can be found. However, note that each set X of objects has possible relations not only with whatever other set of objects Y but also with its complement Y'. For instance, we have both XY and XY' and both $X + Y$ and $X + Y'$. It is not by chance that contradiction (**0**) is part of the algebra. So, $\mathcal{B}_\infty$ is complete but *not consistent*.

The infinite Boolean algebra presents only some very general rules that are instantiated by the relations themselves but no axioms (in the sense of Hilbert's). Here, no decision can be taken nor need to be taken. Decisions follow only when generating particular formal systems. Implementing appropriate decisions, the resulting formal system becomes consistent but incomplete. In fact, whatever formal system we like to build, for consistency we need to choose *certain* relations among classes and not others. These relations are determined by given assumptions. For instance, a class X has one of the two possible relations with another class Y, either with Y or its complement Y' (e.g. an integer is either even or odd). As a consequence, whatever

mathematical system we build, this is necessarily *incomplete*, since there are (many) truths that in fact are not included in the system. This is indeed the result of the Austrian mathematician and logician Kurt Gödel (1906–1978)[43]: there are many true statements (or relations among objects) that cannot be proved in any formal system for the simple reason that its rules are the selective decisions that exclude a large part of true statements. In other words, the incompleteness of formal system is a natural consequence of selecting out of a logical space that is complete in that strong sense. This shows that there is not a single way to do mathematics and that our rules based ultimately on arithmetic are only one of the infinite possibilities.

Selection Out of Infinity

This is also true and even more true when we consider physical universes. In fact, a Boolean algebra or space only represents a category: among its objects are several sets and classes. But nothing is said at this level (I mean at the level of the abstract and infinite combinatorics of all possibilities) about elements of those classes and sets, i.e. about individuals (apart from abstract mathematical sets, like the set of natural numbers). In fact, physical individuals are connected with the specific physical contexts and interactions of each world (showing discontinuous characters) (Sects. 7.1 and 7.2). In other words, those sets and classes taken *sub specie aeternitatis*, i.e. independently from each physical worlds, are like symbols that can receive a specific content (meaning) only when several constraints are specified. Therefore, at this level, these classes should be rather considered as abstract *schemes* of classes exhibiting all possible kinds (this means also incompatible) relations among sets.[44] In other words, Boolean algebra is a pure syntax lacking semantics. C. Peirce hypothesised "a potential aggregate of all the possibilities that are consistent with certain general conditions; and this may be such that given any collection of distinct individuals whatsoever, out of that potential aggregate there may be actualised a more multitudinous collection than the given collection".[45] In order to do that, we need to specify the classes with binary choices relative to all other classes. This is *selection* and therefore, according to quantum Darwinism (Sect. 5.1.3), selection is the hallmark of what is physical and even of the conditions of what is physical, considering the selections generating formal systems. Thus, no world can exist without choices (among relations) and this generates discontinuity (see Sects. 1.1.2 and 1.2.2–1.2.4).

Summarising, the fabric of reality is continuous.[46] but the world is discontinuous. At the limit, we can deal with many situations in our universe by assuming continuity, like for some observables of free systems. All quantum systems taken as free reproduce a combinatorics of all possibilities that must me mapped to the infinite combinatorics of $\mathcal{B}_\infty$. However, in our universe, no system is completely free. Thus,

[43] Gödel (1931).

[44] Carnap tells us that a class does not consist only of its members (Carnap 1928, Sect. 37).

[45] Peirce (1898, p. 247).

[46] See e.g. Peirce (1898, pp. 162–163).

continuity is a mathematical notion and only ideal from a physical local point of view, or at least, given the granularity of space-time at the Planck scale, is based on some coarse-graining.[47] It can even be asked if the whole universe is a completely free system or whether this notion applies only to the multiversal wave function.

Once that we impose specific restrictions on the kind of relations among sets and classes, we generate formal structures that potentially represent specific classes of basic objects and related symmetries. For instance, if there are classes of objects for which only counterequivalence but not equivalence is allowed, this can represent antisymmetric particles like fermions. Once that a self-consistent system of such constraints is built (and there are likely infinite many of such systems embedded in $\mathcal{B}_\infty$), we have generated the potential scaffold of a world. However, the relations among classes are here only of taxonomic kind (hierarchies of classes and subclasses: see Sect. 5.1.2) and not of causal kind, which require specific physical conditions. Although such a scaffold does not represent a world as such, it will constraint the 'space' in which several particular physical systems and even events become possible. Thus, self-consistent selections out of $\mathcal{B}_\infty$ spontaneously generate several (likely infinite) spaces of possibilities that can be indeed possible worlds. Since such constraints determine several symmetries, the laws of nature can emerge as consequences when some specific relations among classes of objects are selected (Sects. 3.2.4 and 6.2.5). Finally, in order to have causal processes and events, some of these symmetries need to be broken. Of course, the logical relations (implications) among classes, i.e. the arrows of the Boolean algebra, are necessary truths that need to be true in every world and in every formal system. This establishes an interesting connection with what is called modal logic.[48] However, I insist that such alternative possibilities concern classes and their relations and not individuals.

Thus, as anticipated in Sects. 3.1.2 and 5.1.5, the fabric of reality or the block universe are in fact simply the Boolean combinatorics of all possibilities. Possible worlds are like dots out of this infinite sea. It is also crucial to understand that, whatever relations will be in fact established in one of these worlds (even when one of them becomes real) must be framed in one (or perhaps many) of the relational schemes that represent the scaffold of possibilities. In other words, according to our analysis of Boolean algebra, relations are *logically predetermined* and cannot be invented in any of the worlds. However, reciprocally, relations and correlations need to be instantiated in physical systems of whatever physical universe in order to give rise to whatever particular contexts. They need to be *embodied*. Such an embodiment explains why types are primary from a logical point of view but secondary from a physical point of view (Sect. 6.3.4). The Australian philosopher D.M. Armstrong (1926–2014) calls this the principle of instantiation and assumes that both universals (types) and laws of nature (as relations among universals: see end of Sect. 5.2.3) need to be always instantiated.[49] This is clearly true for laws of Nature (and symmetries), as far as by selecting them we also select a particular world (and for this reason we

[47] Wheeler (1983), Wheeler (1990). See also Auletta (2000, Sect. 33.1.2).

[48] Chellas (1980).

[49] Armstrong (1983, p. 82 and ff).

are also inclined to think that they are not only possible but real, according to the MWI). However, this is not true for schemes of classes of the Boolean algebra, as far as they represent kind of 'second-level' universals. The principle of instantiation applied here would be the truism that in order to be concrete (or 1st order), they need to be instantiated.

A Game

J. Wheeler introduced the notion of the universe as a *self-synthesising system*,[50] although he often related this concept to the entrenchment between quantum events and their observation. There is an interesting version of the popular the 20 questions game of which Wheeler was protagonist. Let us quote here the words of Wheeler himself[51]:

> That reminded me of Edward Teller visiting in North Carolina at the time was there for a three-year stay. I remember an evening party where the game was played in which a person is sent out of the room and those behind agree on a word, and then the person comes back and has 20 questions to find out—yes or no. But I noticed that when I was sent out and came back and started the usual question: "Is it something in the animal kingdom?" As each successive question was asked, the answers came slower and slower and caused more and more trouble to those who were answering. Finally came to the final word and it was "cloud", after I'd asked a number of questions. Then at last they broke down and told me why the game looked so strange: because they had agreed in advance that this would be one where no word was agreed upon to start with. Every answer, however, would have to be consistent with all the answers that had gone before. So it was really harder for the people playing the game than it was for me. The point was the word "cloud" that had been produced really came more out of the questions that were asked than out of anything that they had agreed upon before the thing started.

The subtle idea of Wheeler is that the order that we observe in Nature can emerge by a complex process in which information takes a central role (Sect. 7.3.3). In other words, in such a perspective, laws (and order) emerge from a physical universe whose starting point is always chance. For this reason, Peirce conjectured that the laws of Nature could be the result of habits acquired by systems,[52] and told: "all things have a tendency to take habits".[53] However, this line of thought cannot be pursued too far: in fact, any level of order in our universe could not have emerged at all if not presupposing the order of the fabric of reality that is the ground of everything.[54]

---

[50] Wheeler (1988). See also Laughlin (2005).

[51] http://www.aip.org/history/ohilist/4958.html.

[52] Peirce (1884, pp. 553–554), Peirce (1891, p. 106). See also Auletta (2011a, Chap. 3). Peirce's view has been also supported in Smolin (2013), although the main thesis there is quite different from that supported here as far as the author rejects conservation laws and symmetries, which are, at the opposite, central to my approach. I remark that in the Introduction, Smolin quotes interesting statements of Dirac that also support the evolution of laws.

[53] Peirce (1887, p. 208).

[54] The late Peirce acknowledged this point to a certain extent (Peirce 1898, pp. 210–211). On such a problem see Auletta (2016d).

Wheeler's view represents the local perspective, out of which indeed structures and relations are built, as it is evident for cosmic evolution. However, from a global point of view, these structures and relations or symmetries (and even more the laws of Nature) need to conform with the eternal constraints of logic. Indeed, although our universe could have emerged from vacuum, the laws of QM are somehow already presupposed.

8.2.4 Is Quantum Logic Non-classical?

Is QM Not Boolean?

It is quite common to assume that the algebra or the logic underlying QM is not Boolean because it is believed that distribution laws (8.110), i.e.[55]

$$X(Y+Z) = XY + XZ \text{ and } X + YZ = (X+Y)(X+Z) \tag{8.134}$$

are not satisfied.[56] The whole trouble here is represented by the fact that, often, it is understood that a superposition state must be represented by a disjunction of the kind $X + Y$. According to Table 3.3, Sect. 3.3.1, disjunction comprehends also the cases in which *only* X or *only* Y (that could stand for paths in an interferometer) is true, while superposition truly means that *both 'ways' are occupied* (if both are open the quantum system will certainly go both ways), although in a weaker sense than that represented by a conjunction XY that expresses joint events. Now, once we interpret a superposition state as a disjunction like $Y + Z$, it is quite natural to consider experimental arrangements for which $X(Y+Z) \neq XY + XZ$. This interpretation is believed to stem from a famous contribution of the American mathematician George Birkhoff (1884–1944) and J. von Neumann.[57] However, the original paper of these authors is much more careful about this:

> It is worth remarking that in CM, one can easily define the meet or join of any two experimental propositions as an experimental proposition — simply by having independent observers read off the measurements which either proposition involves, and combining the results logically. This is true in QM only exceptionally — only when all the measurements involved commute (are compatible); in general, one can only express the join or meet of two given experimental propositions as a class of logically equivalent experimental propositions - i.e., as a physical quality. These facts suggest that the distributive law may break down in quantum physics. That it does break down is shown by the fact that if a denotes the experimental observation of a wave packet ψ on one side of a plane in ordinary space, a' correspondingly the observation of ψ on the other side, and b the observation of ψ in a state symmetric about the plane, then
>
> $$b(a+a') \neq ba + ba'. \tag{8.135}$$

[55] It seems that what follows is not far away from the spirit of Epperson and Zafiris (2013). Anyway, they are among the few scholars to take Boolean algebra as fundamental for QM.

[56] See Beltrametti and Cassinelli (1981, Chap. 12), In fact, one finds this statement still today (Coecke and Paquette 2011, pp. 246–247).

[57] Birkhoff and Von Neumann (1936). See also Landsman (2017, Sect. 2.10).

The statement about the logical conditions of experimental propositions is interesting in the light of what discussed in Sect. 3.3.4 about von Neumann's proof of the impossibility of HVs and Bell's contributions. Moreover, apart from some terminological issues, it is evident that the two authors are speaking about *possible measurement outcomes* and not about the nature of superposition (and therefore neither about the structure of Hilbert spaces). Now, the reason for the above result is due to the incompatibility of experimental arrangements in measuring non-commuting observables. However, as we know, this incompatibility is also true for CM and only concerns the way in which we sample the possible events in classes. Obviously, the possible events when measuring the path are not the same as those when measuring the interference visibility (Sect. 5.2.4). But these two different sets are otherwise fully classical,[58] so that the idea that distributivity fail here is not well grounded.

The fact is that expression (8.135) hides a crucial point. The two possibilities, a (going on one side of a plane) and a', are in fact mutually exclusive. Let us consider a concrete interferometry experiment. Now, facing a binary choice, if a represents the statement that "$|0\rangle$ is observed", it is clear that the statement a' that "$|1\rangle$ is observed" is in fact the negation of a (Sect. 5.2.2).[59] This means that a superposition of two components needs to be represented by a *tautology* [diagrams (8.89) and (8.91)]. In other words, according to Eq. (4.12), beam splitting can be considered as a kind of gate for producing tautologies: although we may be uncertain about the specific nature of the transformation (i.e. about the involved coefficients), in any case it will produce a result that must be represented by a tautology (apart from the limiting cases of pure reflection and pure transmission, which we exclude by assumption, dealing here with a true beam splitter although with variable parameters: see, e.g. the POVM case in Sect. 5.2.4). This is a crucial difference with classical gates. Thus, we can represent the superposition generated by the first beam splitter by the statement $s = a + a'$. Suppose also that the phase is tuned in such a way that only one detector clicks, and we represent this by the statement d. Then, the statement that the photon is in a superposition state and only one particular detector of the two will click can be expressed as

$$sd = (a + a')d, \tag{8.136}$$

Now, there is no fail of distributivity since

$$(a + a')d = \mathbf{1}d = d. \tag{8.137}$$

A similar calculus can be done for Eq. (8.135), so that we have $b(a + a') = b\mathbf{1} = b$. Similarly, for an interferometry of the kind shown in Fig. 5.18, Sect. 5.2.4, which can be represented by

$$a + (b + b') = a + \mathbf{1} = \mathbf{1}, \tag{8.138}$$

[58] As well understood by Ludwig

[59] See also Busch et al. (1995, pp. 25–26).

where a represents the upper path and $b + b' = a'$ the lower path which bifurcates. And analogically when there are more branchings. Of course, we can also have the case of multiple choices. For instance, a photon is trapped in a microcavity and has n possible way out. Let us consider the case $n = 3$ for simplicity. If there is equiprobability, the state can be described by $|\psi\rangle = 1/\sqrt{3}(|x\rangle + |y\rangle + |z\rangle)$. However, logically speaking, the state can, e.g. be rewritten $x + x'$, with $x' = y + z$, what leads us back to the previous cases.

A possible objection could be that, since the state is a superposition, then we are not authorised to treat the two paths as mutually *exclusive*. However, it seems to me that such a reasoning should consequently also deny that we can represent such a superposition as a combination of *different* options, in which case the whole argument fails. As a matter of fact, the two possible outcomes are mutually exclusive. A superposition of mutually exclusive potential events or components is not only reasonable but it comes out in thousands of situations that we have partly discussed in the present book. To assume the opposite brings us to an holism that is not the distinctive mark of QM, at least for multipartite systems (Sect. 6.4.4). Another possible objection could be that states of free systems can be perhaps represented by tautologies, but since the photon has been prepared through beam splitting in a particular state that is represented in a vectorial form, it cannot be represented by a tautology. However, a short reflection will show us that the statement sd expresses a product state: in an ordinary interferometry experiment, the detector knows nothing about the state of the incoming photon before detection, and therefore it cannot 'represent' it as a particular state vector (we cannot have a device being able to do this, otherwise detection would already have occurred at the stage of preparation or premeasurement). Thus, the only possibility is to represent it as a tautology. This shows that quantum-mechanical systems never lose their general character of tautologies. As I have stressed in Sect. 5.2.2, to throw away systems means to consider them as black boxes.

Probabilities and Sample Spaces

The issue of commutativity emerges only when we ask which events can jointly occur. However, as said, also experimental incompatibilities hold for the classical case. The only true difference, due to non-commutativity of, e.g. path and interference observables, is represented by Eq. (2.40). Nevertheless, in the quantum-mechanical case, unsharp observables help use to put this stuff on a correct formal basis that is essentially the same as for classical physics.

First, let us recall that probability, satisfying property (2.24), requires that the collection of all events to which we can assign a probability

- Is closed under countable unions: if $X_1, X_2, \ldots$ are in such collection, also $\bigcup_{j=1}^{+\infty} X_j$ it is;

- As mentioned, is closed under complements (i.e. negation): if X is in this collection, also X' it is. Thus, by De Morgan's law (3.130), we have also closure under countable intersections;
- The null event $\varnothing$ is part of such a collection.

The repartition in sample spaces is crucial. Without such a repartition, we cannot compute probabilities. Clearly, also for a prepared quantum system, we can compute probabilities. However, if no specific (experimental) context is selected, the state of the system can be expanded in different bases and such bases allow different (and incompatible) computations of probabilities. Thus, premeasurement is a fundamental step for sampling the event space in appropriate terms and then to compute specific probabilities. Moreover, the distribution of events in the sample spaces is what allows property ascription, i.e. inferences about the behaviour of the object system. We recover here, on a formal plane, the distinctions introduced in Sect. 6.2.5.

Two Examples

If we consider the sample space Ω of all possible events that could happen when interacting with a quantum system, they can obviously be cast in subsets responding to different experimental arrangements and of some of these sets we can build unions when unsharp observables are involved. Let us examine two examples. First, I shall consider Mandel's experiment schematically depicted in Fig. 5.13, Sect. 5.2.3. We can distinguish among four possible cases (where I recall that the relevant issue is whether we are able to discriminate the two signal photons or not):

- The NDF has zero transmissivity and the BM is arranged in such a way that there is destructive interference: no photon is detected. This is a null event. Since (for simplicity) this is assumed to be deterministic (probability equal to 1), we do not consider this case.
- When the transmissivity of the NDF is 100%, we have a coincidence detection of both photons and we cannot discriminate between s_1 and s_2, nor between i_1 and i_2. In such a case, we have a coincidence event E_{12} which is a coarse-graining of two simultaneous events (detector clicks) E_1 and E_2 (for detectors $\mathrm{D}_s, \mathrm{D}_i$, respectively). This is strong evidence of what anticipated in Sect. 6.2.5: the click of detector as such tells us nothing about nothing if not in appropriate experimental contexts. Note that single photon detections are discarded, and such cases represent the null event E_0.
- When the transmissivity of the NDF is zero (and thus photon i_1 is blocked) but the BM is properly arranged, we can distinguish between two subcases. The first one, is that we have a coincidence, in which case from the geometry of the experiment we infer that we have detected photons s_2 and i_2 (they have in fact symmetric paths). Then, we have again a coarse-graining event E_{34} composed of two events: E_3 when detector D_s clicks (and we infer that it was the photon s_2 to impinge on it), and E_4, when detector D_i clicks (and we infer that photons i_2 was trapped). Here, the null event is represented by no detection.

- The second subcase is when there is no coincidence. In that case, we know from the different length of the paths that photons s_1, i_2 must be involved so that there have been two independent events: E_5, when photon s_1 is detected by D_s, and E_6, when photon i_2 is detected by D_i. Again, the null event is represented by no detection.

Summarising, we have two different sample spaces $\Omega_1 = \{E_0, E_{12}\}$ and $\Omega_2 = \{E_0, E_{34}, E_{56}\}$, with $E_0 = E'_{12}$ for the first space and $E_0 = E'_{34}E'_{56}$, $E_{34} = E'_{56}$ for the second.

The second example is shown in Fig. 5.18, Sect. 5.2.4. Also here, we distinguish among different cases:

- In the case in which we measure the path of the photon ($\eta = 0$), we have possible events $\{E_1, E_2, E_3\}$ whose probabilities are (5.126). Here, the null event is represented by no detection.
- In the case in which we measure the complementary observable, i.e. the interference visibility ($\eta = 1$), we have events $\{E_4, E_5\}$ with occurrence probabilities (5.118). In this case, we have also the null event E_0 when there is destructive interference.
- Finally, when dealing with a POVM ($0 < \eta < 1$), we have kind of 'intermediate' events relative to the first two subsets, namely events $\{E_6, E_7, E_8\}$ given by probabilities (5.131), which could be understood as kind of 'intermediate' probabilities between the other two sets. The null event is represented by detection failure.

In such a case, we have three sample spaces $\Omega_1 = \{E_0, E_1, E_2, E_3\}$, $\Omega_2 = \{E_0, E_4, E_5\}$, and $\Omega_3 = \{E_0, E_6, E_7, E_8\}$, with $E_0 = E'_1E'_2E'_3$, $E_3 = E'_1E'_2$, $E_1 = E'_2$ for the first space, $E_0 = E'_4E'_5$, $E_4 = E'_5$ for the second space, and $E_0 = E'_6E'_7E'_8$, $E_8 = E'_6E'_7$, $E_6 = E'_7$.

8.3 Epistemological–Gnoseological Assessment

Since information does not by itself determine the conditions of its reception, neither of its interpretation (Sect. 6.2.5), it is time to deal with the issue of the nature of inference and knowledge. On this important subject, I like to quote the words of Eddington[60]: "It is actually an aid in the search for knowledge to understand the nature of the knowledge which we seek". Physics is the robust theoretical construct on which every natural science is based. However, such a construct needs logical and epistemological foundations and, as we shall see, the latter, in turn, need to be connected with basic biological functions.

This assessment is even more important for QM given the enormous problems that we have met when dealing with the relation theory–reality or observation–objective

[60]Eddington (1939, p. 5) He further noted that physics is more and more based on epistemological principles as well as mathematics on logical ones. He seems however to forget this when he repeatedly affirms (for instance Eddington 1939, p. 89) that in relativity we observe relations and in QM we observe probabilities. See also Eddington (1939, pp. 87–88).

facts. I shall show that although our theories are the result of human construction, they are not deprived of objective value either, so we cannot say that they are pure inventions of the mind. Following Schrödinger, I have dared to affirm that there must be a connection between physics and mind (Sect. 5.1.5). On the basis of what anticipated, I would rephrase this by saying that there must be some match between epistemology and ontology. Since we are nevertheless the constructors of our theories, as recalled by Einstein (and Bohr), the question becomes: what is the specific connection between ourselves and the world? This connection seems to be twofold. These issues will be explored in the next two subsections.

8.3.1 Physical (and Operational) Connection

On the one hand, we have a *physical connection*. This is necessary since without a dynamical intercourse with the world, no new information could be acquired (and none would have been acquired at the start) (Sects. 6.2, 6.3, and 7.2.1). As I have said following Heisenberg's point of view, *experimental operations* establish this dynamical connection (Sect. 5.2.1). Experimental operations are in no way different from other physical interactions among open systems (operations in a wider sense: Sect. 5.2.5) if not for the fact that they are *controlled*. To throw more light on this issue, we need to consider a specific class of physical systems: the biological ones.[61] The reason is that organisms thank their existence to control and we humans, as controllers, make use of the same basic mechanism.

Few Comments on the Emergence of Biological Processes

In view of what will be treated in the following, it can be helpful to have a short look at the second form of major emergence: the emergence of life out of the physical world (Sect. 7.3.3). Independently from the (big) question of how the first biological macromolecules like RNA may have been formed,[62] once that they appeared on Earth they underwent a terrible (Darwinian) selective fork: either disappear or become stable by protecting themselves through controlling the exchanges with the environment. In fact, a molecule of this kind represents a sophisticated structure that would spontaneously break down due to the second law of thermodynamics if it were not able to concur to the birth of a complex system (the cell) able to control such exchanges. In other words, the system must be able to perform *work* in order to keep internal order—demanding *energy efficiency* (low entropy). However, the maintenance of energy efficiency cannot be itself delivered to chance and thus in turn demands *predictive power* in order to deal appropriately with the environment, and this is ultimately based on the mutual information (again a form of correla-

[61] Auletta (2011a, Chap. 8).

[62] On this point see Auletta (2016a).

tion) between external signals and internal states of the biological system.[63] In other words, the metabolic selective canalisation demands an informational selective canalisation.[64] In this way, we can have a thermodynamic transition from inanimate to animate world: energy efficiency bridges between the physical and the biological worlds allowing the existence of systems that make of the quality of energy their crucial factor. Obviously, basic physical and chemical processes are necessary for the survival of the organism and go on as in the prebiotic world. The novelty is that they are now inserted in a new kind of context and in a (more complex) network equipped with new kinds of operations.

In particular, what is conserved is *metabolic efficiency*, understood as the difference between the work (and the amount of energy) needed for the structural, functional and behavioural needs of the organism and the work (and the amount of energy) needed for allowing the metabolic system to work further; in other words, the system must store energy for allowing the building or rebuilding of its own structures, for letting these work properly, and for acting in a proper way on the environment. Such a process allows the preservation of the homeostasis of the organism that can be taken as the stable state here. Therefore, we can consider biological systems as sort of 'molecular' Maxwell's demons: they build order and release disorder. In such a way, they locally 'violate' the principle of energy conservation since, in order to maintain energy efficiency, they are forced to exchange energy with the environment, especially assimilating free energy and giving back high-entropy energy. Obviously, in a system sufficiently large, there is no violation of the conservation of energy (and neither of entropy). However, the reader should be warned again to keep a sharp distinction between the issue of the conservation laws that are true for whole systems and the *typical behaviour* of certain systems at a certain level allowing new kinds of (emergent) operations. This can be ascertained by considering that also large biological networks and even huge ecosystems reproduce precisely the basic behaviour of organisms.

Control

We have seen that also quantum systems can spontaneously display operations that when well combined together can result in information acquisition (Sect. 6.3.4). As mentioned, the information that they bring with themselves has no reference or meaning and therefore those processes only happen to occur (Sects. 3.3.7 and 6.2.5). At the opposite, biological organisms, like we are, *systematically* acquire and gather information from the environment (included other species) that is crucial to their survival, and thus they need to be referred to specific classes of signals. As said, they use this gathered information for *controlling* the exchanges with the external world in order to have appropriate metabolic reactions: the latter is necessary for keeping their internal order and thus for their survival but cannot work without the former.

[63] Still et al. (2012) Sengupta et al. (2013).

[64] Auletta (2013a).

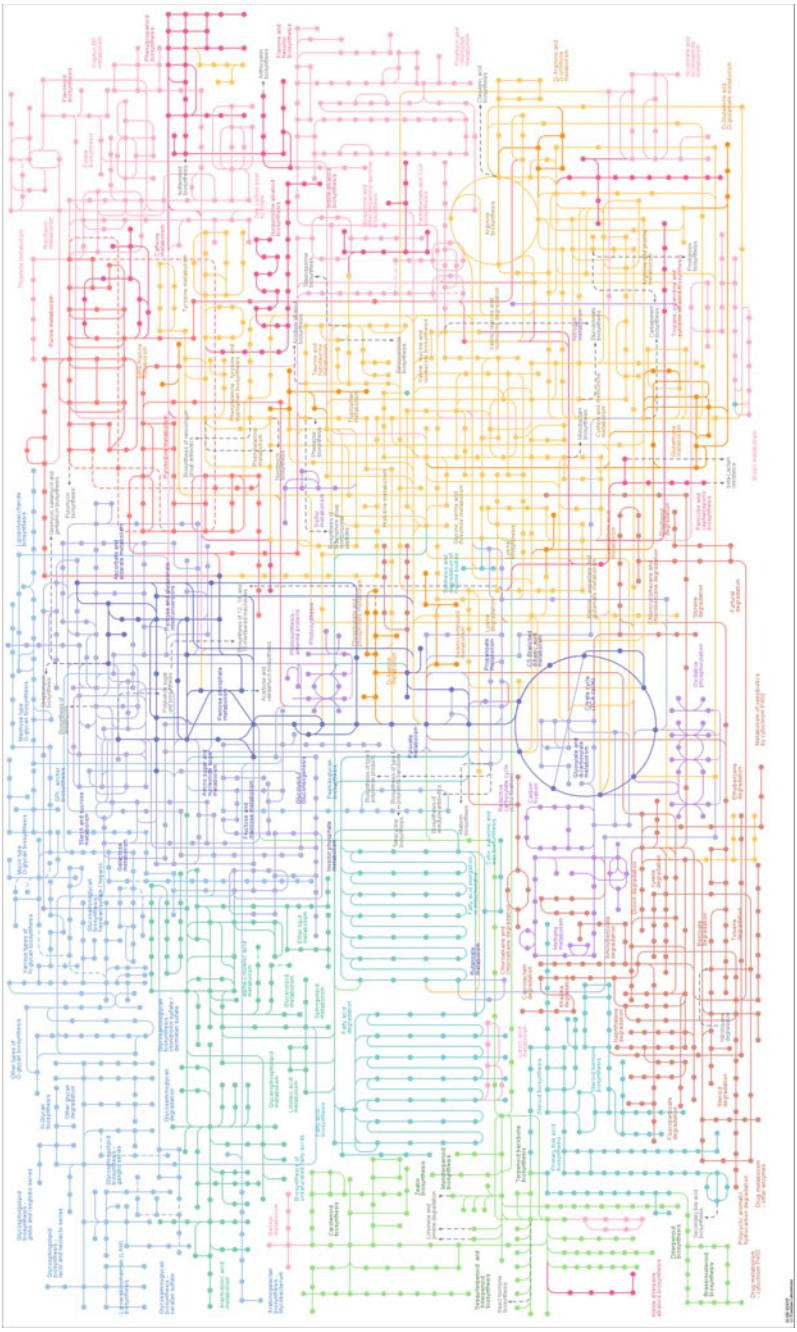


Fig. 8.18 Universal chart of metabolism. Adapted from http://www.genome.jp/kegg/pathway/map/map01100.html

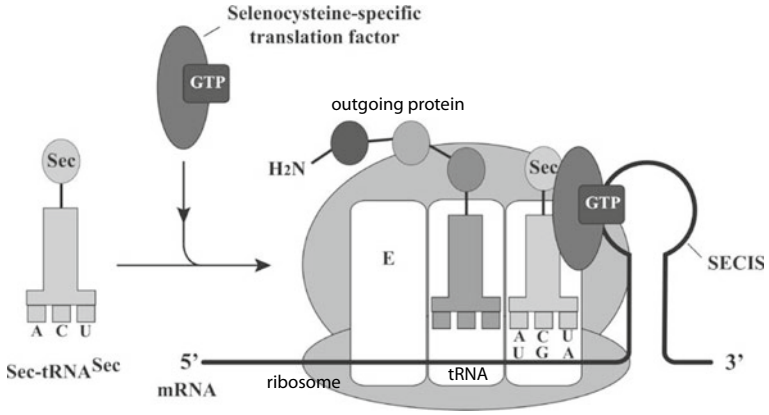


Fig. 8.19 The complex machinery for recoding the stop codon UGA into a code for selenocysteine. The fabric of all proteins is the ribosome (the big grey central body). It happens through tRNA matching with a triplet of the mRNA and having attached an amino acid at the other extreme (the part that becomes a unit of the outgoing protein). Here, recoding is performed by a modified tRNA (having the triplet ACU but with a selenocysteine (Sec) attached) thanks to the influence of (i) a stem loop structure (SECIS element, including a GTPase enzyme) downstream the UGA codon and a (ii) specific translation factor. Adapted from Ramon Fuentes (2016, p. 251)

Thus, life could not exist without control. For instance, living organisms are characterised by a significant amount of chemical reactions (Fig. 8.18). When considered from the point of view of the complexity of the involved chemicals, however, they represent a tiny amount of all the reactions that could spontaneously occur on that basis; for example, the number of proteins that can theoretically exist is likely to be about 10^{390}, yet the number of those that are produced by living organisms is a tiny part of that: about $10^6 - 10^8$ in the case of humans.[65] This means that chemical reactions are *driven* through sophisticated check and repair mechanisms.[66] A splendid example of such functional control is given by the production of selenocysteine, a crucial component for the selenoprotein, which is a biomolecule necessary for life. Since selenocysteine is an amino acid not codified by the genetic code, in order to produce this protein without introducing any change in the code, the cell has developed a very complex molecular machinery that allows the 'reading' of one of the stop codons (UGA) as a triplet codifying for the selenocysteine (Fig. 8.19). The reason is that codes (Sect. 3.2.3) are upstream of all regulations in the organism and so it is extremely dangerous to change them whilst it is safer to intervene downstream.

[65] Uhlen and Ponten (2005).

[66] Alberts et al. (1983, Chaps. 2, 7).

Sensory Activity

For understanding the role of *codes*[67] (Sect. 3.2.3) in control processes, let us consider the sensory activity[68]: as mentioned, all organisms are able to catch specific classes of signals (according to their environment and needs) in order to monitor environmental variations and act accordingly. A cybernetic theorem tells us that there is no controlling instance without a *model* of the system to be controlled.[69] The model represents the crucial elements of the controlled system and their functionality: in other words, it instantiates the default or expected functionality of the controlled system with which the current state of the latter needs to be compared. This sets from the start cognition in strict connection with action. In other words, the 'representational' models serve the purpose of practical operations on the (internal or external) environment. In particular, it allows error correction. Now, no model and controlling instance is possible without codes, in particular, without imposing *endogenous* codes on the system (e.g. the environment) to be controlled. Being such codification an endogenous activity, it fulfils the quantum-mechanical principle of local codification (Sect. 3.3.7). The necessity for organisms to use codes is simple: if the reactions or actions of the controlling instance depended on the physical–chemical nature of the signal (whether external or internal) received, the former would depend on the latter and become in fact controlled, as happens for most devices of our technology which are controlled, directly or indirectly, by ourselves.[70] At the opposite, imposing a code on the controlled instance allows the biological controller to attribute functional and vital meaning to the different signals according to its *intrinsic needs* (instantiated in the model). This could be considered to be a consequence of a quantum-mechanical no go theorem: no source determinism (Sect. 6.2.5). Note that such a process is necessarily *top-down*, where with such a term in such a context I understand the capability of the controller to canalise biochemical reactions occurring spontaneously.[71] In other words, no top instance can generate those reactions since this represented a violation of the principle of causal closure (Sect. 2.3.3), but can easily mould and canalise them, once that they have already happened, in a network-like form according to the constraints present at that top level. Thus, such a causal influence is only *ex post* and *indirect*.

Functional Equivalence Classes

We speak of control when an organism is able to treat several environmental factors as *functionally equivalent*, that is, satisfying some functional need plus or minus in the same way. This is necessary, since those classes represent the grid (the code)

[67] Barbieri (2003, 2015).

[68] Auletta (2016b).

[69] Conant and Ashby (1970).

[70] Auletta and Jeannerod (2013).

[71] Auletta et al. (2008), Auletta (2012).

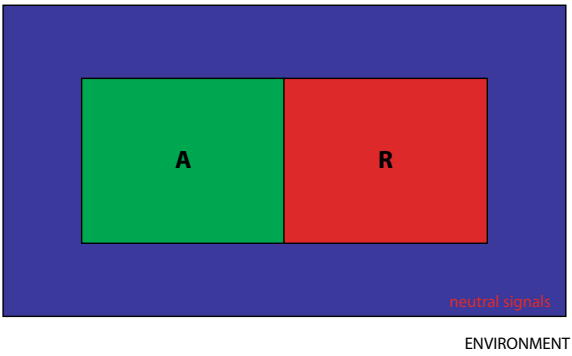


Fig. 8.20 The world of the *Escherichia coli*. Note that in such a way the organism establishes functional equivalence classes between different kinds of signals

that the organism 'imposes' on the world in order to be able to exert control at all. For instance, the bacterium *E. coli* divides its whole universe of signals in attractants and repellents on a background of neutral signals, as displayed in Fig. 8.20.[72] It is attracted by various sugars and amino acids and repelled by fatty acids, alcohols and other potentially noxious compounds. Note that, the two classes of signals regroup chemicals of different sorts, whose common denominator is represented only by the needs of the organism. Obviously, during its evolutionary history, some, possibly random, associations that for one reason or the other turned out to work well were selected.

Autarchy and Openness

The second crucial aspect is that the emerging and the manipulation of codes are, in fact, impossible without mechanisms that allow the organism to be partly *shielded* relative to the external (but similar arguments are true also for the internal) environment. Speaking in general, external signals (known as *first messengers*) never reach the genes. They are always transformed (in a process called *signal transduction*) into a different class of internal signals (called *second messengers*) and only these, or their derivatives, reach the genes (and we speak here of *signal integration*). In most cases, the molecules of the external signals do not even enter the cell and are captured by specific receptors of the cell membrane, but even those that do enter (some hormones) must interact with intracellular receptors in order to influence the genes.[73] Note that there is no necessary (or mechanical) connection between first and second messengers, because it has been proven that the same first messengers can activate different types of second messengers, and that different first messengers can act on the same type of second messengers. In general, the state of the cell is fundamental in order that a signal is 'interpreted' in this or that way, what is even

[72] Auletta (2011c, 2013a) and literature quoted there.

[73] Alberts et al. (1983, Chaps. 13–15), Barbieri (2015, p. 43), Auletta (2011a, Sect. 7.6.2).

more evident during the developmental stage of multicellular organisms.[74] What this example shows is that even a classical physical system like a cell (or an organism) is informationally shielded and nothing enters its informational-codified system. What enters the cell are at most certain chemicals that are necessary for metabolism, but the coming in and out of stuff is again strictly controlled by the informational system of the cell.[75] Thus, biological systems satisfy also the no-collapse determinism requirement of QM (Sect. 6.2.5).

The organism needs to be shielded only partly because it needs also to be *sensitive* to changes occurring in the exterior in order to act appropriately. Therefore, it needs some degree of plasticity. In other words, a too rigid structure, as a consequence of autarchy, would endanger the organisms's survival when facing significant environmental changes, while an exceeding plastic organisation, due to openness, would also endanger the organisms's survival due to lack of control.[76] This means that in general these two aspects (i) respond to different mechanisms but (ii) need to be integrated for finding a quasi-optimal trade-off (optimality is in fact rarely attainable and even dangerous for survival). This is accomplished through the fact that any organism (and any significant part or subsystem of an organism) can be represented as a *Markov blanket* (Fig. 8.21), after the name of Andrej A. Markov: it is a concept derived from the Markov chain (in Sect. 6.2.5 I have mentioned that

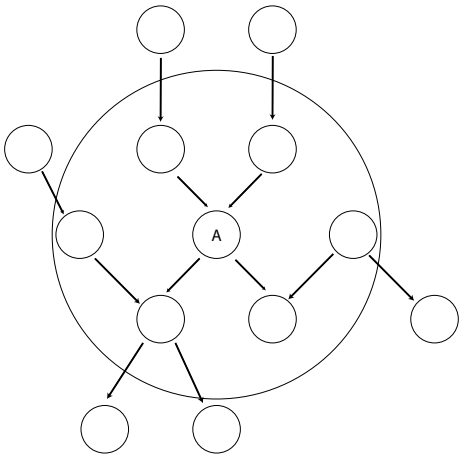


Fig. 8.21 The Markov blanket of node A is displayed by the large circle. It encompasses the node A itself, its parents (the nodes that influence it), its children (the nodes that it influences) and other parents of those children (nodes that have influences on the latter). In other words, the function of a Markov blanket is to introduce a protective belt around its node. The transition from prokaryotes to eukaryotes, recalled in Sect. 7.3.3, can be understood in terms of the creation of internal Markov blankets where there was none

[74] Gilbert (2006) West-Eberhard (2003). See also Auletta (2011a, Sect. 11.2).

[75] Auletta (2013a).

[76] A first, still immature, analysis in Auletta (2008a).

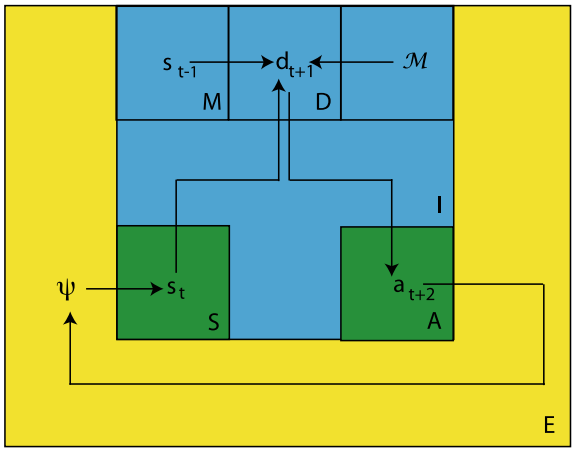


Fig. 8.22 The organism as a Markov blanket. In blue-green, the internal system I (the cell in the easiest case), in yellow, the external part (the environment E giving rise to a signal ψ). The internal blanket I has two sub-blankets that are shown in green: S stands for sensory subsystem, A for action (motor) subsystem. M stands for the memory slot, $\mathcal{M}$ for the model, and D for some 'decisional' subsystem: $\mathcal{M}$ represents the endogenous input that is necessary for control and it may be assumed to depend on genetic regulation. The memorised state is the result of a previous input. The decisional system compares the two sensory inputs at different times (s_{t-1} and s_t) and verify whether there is discrepancy with the model (which, for the *E. coli*, essentially tells: "approach the maximal sugar concentration"). Then, it produces the output d_{t+1} that gives finally rise to an action a_{t+2}. Relative to the original model of Friston et al. (2014) I have introduced several changes

each step in a Markovian dynamical sequence depends only on the previous step). The structure of the Markov blanket allows a *filtered sensitivity* to the exterior, i.e. a limited-range sensitivity such that the information is received, coded, re-elaborated (often in connection with endogenous signals), and canalised through the sensory–metabolic network of the organism giving finally some output (often some kind of action). Thus, living cells are in a permanent 'dialogue' with several environmental signals that are crucial for their survival. There are both genetic and epigenetic mechanisms for dealing with such signals, where "dealing" means to (i) have appropriately codified the signal, (ii) evaluated it in its vital 'significance', and to (iii) give rise to an appropriate series of reactions and actions.

The basic idea here is that the internal states of a biological system that persists in a changing environment can always be interpreted as representing or inferring the external states beyond the biological system's boundary. Recent (informational) free-energy models show that fundamental physical and chemical processes may have spontaneously generated this kind of structures whose prototype is the external bacterial membrane.[77] In this way, even the humble bacterium can display a complex regulatory system as shown in Fig. 8.22. Here, the input environmental signal ψ (that has been influenced by some previous action of the organism) is codified by the

[77] Friston (2012, 2013), Friston et al. (2014).

sensory system S, which is now in the state s_t and delivered to the regulatory system R that simultaneously receives inputs from the memory slot M represented by a previous sensory input s_{t-1}, and the decisional system D in the state d_t. Both the decisional system and the regulatory system contribute to the final action a_{t+2} of the action system A.

The Mechanical View

It may be noted that, according to the traditional view stemming from CM, the situation should be quite the opposite. In the nineteenth century, even in the case of more developed organisms, it was believed that perception and broadcasting of the information to the animal's brain were *mechanically* produced through the propagation of an electric impulse from the exterior or the peripheral sensory system to the brain. However, it was shown by von Helmholtz that the actual speed of transmission is too slow, and this in turn suggested that each neuron somehow mediates and regulates the impulse it receives before further transmitting it.[78] Today, we know that this delay is due to internal information processing and broadcasting (see also Sect. 6.3.4). In fact, the basic idea of CM that our representations and physical notions are copies of reality (that can be acquired in pure efficient mechanical way) simply contradicts the Darwinian theory, which has to be included here as integral part of our interpretation and fits very well with quantum-mechanical principles. In fact, the latter presupposes a strict separation and reciprocal independency between the external environment and its selective mechanisms, on the one hand, and the solutions (at a genetic, epigenetic, neural, behavioural, mental level) that the organism can give rise to (often as a combination of random events and constraints), on the other.

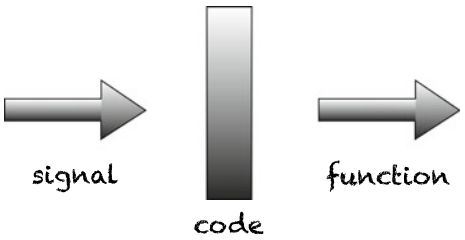


Fig. 8.23 The semiotic mechanism: the code as codifying a signal as a sign for a function or a functional action. Such an action or operation is performed fo its functional relevance, and in fact several operations can equivalent for giving rise to the same functional result

[78] von Helmholtz (1883, pp. 663–679, 881–885). See also Ledoux (2002, p. 44).

A Semiotic Mechanism

Thus, the complex mechanism through which a physical–chemical signal is classified according to the needs of the organism and evaluated for possible responses is in fact a *semiotic mechanism*,[79] as shown in Fig. 8.23: a signal is taken as a sign for something else (like something that either is good if food source or bad in other cases) and therefore as an index of appropriate action (in the case of *E. coli*, either swimming in that direction or tumbling for moving away: see Fig. 8.24).[80] Note that to tumble determines a pure random 'choice' of a new direction. Nevertheless, a combination of subsequent tumblings interspaced by comparisons of sensory inputs (the current one with the previous one occurred about 3–4 s before) allows in most cases to get the food.[81] It is interesting to remark that the sequence tumbling–swimming–tumbling could be considered as a particular realisation of the general quantum–mechanical mechanism of interaction–propagation–interaction. Note also that the first scholar to have pointed out this kind of semiotic processes is C. S. Peirce.[82]

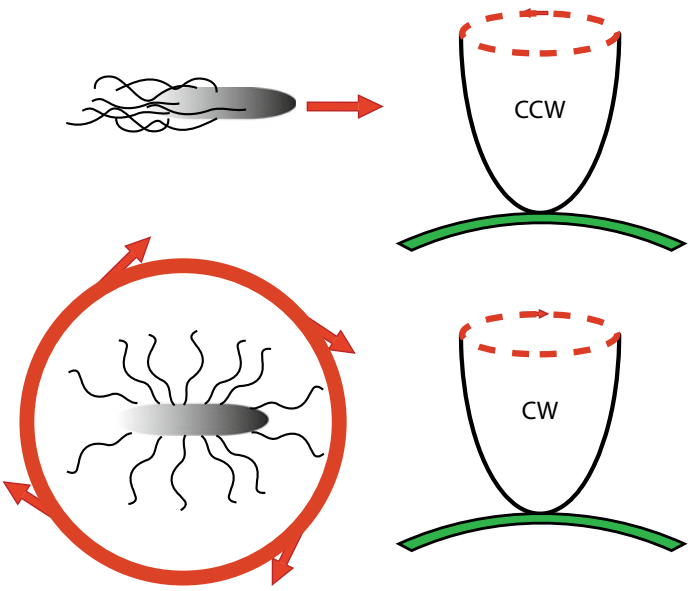


Fig. 8.24 *E. coli*'s motion. Above: Straight swim. In this case, the flagella turn counterclockwise. Below: Tumbling. In this case, the flagella turn clockwise. Adapted from Auletta (2011a, p. 267)

[79] Deacon (1997), Lemke (2000).

[80] Although bacteria do not possess a brain, molecular information processing performs not badly (Adleman 1994, 1998; Bourret/Stock 2002).

[81] Berg and Brown (1972), Jurica and Stoddard (1998), Shimizu et al. (2010). See also Auletta (2011c).

[82] Peirce CP (2019, 2.228, 2.247–48, 2.304, 1.540), Peirce (1907).

Invalidating Feedback

In such a context, for most organisms, the most important aspect of dealing with information is detection of a signal or, in more general terms, information selection (Sect. 6.3.4). Note that to detect a signal means in principle to receive *invalidating feedback*.[83] In fact, the more discrepant is the information selected relative to our (or other organisms') expectation, the more precious it is. This is the content of the notion of *surprisal* (or surprise) in information theory, where I recall that surprisal is expressed by a single component (e.g. $-\lg \wp_x$) of the entropy (3.43). Let us call this the negative side of information acquisition. This can be again framed in quantum theory of information, since I have shown that quantum systems are sensitive to environmental details (Sects. 5.2.6 and 7.2.1), and this can be considered as an invalidating feedback of the environment on them. Note again that all connections between systems in our universe are only *indirect*, in agreement with the basic quantum-mechanical constraint that codes can only be paired (Sect. 3.3.7). This will be true also for the processes of human cognition.

Adaptation

I have already pointed out that information selection would be completely unhelpful without information sharing with the exterior, and only the combination of these two aspects constitute information acquiring (Sect. 6.3.4). This is also true for organisms. In fact, all receptor organs are tuned on specific kinds of signals. How is it possible? This is the result of the long history of evolution (and here comes in play the Darwinian explanation). When some invalidating feedback is received by the organism (and this necessarily represents a source of instability), for the sake of survival this is obliged to correct its responses, which in the majority of cases are generated randomly. If the new responses lead to solutions that are able to deal better with the exterior, they are fixed through natural selection (those organisms that were unable to generate them are excluded from the game of life). Obviously, some individuals that had already expressed some mutations may be in a better situation for survival. In such a way, although the organism has not been instructed by the environment, as a result, it comes to share information with the latter and therefore becomes *adapted*.[84] Through this adaptation, it has also acquired some new information that is stored in its genome, functions, morphology or epigenetic pathways, and therefore most of those signals and environmental effects that played the role of invalidating feedback at the start are now inserted in the vital network (included environmental factors) of

[83] As pointed out in Peirce (1903a, p. 167). Note that here and in the following, I use the term *invalidating feedback* for denoting a signal that contradicts the expectations or indicates that things do not proceed in the right way. Sometimes scholars, included myself, lacking a generally acknowledged term, use *negative feedback* to this purpose, although the latter term has a different technical use.

[84] See Schrödinger (1958, Chap. 2).

the organism.[85] This is evident when we consider that all terrestrial animals have a respiratory system based on oxygen (a potentially noxious element) only because it was produced by the plants when they occupied the emerged ground. Let us call this the *positive side* of information acquisition. I do not need to insist on the relevance of information sharing for quantum systems. I remark that in both cases, selection mechanisms can give rise to correlations between systems that are blind to each other.

Knowledge as Adaptation

Resuming, control perfectly fits with the basic principles of quantum information: it (i) presupposes some kind of previous adaptation (trade-off between openness and autarchy, filtering) in order (ii) to catch (select) specific environmental signals, and, (iii) thanks to a model, (iv) interpreting them as signs for specific (equivalence) classes of actions provided of functional meaning. Humans are the result of billions of years of this evolution, and, again, without such a biological substrate, no operations on the external world, and thus no knowledge of the exterior, would be possible. Indeed, the basic mechanisms of control depicted so far are precisely the same in humans, also when we control our experiments. In this respect, we humans are not different from any other organism on the Earth.[86] And precisely for this reason, without our biological substrate, we would be unable to deal with the external world even for the goal of knowledge. What I am affirming is that our operations in our laboratories are not in principle different from the many activities that we attend in ordinary life and the latter are not very different from those that organisms exhibit in their exchanges with the surrounding environment. Essentially, what we do when we like to know something is to monitor a system catching signals that can tell us something about its state or behaviour (as it happens in the measurement process). It was Ernst Mach to hypothesise for the first time that knowledge should be considered as an extension of the biological dimension.[87] Obviously, the only difference relative to other organisms is that we humans perform our experimental operations for the *sake of knowledge*, or at least also for the sake of knowledge, while all other organisms deal with practical matters connected with their immediate survival. Somebody could object that also humans act for their survival. This is true and that is precisely what makes the connection between us and other organisms. Nevertheless, for humans, the issue of knowledge has also acquired a relatively autonomous dimension, as I shall argue now.

[85] See Auletta et al. (2013).

[86] Also Deutsch seems to agree on the relevance of control (Deutsch 2011, Chap. 3).

[87] Mach (1905).

8.3.2 Inferential Connection

Up to now, we have dealt with the physical (and biological) connection. However, humans have a mental dimension too, otherwise it would make no sense to speak of cognition. There is a fundamental distinction between *knowing* the solution to a certain problem and being able to *recognise* the solution: the ability do to the latter does not necessarily presuppose the capability to do the former (see also Sect. 2.1.1).[88] This distinction is strictly connected with the fact that there are two main aspects of observation and perception in general: the *segregation* of an object on its background (or also its individuation), and the *discrimination* among two or more objects or solutions.[89] Then, we can say that there is no knowledge without mental processes (whose inference is the most important component). This is constituted by our conceptual system and the element of consciousness recalled by Einstein (Sect. 5.1.2). The crucial point is that this conceptual system would be totally separated from the world without the mentioned biological and physical dimension in which it is rooted (that allows us to physically interact with the world and being part of the latter). In other words, cognition is necessarily *embodied* cognition.[90] In fact, in the same way in which correlations and information are and need to be physically embodied in our (and any) universe (Sect. 8.2.3), also our cognition is embodied. This is what establishes a connection between physics and mind avoiding the speculations about the intervention of the mind in the physical world or the psycho-physical parallelism (Sects. 2.2.3 and 3.1.1), speculations that would clash with our current understanding of how humans have evolved on our planet (Sect. 5.1.5). It is such an embodiment that gives a perspective-like character to our cognition and makes inferences necessary. In other words, embodiment is what both ensures and needs inferences for the sake of cognition. Let us first consider what is peculiar to the human mind.

The Mind

When we climb the ladder of animal evolution, we can see something quite extraordinary: most vertebrates use 2–8% of their basal metabolism for their central nervous system (CNS), primates use at most 10%, but humans use 20–25%.[91] Now, what is even more remarkable in the case of humans is that most of the brain activity is not addressed to operations that possess a direct survival value. In fact, when we perform an experiment in a lab, when we write a paper, when we read a book, when we compose or listen to music, when we go to the cinema and so on, we perform activities that could be considered meaningless from a biological point of view. It is true that some of the mentioned activities represent works or jobs and so help people to survive, nonetheless they do not have a direct biological meaning (and in fact all

[88] See Nielsen and Chuang (2000, p. 249).

[89] On this point see Auletta (2011a, Sect. 4.1) and references therein.

[90] Merleau-Ponty (1942, 1945), Schrödinger (1958, Chap. 3), Clark (1997).

[91] Mink et al. (1981), Raichle and Gusnard (2002).

of them can be understood only in the context of a human society) and it could even be asked whether they are performed for the sake of survival only, at least in some cases. Of course, a significant segment of our evolutionary history is occupied by the production of tools that had somehow a more direct survival value. However, also in this case, at least from the production of symmetric bifacials (about 800 kya) onwards, the conception and the making of objects require a significant time lag and a certain amount of cooperation, which means that these tools were conceived and planned in advance, what demands some degree of detachment from immediate practical needs.[92] It is also interesting that they show double symmetry, which confers to them even aesthetic value. Moreover, across human evolution activities have lost more and more their connection with biological functions and processes. In other words, this kind of activity means a 'violation' of the metabolic efficiency of the organism.

Then, the question arises about the significance of this evolutionary step and how can humans have survived and even evolved in those conditions.[93] The answer to the latter question is relatively easy: humans have used the part of the neural activity that is devoted to the intellectual activity for inventing and producing the means for controlling 'energy outside themselves'.[94] Thus, the emergence of the mental capabilities is connected with an important transfer of resources from the interior to the exterior, to technological means and devices. A nice example of this transfer of resources in the cognitive domain is represented by our use of diaries and mobile phones to help our mnemonic faculties. Moreover, instruments like a stone, a cart, a hydraulic pump substitute the efforts of many humans and therefore represent a significant economy of energy, so that the final balance between the energy employed (often by a small group of individuals) for conceiving and producing any of these and similar tools and the economy in energy that it represents (for a large population) is favourable to the survival of humans and to the propagation of culture as the specific mode of life in which these processes occur. This means that the intellectual life can endure only if it helps the self-maintenance of the organism and the human species by artificial means.

The crucial point is that all intellectual operations need to be, and in fact are, judged for their *rational value*. What is stable here are those results that are considered to possess a universal intellectual value and are therefore transmitted to subsequent generations or even spread geographically. They are in fact what remain across the generations as a good shared, at least potentially, by the whole humanity, while the often difficult and tortuous path that has led to them is only studied by a few historians. How can one judge this rational value? The most important aspect is the connection: rational plans, models, hypotheses and so on cannot be clusters of disconnected subunits but must instantiate *logical coherence*.[95] Another way to say this is that the mind does not tolerate breaks in those connections. This is again a

[92] Stout and Chaminade (2012). See also Auletta (2015b).

[93] Auletta (2015a).

[94] Crile (1941, p. 211).

[95] "Les vérités ne sont fécondes que si elles sont enchaînées les unes aux autres" (Poincaré 1897b).

distinctive difference relative to biological operations that can operate in parallel in a modularised form and do not need to be integrated as a whole. However, this confers to humanity an openness and a self-increasing capability that is unknown to any other biological species on Earth.

Comparison Between Humans and Other Animals

The issue of cognition is not to exert control on other systems (like for semiotic activity) but rather to verify and improve the matching between the Bayesian hypotheses generated (not always consciously) by the brain and external (both natural and social) reality (Sect. 6.2.5).[96] In all organisms but humans, this matching has only been a consequence of natural selection and is in general biologically hardwired for what concerns the behaviour (but less for the perception and the representations that the organism builds, which need to remain essentially bound to the individual ontogenesis).[97] Although higher animals like mammals display, e.g. motor schemes that are plastic, a closer look shows that only the assembly of motor subsegments is plastic (typically during a battle) while the subsegments themselves (like showing teeth or inarching the back for a cat) are again hardwired. Apparently, only the other primates show behavioural schemes that can be considered as provided of some plasticity as a whole[98] (although plasticity in biology is a wider notion, as stressed in the previous subsection), as displayed by their ability to reproduce (emulate) the action of other agents, as it is evident through the results of the research program on mirror neurons.[99] In other words, while models generated by other organisms have uniquely a biologically adaptive significance, the models generated by humans need to be *faithful* to the external reality, and here the comparison between expectation and data allowed by Bayesian inferences, especially when cognitive activity is involved, is crucial (see also Sect. 6.3.4).

Improvisional Intelligence

Let us understand the reasons for this need. Characteristic of the human brain and mind is what evolutionists have called *improvisional intelligence*: any environmental resource or signal is evaluated and used on the ground of the contingent and variable necessities of the human culture.[100] This implies, as it is notorious, that the functional needs of the human being are in part detached from the biological substrate. In other words, the same stuff could be a poison to certain purposes but a drug for others. This is why our hypotheses need to be faithful: the enormous variability of the

[96] Friston (2005), Friston et al. (2006), Friston and Stephan (2007), Friston and Kiebel (2009).

[97] Deacon (1997), Auletta (2016c).

[98] Marshall-Pescini and Whiten (2008).

[99] Rizzolatti et al. (1990), Fogassi et al. (2005).

[100] Cosmides and Tooby (2000), Cosmides and Tooby (2002). See also Auletta (2015b).

possible representations of the same object would generate disastrously contradictory behaviours if these representations were not 'theoretically' organised. And this theoretical organisation is represented by *concepts*, which are therefore second-level representations (they are organisers of first-level representations).

A couple of remarks appear appropriate here: (i) we can see how the transition from ordinary survival-centred mode of life to systematic inquiry about the world may have happened on the outline of E. Mach's idea of a continuity: the same expansion of needs and relations with the environment leads spontaneously to rational inquiry and so to science. (ii) Our ancestors have in this way discovered the world of *aspects* (Sect. 8.1.2), that is, the possible ways to consider an object through relations and interactions with it. Note that, when this relation is univocal and hardwired, as it happens for all other animals that we know, we cannot have such a step. In the case of other animals we speak in fact of *affordances*, that is, the species-specific ways in which a certain environmental detail is 'considered' by a species (like the trees as nesting places for birds).[101] Of course, also humans display affordances. Older stratifications are never wiped out by evolution. We note here the peculiarity of human's embodied cognition: on the one hand, it is cognition and therefore it goes towards a general understanding of the objects's aspects; on the other, we need to infer most of these aspects since they are not immediately accessible to our biological perspective.

Symbolic Systems

Now, such a new organisation of life implies and simultaneously requires a new form of dealing with information. While all organisms (included humans) display a semiotic activity,[102] only humans (at least on the Earth) have built symbolic systems, like mathematics, natural languages, music, and so on.[103] A *symbolic system* is a system that is *conventional* and capable to attach a *label* (a denotation like a name) to the different aspects and objects. There are some basic characters of all symbolic systems:

1. *Symbols are not bound to experience*: concepts like infinity in mathematics or force in physics (force is in fact never directly perceived but inferred given some effects that we take as induced by its presence) or even beauty or truth have no immediate empirical substrate. This is crucial, since improvisional intelligence requires the capability to extract information from the environment that is most of the time *hidden* to direct experience (need to be inferred), and this is a crucial part of our way of life, otherwise humans would still depend on hardwired perceptions (on the sensory manifestations of the objects), and thus on affordances. In other words, by looking at the external aspect of a mushroom often we cannot say if it is comestible or poisonous. Obviously, also many animal can make this distinction

[101] Gibson (1979).

[102] Hauser (1996).

[103] Auletta (2011a, 3rd part).

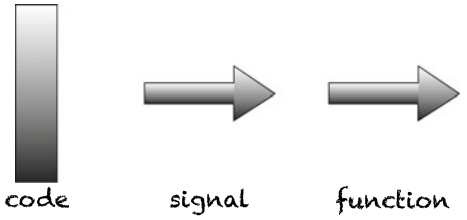


Fig. 8.25 The way in which a symbolic system works. Compare this mechanism with that displayed in Fig. 8.23. The basic elements are the same (more complex biological systems rarely change basic systems). They have only been remoulded conferring to the codes the leading role

but again thanks to mechanisms that are hardwired. At the opposite, we humans need to *learn* such situations.

2. This capability to expand the universe in which we live far beyond the immediate perceptual limits makes of all symbolic systems means that are *potentially infinite*[104]: with our ordinary language or mathematics we can in principle express any possible concept. This is crucial, since there are no a priori limits on the possible understanding and manipulation of an object. This is also physically true: the possible transformations that a system can undergo are indefinte.
3. Conventional systems of such a complexity cannot exist without the capability to share these codes with our conspecifics. This is crucial: since there is no longer a biological 'guarantee' of the correctness of our models, only their sharing in a common culture or social practice ensures the survival of our species. Thus, all symbolic systems have necessarily a *social–pragmatic dimension*.[105] This is why also science is practiced in academies, schools, labs and so on, in other words within a scientific community.

This makes the mechanism of symbolic dealing with information very different relative to those of semiotics (Fig. 8.25). In fact, the starting point here is not an (external) signal but the endogenous code that is imposed on a signal in order to give rise to an appropriate (second-level) function. For instance, we impose our phonetic and syntactical codes on the emission of sounds in order that a speech is produced (see also Sect. 8.1.4). I speak of *second-level* functions because any of the functions and processes of the mind are in fact equivalence classes of biological (first-level) functions, which are in turn equivalence classes of physical–chemical pathways, as seen in the previous subsection.

[104] Chomsky (2000), Hauser (2009).

[105] Auletta (2011a, Chap. 19).

Inferences

Peirce defined *inference* as "the power of judging the unseen by the seen".[106] This could also be said to be the power of judging the less known from the more known. From this point of view, inferences represent a refined form to deal with the problem of how to acquire information, which is already present universally in the biological domain. From this perspective, we could consider the whole of the biological evolution as a giant experiment for testing the rightness of the responses of organisms to new environmental challenges. It is not by chance that, also for other organisms, perception is essentially a hypotheses generator and testing.[107] For this reason, C. Peirce has stressed that perception is a kind of perceptual inference.[108] Of course, animals show inference-like processes, and it is likely that there are many gradations and intermediate steps between humans and advanced organisms like the other primates. However, there is a fundamental difference between inferences in the strict sense of the word (which require classes of objects and their aspects and therefore involve concepts) and kind of proto-inferences (developed by other organisms), which have the same general structure of inferences but are deprived of that characters and are rather based on basic biological mechanisms (and first-level representations).

General Structure of Inferences

This analysis allows us to settle the general structure of reasoning on a correct basis:

- Any inference must necessarily deal with a *rule*.[109] Indeed, the absence of a rule would compromise the very essence of inference, since in that case we would deal with a free association of ideas. We may perhaps distinguish here between laws, a general concept that also applies to the laws of nature (like the law of gravitation or the Schrödinger equation), which in most cases are in themselves mathematical formulae, and the logical rule, that is, the way in which we assign a general scope to such a law, whatever it is.
- However, an inference cannot consist in a rule alone, since this would be at most the definition of a rule but not an inference. Therefore, we must also be able to consider some *application domain* of the rule, that is, some case to which, even hypothetically, such a rule could be applied.

---

[106] Peirce (1866, pp. 405–406).

[107] von Helmholtz (1867, pp. 586–593, 601–602). See also Kandel (2006, p. 302), Auletta (2011a, Sect. 12.2). Peirce dared to say that Nature also makes inductions and retroductions (abductions) (Peirce 1898, p. 161).

[108] Peirce (1868b, p. 214), Here, not by chance, he says that perceptions represent premises for our further reasoning. It is also true that he did acknowledge that to perceive represents also a kind of discontinuity (Peirce 1903b, p. 191–94), in accordance with our interpretation of information selection. See also Margenau (1950, Sect. 4.1).

[109] Peirce (1865, pp. 259–261), Peirce (1866, pp. 365–369).

- Nevertheless, if a set of statements were constituted of these two statements only, it would not be an inference but at most the enunciation of a law or rule and the exhibition of a possible example or application domain. From an inference, we expect also some *consequence* out of the first two statements, i.e. a derivation of some conclusion from some premises. Note that in such a case we are assuming that the premises *are true*, while in any logical proposition nothing of this kind is assumed so that we can neither speak of premises. For instance, $XY \rightarrow Y$ is a tautology, but nothing is said about the truth value of XY or Y (I recall that all relations of this kind are in fact contingent).

Then, an inference is necessarily made of these three propositions, which may also be called thesis, condition and conclusion or even rule, subsumption under the rule, result.[110] I have shown that we cannot have a smaller number of propositions, but it is also immaterial to have more since this would only be a multiplication of one of the above statements. Note that already Aristotle explicitly affirmed the necessity to have three propositions in any proof.[111]

Scientific Inferences

According to the improvisional–intelligence explanation, the inferences that we make in science are only a specific form of the mentioned general kinds of human inferences. Following Einstein's insight (quoted in Sect. 5.1.2), we could define *scientific inferences* as follows[112]:

> A scientific inference is a rationally grounded connection between some principle or law and some experience, resulting out of a systematic enquiry with the goal to promote knowledge.

At the opposite, many everyday inferences show still a kind of connection between some general or universal statement and individual experience but occasionally and, at least in most cases, for practical purposes only. Moreover, scientific inferences need to be Bayesian hypotheses while everyday inferences sometimes lack such a character. The problem here is represented by the difficulty to perform correct explicit and conscious inferences. Therefore, I assume that both everyday and scientific inferences are particular instances of a larger genus of inferences that could be called *natural inferences* (as distinct from the pure logical propositions, i.e. tautologies). In other words, I am not considering here the so-called formal inferences or formal deductions, but inferences dealing with empirical sciences and empirical knowledge.

---

[110] Peirce (1865, p. 259), Peirce (1866, pp. 362–363).

[111] Aristotle An. post. (2019a, 75a38-75b2 and 76b13-16).

[112] Auletta (2013d).

Extension, Intension, Information

In order to treat adequately the issue of inferences, we need to recall the distinction made by Peirce between *extension* (the class of objects that are collected according to some criterion) and *intension* (the properties or characters that are shared by the members of a class and that therefore conceptually determine the class).[113] I also recall that extension is denoted (and therefore we can also use the word *denotation* for indicating extension) whilst intension is connoted (what justifies the word *connotation* for indicating intension). Sometimes, Peirce prefers to use the words breadth and depth, respectively, although the meaning is slightly different.

I like to consider here an interesting conceptual step done by Peirce, when he distinguishes among connotation $\mathcal{C}$, denotation $\mathcal{D}$, and *information* $\mathcal{I}$, and notes that connotation is the inverse of denotation only if the information remains the same.[114] Information, on the other hand, is defined by Peirce as connotation times denotation, but could be better defined by saying that it represents the information needed for correctly denoting, and ascribing correct connotation to a collection of objects. Therefore, the propositions involved in any inference must always be a connotative statement, a denotative statement and informative statement. In the case in which we do not have such information (as witnessed by some experience representing counterevidence), we need to improve our knowledge.[115] This implies that we can increase the knowledge of a class of objects subsuming it in a larger class, thus by losing its connotation without appreciably changing its proper denotation; or vice versa, we can increase its determination by narrowing its denotation, without appreciably changing its proper connotation.

Deduction

This allowed Peirce to cast all forms of natural inferences in deductions, abductions, and inductions.[116] *Deductions* cross hypotheses for deriving specific *expectations* about facts. These hypotheses are the result of *previous* experience and knowledge. Any hypothesis that we have ground to assume to be true, like a scientific theory that has already received some credit, is crossed with more specific kinds of hypotheses for understanding what would be the *consequences* of this theory and therefore also for opening new domains of application.[117] As said, the most general form that the general hypotheses take is in terms of the laws or principles of nature. In particular, when dealing with deduction, we have the following three statements[118]:

Conditional knowledge Resulting from previous hypotheses and experience

[113] Peirce (1868a, pp. 72–74). See also Auletta (2013d).

[114] Peirce (1865, pp. 187–189 and 284–286).

[115] Peirce (1868a, p. 83)

[116] Peirce (1878)

[117] See Margenau (1950, Sect. 5.6).

[118] Auletta (2009).

Identification An actual procedure for singling out a class (a domain)
Result Expected future events or consequences

Einstein was very clear about this structure of deduction, since he says[119]:

> But if experience is the beginning and end of all our knowledge about reality, what role is there left for reason in science? A complete system of theoretical physics consists of concepts and basic laws to interrelate those concepts and of consequences to be derived by logical deduction. It is these consequences to which our particular experiences are to correspond, and it is the logical derivation of them which in a purely theoretical work occupies by far the greater part of the book [of Euclides]. This is really exactly analogous to Euclidean geometry, except that in the latter the basic laws are called 'axioms'; and, further, that in this field there is no question of the consequences having to correspond with any experiences. But if we conceive Euclidean geometry as the science of the possibilities of the relative placing of actual rigid bodies and accordingly interpret it as a physical science, and do not abstract from its original empirical content, the logical parallelism of geometry and theoretical physics is complete. We have now assigned to reason and experience their place within the system of theoretical physics. Reason gives the structure to the system; the data of experience and their mutual relations are to correspond exactly to consequences in the theory.

Abduction and Induction in General

By paraphrasing Peirce's (and Einstein's) formulation, I can say that the problem of how to make an abduction or induction is how to make any general statement with reason (in good agreement with logical laws).[120] According to Peirce, the whole point of induction is whether or not it is better to prefer an increase in extension that could involve some error.[121] However, the non-apodeictic nature of induction does not derive from the inferential form itself (which is analogous to that of deduction) but is related to the initial decision to prefer a certain solution to another.[122] This is also a matter of evaluation and rational choice, where obviously the estimation of probabilities plays a role, as it must be for Bayesian inferences (Sect. 6.2.5). In fact, we deal here with a basic incertitude that cannot be fully wiped out. Similar considerations are true *mutatis mutandis* for abduction. It is suggested that we have here to deal with a new term that has more intension than the original term. The possible increase of formality represents an increase of truth but the increase in intension a chance of error.

Having formulated the terms of the problem, the natural question arises: how can we have a solid and consistent advancement of learning when our deductive reasoning is impaired? To solve this difficult conceptual problem is not easy. It is again Peirce that gives us an important hint about the way to consider it.[123] Deductions are explicatory. What explicitly appears in the conclusion is considered as implicitly

[119] Einstein (1934, pp. 164–165).
[120] Peirce (1865, p. 179).
[121] Peirce (1865, pp. 187–189 and 284–286).
[122] Peirce (1865, p. 292).
[123] Peirce (1866, pp. 458–471).

contained in the premises. In other words, any direct (positive) explanation consists in saying that what a word denotes is what is meant (connoted) by the word or also whatever is contained in the content of a term belongs to whatever is contained under it. The principle of induction is that whatever can be predicated of a specimen of the extension of a term is part of the content of that term; but this specimen must be taken at random and not selected as belonging to a narrower extension (some subclass of the original class). Precisely for this reason, induction consists in denying that the specimens can be confined by a certain connotation $\mathcal{C}$, since this would result in an ungrounded limitation of extension. Therefore, we need to find an appropriate connotation due to the assumed failure of a previously acknowledged rule. The only point is that we cannot derive this law or connotation directly by logical means alone. On the other hand, abduction shows that the failure of the previous expectation is due to a definition of a previous denotation (the chosen sample) $\mathcal{D}$ that is not correct. We need to find a new denotation with a more specific connotation. However, we cannot positively find through pure logical means which will be the new denotation enabling us a correct identification procedure.

We should take care of at least of the following criteria, which hold for abductions[124]:

> Facts cannot be explained by a hypothesis more extraordinary than those facts themselves. And
>
> Of various hypotheses, the least extraordinary must be adopted.

In other words, we cannot introduce a hypothesis if not because we assume that the explanation is more readily acceptable than the explanandum. Moreover, we must try to lower this extraordinarity as much as possible. We could similarly say for inductions that

> One cannot build sets of objects sharing some character without this resulting in some consistent advance in knowledge.

Summarising, both induction and deduction are reasoned forms of inference which take into consideration some consequences having made some partial (but negative) experience. Let us consider these processes in detail.

Abduction

Let us consider a deduction of the form[125]:

$$\text{If } X' + Y \text{ and } Z' + X \text{ , then } Z' + Y, \tag{8.139}$$

where X and Z are classes of objects (the latter still to determine) while Y, expressing an aspect, is a set. As expected, deduction has the form of a logical implication and is

[124] Peirce (1866, p. 452).

[125] Auletta (2017).

therefore a tautology (Sect. 8.2.1).[126] Let us now assume that experience contradicts the expectation expressed by this inference, so that we have ZY'. This is now a *fact* (Sect. 6.3.4). Since experience has falsified our inference, one of two premises must be wrong (I do not consider the extremely rare case in which both are wrong, what would denote a serious inconsistency in our previous knowledge). In other words, from the denial of a connotative statement, we infer either the denial of the informative premise (the first one) or of the denotative (the second) one.

Which one? There is no logical means for knowing that. The only thing that we can make is a Bayesian *evaluation*, what means trying to find the likelihood of certain hypotheses and thus accepting a certain incertitude (Sect. 6.2.5). It seems that we pass from the logical certitude to incertitude. However, that certitude was illusory, since deduction already expressed an *expectation* and nothing more, as it is characteristic of all natural inferences. Incidentally, I note that this fits well with the fact that a quantum state provides probabilities for certain outcomes and therefore represents a collection of expectations or expectation values.

If the counterevidence displayed by experience is a single one, it is more likely that the error is in the denotative statement or at least it is more convenient to assume so since it does not represent the sum of our knowledge so far. In such a case, we perform the first step of an abduction which takes the form[127]

$$\text{If } X' + Y \text{ and } ZY' \text{, then } ZX'. \tag{8.140}$$

Why only the first step? Because the conclusion is evidently *negative*: its states that the objects Z are not well denoted by the equivalence class X. Why? Because their behaviour shows some additional features that the ascribed denotation cannot account for. However, the problem now is: what is the appropriate denotation for solving the problem? It is here that logic cannot help. The only possibility is to be able to guess such a solution. How? Through a pure association. When this association is insightful we call it *insight*.[128]

First, let me clarify why an association is not determined by logic, as clearly understood by Hume.[129] It is the expression of randomness in the context of mental–neural processes. The most associative faculty of the mind is memory, so it can help to analyse what are its distinctive characters. Memory works essentially along two possible ways (as also acknowledged by Hume himself[130]):

---

[126]Deduction (8.139) has the form of a Aristotle's 1st–figure syllogism, in particular of (what in the Middle–Ages was baptised) *Barbara* (Aristotle An. pr. 2019b) See also Auletta (2013d).

[127]The following inference has the form of a Aristotle's 2nd–figure syllogism, in particular of (what in the Middle-Ages was baptised) *Baroco* (Aristotle An. pr. 2019b). For people interested to these issues, see again Auletta (2013d).

[128]Poincaré (1905, p. 32). Poincaré spoke also of a "creative virtue", distinct from logic, even in mathematics (Poincaré 1902, p. 32).

[129]Hume (1739, Book I, Sect. 6), Hume (1748, Sect. 4, Part I).

[130]Hume (1739, Book I, Sect. 4), Hume (1748, Sects. 3 and 5, Part II); see also Auletta (2011a, Sects. 20.5, 20.6).

- Either we collect events or objects together because they happened in temporal or spatial contiguity (as when we recall a certain situation because we were thinking at ninth eleven and that situation was the context in which we received that news first time). This kind of association is of *metonymic* kind.
- Or recalling a detail of an object brings to our memory a similar character of another object (as when we recall the red of a dress by looking at the red of the wine that we are drinking). This kind of association is called *metaphoric*.

I recall that I have already introduced these distinctions in Sect. 5.1.2 when dealing with inferences: metonymic and metaphoric inferences are grounded on the respective associations. Thus, metonymic association is centred on contextual interconnections, while the metaphoric one on taxonomic relations. It is clear that none of these associations occur according to some rule of logic (I do not consider causal relations as associations, as Hume does, as recalled in Sect. 2.4.1: I assume that are in fact the result of inferences). When dealing with abduction, we deal with the first one.

New Denotation

Now, suppose that we get a good idea (it does not matter how by now). In general, especially if experienced, we know from the start if an idea is viable or at least we know that most of those that occur to us are not.[131] How need such an idea to be? It is here that logic selects and moulds the spontaneous associations that we may have produced, and here Bayesian inferences play again a role. If we are looking for a new denotation, logically this needs to represent a *narrower* (more specific) denotation than X, say $\mathcal{X}$. In fact, only this can guarantee to us not only a major precision in considering the denotation of the collection Z but also the correct relation with the set Y. So, if we are fortunate, we can replace the old denotation by the new one and proceed in our inference as follows:

$$\text{If } \mathcal{X}' + Y \text{ and } Z' + \mathcal{X}, \text{ then } Z' + Y. \tag{8.141}$$

Note that the first premise (expressing the rule) has not been touched, since, whenever $X' + Y$ is true, also $\mathcal{X}' + Y$ is true due to the fact that $\mathcal{X}' + X$. Note also that the inference (8.141) is in fact a new deduction, that is, the generation of the same expectation (that the objects Z are Y) but having assigned a narrower denotation to Z. This explains why deductions deal with expectations, being the result of a previous abduction (or of a previous induction). Summarising, the whole process of the abduction consists of three steps, two logical-formal parts and an associative part:

1. First we infer that it is not true that the collection of objects Z has the property Y under the condition X, as expressed by the (8.140).
2. Then, thanks to an association, we get a narrower equivalence class $\mathcal{X}$.

[131] Kuhn (1962).

3. Finally, we perform a new deduction by inferring that the collection of objects Z has the property Y under the new condition $\mathcal{X}$, as expressed by the (8.141).

Induction

Now, suppose that, in the course of our scientific research, we gather more and more counterevidence such that, for several collections of objects $Z_1, Z_2, Z_3, \ldots$, and the same also for several conditions $X_1, X_2, X_3, \ldots$, our expectation that the former collection of objects display the property Y is falsified. Then, we are driven to the conclusion that the problem is not with the chosen objects and their denotations but with the law and in particular with the assigned property.[132] In other words, from the reiterate failure of deduction (8.139) we infer[133]

$$\text{If } ZY' \text{ and } Z' + X, \text{ then } XY', \tag{8.142}$$

where obviously we usually make several inferences of this kind with the different collections of objects. Now, this result is *inconclusive*. In order to get a conclusive solution to our problem (a new law or a new theory), we need again an insight and therefore to go through an associative moment, in this case a metaphoric extension. Now, suppose that we get such an insight, i.e. a new set $\mathcal{Y}$. How needs this set to be? For logical reasons, larger than Y. In fact, only this allows us a generalisation in order to solve the problem and also the correct relation with the class X. In fact, having $Y' + \mathcal{Y}$, we can infer

$$\text{If } X' + \mathcal{Y} \text{ and } Z' + X, \text{ then } Z' + \mathcal{Y}, \tag{8.143}$$

where I attract the attention of the reader on the fact that the second premise remains untouched. We have again a new deduction but this time generating a new expectation. Summarising, we have

1. An inconclusive inference through which we conclude that the objects X are not Y without specific conditions, as expressed by (8.142).
2. Through an associative moment we get the insight of a new, broader, property $\mathcal{Y}$.
3. An inference establishing a new law according to which it is the broader property $\mathcal{Y}$ to be the right connotation for Z, as expressed by (8.143).

Examples

A splendid example of deductive inference is, not by chance, displayed by EPR (Sect. 3.3.1): from the laws of QM, the criterion of reality and the separability prin-

[132] Auletta (2017).

[133] The following inference has the form of a Aristotle's 3rd–figure syllogism, in particular of (what in the Middle-Ages was baptised) *Bocardo* (Aristotle An. pr. 2019b). See also Auletta (2013d).

ciple they tried to deduce certain new consequences, in particular the possibility to ascribe reality to non-commuting observables and therefore infer the incompleteness of quantum physics. On the other hand, Bohr and Schrödinger have preferred to employ in this debate inductive and abductive inferences, respectively (Sects. 3.3.2, 3.3.3). Bohr inferred that the reality criterion of EPR cannot apply to QM, while Schrödinger inferred a new character of quantum systems, namely entanglement, that leads to the rejection of the separability principle, which is less general than the criterion of reality. Other examples are represented by Planck's abductive style, according to which there is a new class of objects (light with discontinuous properties) but no rejection of classical laws (Maxwell's electromagnetic theory), and Heisenberg grounding induction, that using operators as a formal generalisation relative to variables and functions would solve the problem of the several discontinuities found at experimental level (Sects. 1.1.1 and 1.2.2).

Inductions as a whole (from a previous deduction to a new deduction) can be understood as Bayesian inferences, which, I recall, tell us what is the probability that we choose a correct hypothesis given a certain set of data, while abductions as a whole (from a previous deduction to a new deduction) can be understood as inversed Bayesian inferences, which tell us what is the probability that those data are in fact observed (that is, that they are real) supposing that our hypothesis is true (Sect. 6.2.5).[134]

Controlled Operations as Bridge

Biological operations, inferences and the measurement process have interesting commonalities. In the case of biological operations, we have

- The response to a stimulus,
- Thanks to the amount of surprisal represented by the latter and
- Some previous adaptation to (correlation with) the environment.

On the other hand, we have conceived all inferences as dealing with

- Inferring a (possible) outcome,
- Given facts that we know and
- Some previous knowledge that we have about such facts.

Note that the facts that we know are experienced through a basic mechanism of invalidating feedback (and therefore surprisal) as for all other organisms. Moreover, as recalled, evolution through natural selection and adaptation is a kind of giant experiment cumulating information on the environment. The more species become complex, the more they are able to deal with a wider spectrum of environmental stimuli.

On the other hand, we have defined measurement (as the paradigmatic form of physical operation, therefore representing a specific but insightful case of information

[134] Auletta (2011a, Sect. 18.4.4), Dehaene (2014, pp. 117–118).

acquiring) as a procedure that consists of the following elements (see Sects. 5.2 and 6.3.4):

- We are able to attribute a property out of an initial state of a system representing unknown information,
- Thanks to a detection event representing information selection and
- A coupling between an apparatus and this object system that displays information sharing.

The crucial point is that also in the case of inferences, we have a conditional outcome given certain selection events (the point-like experience that we make with certain facts) and information sharing with the conditions in which that outcome should happen (the knowledge that we have already acquired thanks to previous experiences and inferences). And I have already said that any measurement could be considered as an inferential process, since the attribution of a property is essentially an inference to be kept distinct from the physical operations in a strict sense. Moreover, the fundamental structure of semiotic and symbolic processes is essentially the same information acquiring. In semiotics, we have that a codified signal is taken as sign of something else and is an index for a functional action (Fig. 8.23). Thus, an action is selected on the basis of the reception of a variety (the signal) which can be put in connection with an information that the organism shares with the environment. With symbols (Fig. 8.25), we select a code for influencing a signal able to give rise to a function–action with a certain significance in that code. Thus, across the domains of Nature, we have essentially the same structure of information acquisition that we have discovered with QM.[135] What does change are only the relations among the fundamental terms. But this is something that is already true at a pure quantum-mechanical level (Sect. 6.3.4).

In fact, the previous considerations make us understand that the whole process of knowledge (and of scientific progress) is only a particular case of *error correction* (Sect. 6.3.2): both abduction and induction are in fact sophisticated and rational (inferential) forms of error correction. Let us examine this point.

Hypotheses and Experience

The problem of knowledge is how to have an opportune mediation between general hypotheses or statements and experience, which by definition is constituted by perception of individual events and objects or tokens and even by basic stimuli (and is also made by individuals).[136] As seen, direct experience, ultimately, is indeed only a invalidating feedback relative to some of our assumptions but tells us positively

[135] This is the main thesis of Auletta (2011a).

[136] Poincaré (1897a).

nothing about further assumptions to formulate. Nevertheless, this function both is the sole source of novelty and establishes empirical certitude.[137]

I think that one of the greatest contributions of the Austrian–Jewish–British epistemologist and philosopher of science Karl Popper (1902–1994) is to have fully accepted this conclusion, although he could only admit deduction as a scientific form of reasoning.[138] Now, our general statements or hypotheses, although representing the possible responses to the solicitations coming from experience, are never able to fully account for facts, and sooner or later need to be abandoned as a consequence of further experience. Generally speaking, we can say that our theories and hypotheses are *underdetermined* relative to facts or factual assertions, as far as they are not able to single out an appropriate class of facts if not in conjunction with some further experience.[139] On the other hand, facts or factual assertions are *overdetermined* relative to theories, as far as they are too specific in order to reach some general conclusions without the help of further theoretical assumptions, which by definition need to be less determined or more general than experiences.[140] We can see in this duality also a restatement of the dichotomy globality/locality (Sects. 3.1.3 and 5.1.3–5.1.4).

A Circle

Resuming, when we face negative experience, a good insight can help us to solve in a positive manner either an abductive or an inductive inference: we have either found the proper denotation (a new class of objects) we look for or we have found the law or the theory that we were missing (and therefore a new connotation). If this process is occurred in a satisfactory way, avoiding false solutions consisting in a bare abstract split and recombination of classes, we likely get the idea and then shall try now to test it. This is the business of a new deduction (Fig. 8.26). So, the process of inference-making starts again.[141] If we have found the solution to the problem by abduction, what we try to do is to update the theory and try to see if this new character (and the relative class of objects) accounts for predictions that we hope will be helpful and fertile. However, as time goes on, this will have an end, and any theory cumulates in the long time span such an amount of drawbacks that we need to perform an induction (to become aware that we need to find a new theory), avoiding

[137] Poincaré (1902, p. 157). Although the great scientist seems to consider experience in positive terms according to the traditional (and erroneous) inductive explanation of knowledge (Poincaré 1902, p. 158).

[138] Popper (1934, pp. 16–17). See also Deutsch (1997, p. 62). At the opposite, C. Peirce, although having much more clarity about the logical and inferential background of science, showed still some incertitude about the non-positive (non-instructive) character of experience. On the issue of induction in Peirce see also Shimony (1970, pp. 231–235).

[139] Margenau (1950, Sect. 5.1). See also Friedman (1983, Sect. 7.1).

[140] Auletta (2013b).

[141] Auletta (2017).

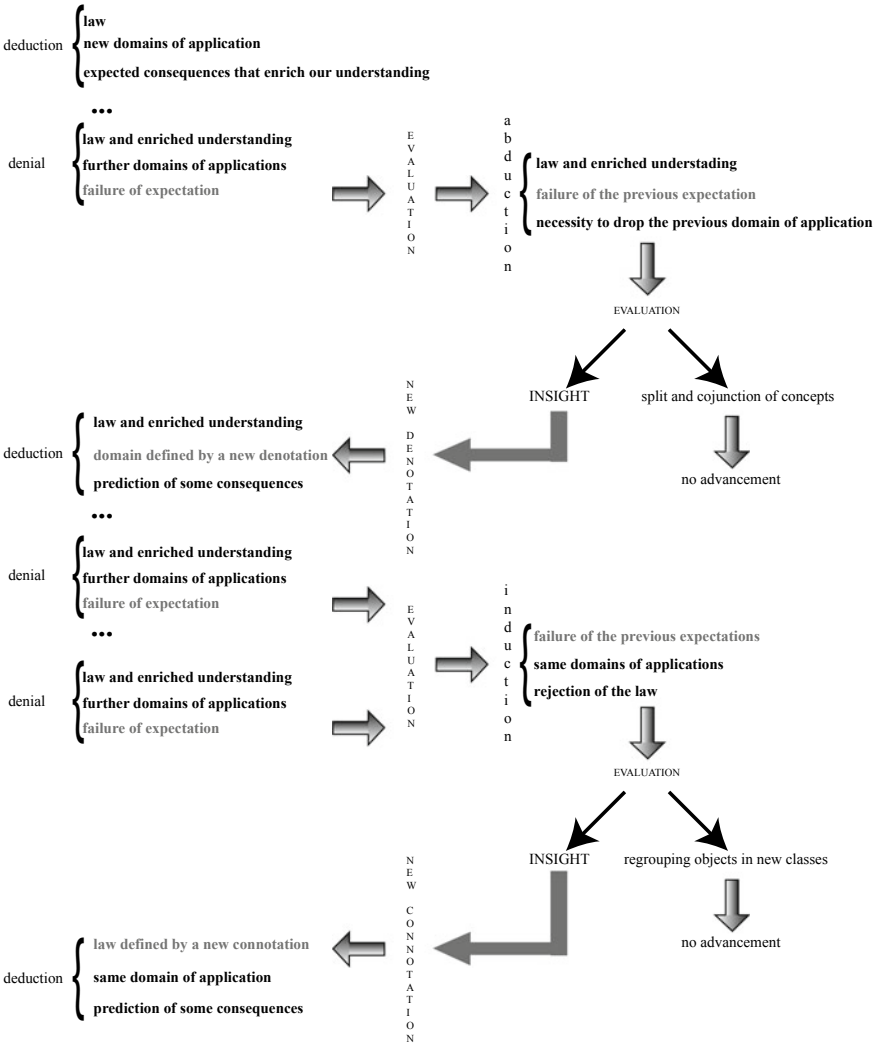


Fig. 8.26 The circle of inferences. We always start with some form of deduction in order to enrich our knowledge with some interesting consequences, and it is only when contradicted by experience that we are forced to undergo an evaluation process in order to correct the error. The latter con be very complex and all forms of inference can be used. The issue is finally to evaluate if there are reasons for dismissing one of the two premises. In general, when negative experiences are few, it is better to reject the chosen sample of objects or the application domain of the law (to perform an abduction). However, we need always to evaluate if it is convenient to step towards the search for new laws, principles, theories, what is the business of induction. Note that all alternatives to insights are attempts at positively solving the problem by logical means but lead to failure, because neither the logical manipulation of concepts for building artificial concepts, nor regrouping objects in new artificial classes can produce new knowledge

again false solutions consisting here in abstract regrouping objects in new classes. If we succeed, the process starts again but a higher level. Peirce was very well aware of the circular character of inferences.[142] From this examination, it can be said that we start our knowledge with doubts and problems (and bad formulated hypotheses), we go through inferential processes and thanks to the test of the experience (and therefore the interaction with a real world) we correct our initial assumptions and give our assent to certain consequences or conclusions.[143]

Nevertheless Knowledge Progresses

In sum, deduction is by its very essence *anterograde*, because it starts with assumptions and focuses on the possible experiences of new facts that we expect. Not by chance, this was Einstein's preferred inference. On the contrary, both abduction and induction are *retrograde*, because they start from experiences that have contradicted previous expectations and go back to the foundations of knowledge in order to see what assumptions need to be rejected with the aim to make new discoveries (either new properties or new laws able to fill the gap). The three forms of inference reinforce each other, with a sort of dynamical–cyclic evolution across the historical time with bursts of innovations followed by periods of consolidation and rearrangement, as it was well understood by the American physicist and probabilist Edwin T. Jaynes (1922–1998) and Thomas Kuhn,[144] although I would avoid a too sharp contraposition of these different stages of the evolution of a theory. Indeed, T. Kuhn distinguishes two fundamental and recurrent phases in the development of scientific research (and of human culture in general): one called normal and the second critical or revolutionary. In the normal phase, the scientist draws some consequences from a theory that is reputed correct and is accepted by the scientific community. It corresponds to what I have called the deductive stage. The phase of crisis of a certain scientific paradigm leads to revolutionary bursts in which different alternative theories are proposed to the scientific community as a solution to the problem which the old theory (or paradigm) is incapable to solve. This corresponds to what I have called inductive stage. The scientific community, which is in general conservative, in these periods becomes more willing to accept new solutions. When the community gives its assent to a certain theory (or paradigm), this becomes dominant, and so a new normal phase starts.

There are several points in Kuhn's approach that deserve a critical scrutiny. First, he supports the idea that different paradigms are not comparable as far as even the criteria that individuate a scientific field may change, so that problems and concepts can be fully redefined. This may be in part true, especially if we consider the passage from the Aristotelian–Ptolemaic view of cosmos to the Galilean physics (and this

[142] Peirce (1901, pp. 96–97). See also Peirce (1893, p. 196).

[143] Auletta (2011a, Sect. 6.1, Chap. 12), Auletta (2011b, Sect. 2.2). See also Deutsch (2011, Chap. 12).

[144] Jaynes (1967), Kuhn (1957, 1962).

only to a certain extent), but this is certainly an overstatement when applied to the developments of modern science. The image that we receive from Kuhn is a science that proceeds through jumps without an effective accumulation of knowledge, while in fact such an accumulation seems distinctive of the scientific research. Obviously, Kuhn has contributed to destroy the idea of a linear and continuous progress, but this does not imply the absence of progress. Moreover, according to Kuhn, the last instance that decides about the value and the acceptance of the different paradigms and theories is the scientific community. This is true and has been important to have stressed this. However, the (tortuous) progress of knowledge is not an issue of sociology of knowledge only. In fact, the scientific community takes its decisions also by judging on the basis of criteria that are *intrinsic* to the scientific issues, otherwise developments that are initially marginal could never be subsequently accepted by everybody (as it often happens). We can say that the objective force of scientific truth is finally stronger than personal prejudices. Moreover, at the first stages of a theory, the consequences that are drawn from it are themselves revolutionary as far as they open new fields that were often not foreseen in its original formulation. For instance, the idea of a universe in expansion that was drawn from Einstein's relativity was so new that this scholar was even unwilling to accept it. Likely, the most important limitation in Kuhn's view is not having considered that in the maturity phase of a theory substantial amendments are introduced. This so-called abductive stage is a kind of middle situation between the two phases considered by Kuhn.[145] It is likely that Kuhn would object that such a phase is itself conservative. Nevertheless, in this way, we would miss its very nature that is addressed to introduce significant changes in a process of error correction. Finally, even when a theory is dismissed and substituted with another one, parts of the old theory are imported in the new theory (as it is evident for QM relative to CM) and the laws of the former one still possess explanatory and predictive power at their specific scale and in their domain (special and general relativity have not abrogated the use of pre-relativistic physics for explaining physical phenomena, e.g. happening on the Earth).[146]

Thus, the progress of knowledge is not linear but nevertheless it is a progress through different kinds of correction. The main directions are towards (i) more and more general laws and principles (it is a forgetful or a right adjoint, that is, abstraction) along inductive lines and (ii) a growing specification of its objects (free or left adjunction) along abductive lines (Sect. 8.1.4). It seems to me that the biggest problem with current epistemology is the wrong idea that theories are conventional (Sect. 3.1.1). Obviously, there are also conventional aspects. However, the progress of knowledge is measured on the capability to get results that are less and less conventional coming out of a situation where conventionality is more dominant. Thus the cyclic–dynamical process of error correction allows, in fact, for sharing more and more information with the world. I think that Poincaré aimed at something in this direction when he said that our scientific principles are both (i) based on experience

[145]To a certain extent, this stage corresponds to the so-called auxiliary hypothesis that does not change the core of the theory, as proposed in Lakatos (1976).

[146]Poincaré (1902, Chap. 10).

and (ii) are postulated. He stresses their conventional character but nevertheless adds that they are not fully arbitrary since they have been assumed in our confrontation with experience.[147]

Again on the Principle of Sufficient Reason

This allows us to try to lead all the three forms of inference under a common heuristic principle, which is derived from the above examination of their procedures.[148] We might even dare to call this a restatement of Leibniz's famous principle of sufficient reason (Sect. 2.4.1), which could be enunciated as: everything has a reason to be as it is. We have in fact seen that this formulation does not work for QM. The reformulation that I propose is then the following one:

> Of everything a ground can be found that enables some results that we hold to be true.

It is not difficult to see that this formulation covers both anterograde (deductions) and retrograde (abductions and inductions) inferences. Note that this reformulation of the principle of sufficient reason is made in such a way to avoid any restriction to causal–mechanical explanations. We can summarise this principle in a very schematic way as follows:

Abduction Of everything a ground can be found
Deduction That enables some results
Induction That we hold to be true

Somebody could object that this cannot be true for QM as far as we cannot find a ground of any random event such that the latter be enabled by it. This is true, but is not an objection to the principle, for the simple reason that inferences (and theories) deal with general aspects and not with particular or individual ones. It is also true that in most experiments we like to know precisely what happens to the single system under observation. However, in order that this represent an experimental test at all, we implicitly assume that, in the premeasurement stage, any other system will behave in a similar way in similar conditions: e.g. that it will again and again show undulatory behaviour in that conditions or corpuscular behaviour in the complementary experimental arrangement. Thus, what the principle allows us to say is: given such preparation of the system and such selection of an observable through coupling with an apparatus, we shall get any of the eigenstates of that observable with probability such and such. Thus, the results are the observable's eigenstates and the ground is the experimental set-up.

We may say that the above heuristic reformulation of the principle of sufficient reason could represent the retrograde foundation of inferences. The relevance of this foundation lies in this: due to the circle of inferences that has been sketched (and that can be applied to all forms of reasoning, even the mathematical or logical ones), the foundation of knowledge can never be an enterprise a priori so as to proceed only in

[147] Poincaré (1902, p. 151), Poincaré (1905, Chap. 10).
[148] Auletta (2011b, Sect. 5.4).

an anterograde way. Quite the opposite, we need to rationally justify our assumptions and, when we succeed in doing this, we have also

1. Retrogradely found a ground of bigger generality of that initial and necessary immature assumption, and
2. From that ground certain consequences can be anterogradely derived.

As recalled in the Introduction, something similar was affirmed by Einstein, when he said that physics is a theory in evolution "going in the direction of increasing simplicity of the logical bases"[149] (and not the other way around, what would be the business of deduction), a process that deals with the *justification* of the system of knowledge, provided by the proofs and generalisations of the theorems we formulate on the basis of the experience. I stress that this is in agreement with category theory (Sect. 8.1) and hopefully justifies the non-deductive approach that I have chosen for dealing with foundational issues.

8.3.3 Interpreted and Uninterpreted Ontology

In the following, I shall consider the categorial structure of QM. We are interested in particular into two different issues:

- On the one hand, we ask about the categorial relation between reversible (unitary) and irreversible dynamics.
- On the other, we like to set in the right categorial terms the relation between the basic theoretical entities of the theory and what we call reality.

I shall treat these two problems in such an order.

Interpreted Ontology

Let us start with an analysis of basic QM, i.e. of the non-relativistic theory without consideration of quantum fields (nor of gravity). This analysis will tell us something that is fundamental also for the other cases. Properties are customarily represented by sets while states and observables by classes in strict sense. Properties have an ontological substrate in the way in which systems interact or are related with other systems (a subcase of which is the way in which we measure or observe them). This is congruent with the notion of aspect of category theory (Sect. 8.1.2). Now, when we attribute a property in a classical mechanical context, we forget these relations and consider the property as intrinsic, i.e. as a characteristic that the system possesses in itself, and therefore we represent it by a variable that in fact turn out to be a set (of values). It is a simplification that is very helpful in many (especially classical) situations but it has no direct ontological substrate, and to assume the

[149] Einstein (1936, p. 96).

opposite determines the known troubles in quantum theory (Sect. 3.3.4). Of course, in Boolean algebra we have basic sets. Nevertheless, they also are derived from relations (e.g. $X = XY + X$), but, since in a logical context we have completeness (i.e. any set is related to any other set), such relations 'collapse' and all of these expressions are resolutions of identity (Sect. 8.2.3). Of course, this is not the case for any physical context in our (and any possible) world(s).

Thus, we can say that classes cannot be reduced to sets apart from a forgetful relation (Sect. 8.1.4).[150] This is clearly in agreement with the relational point of view supported here (Sects. 5.2.6 and 7.2.1).

QM suggests us to widen the notion of aspect and to apply it to the three basic classes involved in the theory. In fact, all basic aspects (state, observable, property) are equivalence classes of some kind of operation (Sects. 5.2 and 8.1.6) and therefore are types and present the relational aspect that is characteristic of aspects. Now, equivalence classes are rooted in the nature of information (Sects. 3.3.6, 3.4.2, and 6.3.4). Since in any information protocol the physical details of a system are irrelevant, we can extend the notion of equivalence to *operationally equivalent systems*.[151]

From this, it is clear that our theoretical categories are somehow at the cross of algorithms and operations. In other words, QM suggests us that a physical theory needs to be composed of formalism, ontological ascriptions and measurement theory

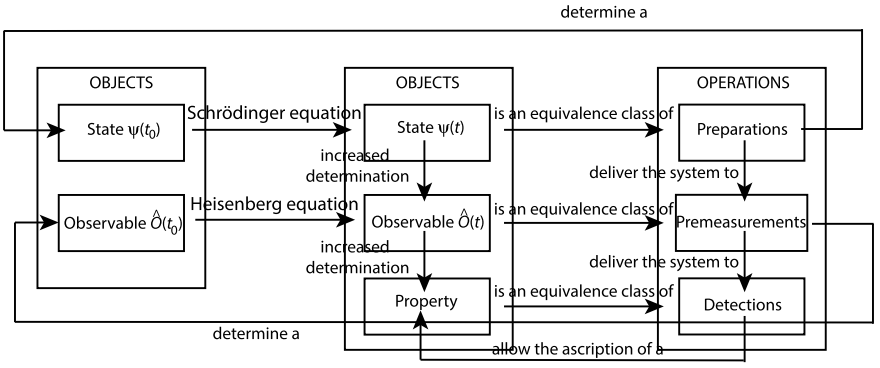


Fig. 8.27 Olog of basic quantum theory. QM as a theory is composed by the mathematical formalism (morphisms) and the theoretical entities (states, observables and properties). These morphisms (Schrödinger and Heisenberg equations) constitute a group and rule the evolution of states and observables. Experimental operations are concrete interactions with the physical world guided by the theory. Preparation, premeasurement and detection can be understood as three successive degrees of dynamical determination of the involved physical system. Also, these three steps are related by morphisms. States, observables, and properties are equivalence classes of preparations, premeasurements and detections, respectively

[150] A first hint in Peirce (1898, p. 198).

[151] D'Ariano et al. (2017, Sects. 3.1, 3.2 and 4.2).

that defines the kind of operations that are necessary for acquiring information from the system.[152]

Let us have a look at Fig. 8.27 as a kind of summary of what has been said so far. The scheme is made up by the theoretical entities understood as aspects of systems or objects in a wide sense.[153] It represents the way we refer to reality in a non-mirroring way by making use of theoretical elements defined by the theory and within the mathematical formalism (represented by the algorithms, i.e. the morphisms). In other words, being state, observable and property equivalence classes, they do not mirror any specific individual reality, although are still referred to their ontological substrate represented by experimental operations (other morphisms, not necessarily reversible) that we perform on the system. These three notions constitute the *interpreted ontology*. This theory guides us and thus also describes the concrete operations that we perform on the external reality (the three steps of the measurement process). It is clear that also (experimental) operations can be understood as morphisms. However, we are here interested in the equivalence classes that they give rise to and in this sense I treat them as objects for simplicity: category theory allows indeed for building classes of morphisms.

This theory describes well the global dynamics of quantum systems. However, as it is clear from the above scheme, properties are not ruled by a morphism and remain so as a kind unjustifiable addition to the theory. In order to account for the global–local dynamics described by amplitude and effects operators (Sects. 5.2.4 and 5.2.5), we need to introduce the generalised dynamics displayed in Fig. 8.28: a propagation is followed by a jump (event) and this again by a propagation, and so on. Thus, taking the Kraus equations (5.153) for states and observables as models (replacing Schrödinger and Heisenberg equations), we deal with propagators for states (as the ordinary Green's functions) and relative propagators for observables. In fact, the formalism of path integrals or quantum trajectories (Sect. 6.2.4) is capable to incorporate all the three kinds of morphisms. Since such a dynamics is based on

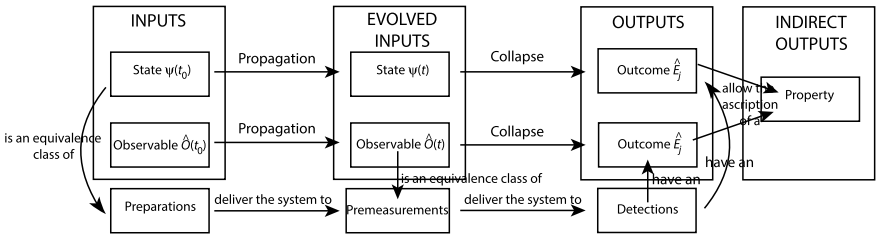


Fig. 8.28 Olog of the generalised quantum theory. Such generalised morphisms constitute a monoid. Note that the reversible morphisms are particular cases of the two Kraus equations. Thus, the monoid contains the previous group. Note also that the previous equivalence relations, as displayed in Fig. 8.27, remain the same (but they are no longer explicitly indicated here) This generalised theory describes the global–local dynamics

[152] As analysed in Ludwig (1978, pp. 8–9).

[153] Auletta and Torcal (2011).

propagators, it is also suitable for describing, e.g. the propagation of the electric field or the quantum gravity dynamics (Sects. 7.1.2, 7.1.3, and 7.2.2). As said, the sequence propagation–event–propagation can well represent the diffusion of a signal through (subsequent) event(s), in agreement with the local–causal structure of relativity.

All requirements of a category are satisfied (Sect. 8.1.3). Note in particular the requirement of composition (8.24). If we denote by P the functor leading from active to passive transformation and by A the functor leading from passive to active transformation (being one functor the inverse of the other), in the case of the olog 8.27 we have

$$\begin{array}{ccc} \psi(t_0) & \xrightarrow{\text{S.eq.}} & \psi(t) \\ {\scriptstyle P,A}\updownarrow & & \updownarrow{\scriptstyle P,A} \\ \hat{O}(t_0) & \xrightarrow{\text{H.eq.}} & \hat{O}(t) \end{array} \tag{8.144}$$

Also the identity morphism does exist, as far as, e.g. energy eigenfunctions (for the Schrödinger picture) or Hamiltonians (for the Heisenberg picture) do not evolve with time. We have a similar scheme with the olog Fig. 8.28:

$$\begin{array}{ccccc} \psi(t_0) & \xrightarrow{\text{Propagation}} & \psi(t) & \xrightarrow{\text{Collapse}} & \hat{E}_j \\ {\scriptstyle K_1,K_2}\updownarrow & & \updownarrow{\scriptstyle K_1,K_2} & & \updownarrow{\scriptstyle \text{identity}} \\ \hat{O}(t_0) & \xrightarrow{\text{Propagation}} & \hat{O}(t) & \xrightarrow{\text{Collapse}} & \hat{E}_j \end{array} \tag{8.145}$$

With the functors K_1, K_2, I understand the transformations between the two Kraus pictures (5.153). Moreover, we pass from one kind of olog to another thanks to natural transformations (Sect. 8.1.5). Essentially, these natural transformations, α, α', are based on the condition (5.134): when we get equality sign, we recover the transformations of the first olog from those of the second olog. Thus,

$$\begin{array}{ccc} A & \longrightarrow & K_1 \\ {\scriptstyle \text{olog 1}}\updownarrow & & \updownarrow{\scriptstyle \text{olog 2}} \\ P & \xrightarrow{\ '\ } & K_2 \end{array} \tag{8.146}$$

Uninterpreted Ontology

As we have seen, we can have in QM maximal knowledge of the whole (being in a pure state) with minimal knowledge of the parts or subsystems (whose reduced density matrices are mixtures). This can be called the quantum-mechanical *holism*. This contradicts local (classical) realism as defined in Sect. 3.3.4. Nevertheless, such a holism cannot be brought up to the point that we can have access to the global state of the system only if we let the involved subsystems interact in a joint measurement (like

Bell measurement (Sects. 3.3.6 and 6.4.4))[154]: in fact, this would ultimately imply that we should computationally run transformations with the input system entangled with any possible additional system of the whole universe. At the opposite, QM ensures us that the information extracted from two entangled subsystems considered separately is sufficient to reconstruct the whole information about the compound entangled system. This shows that quantum physics is also local and therefore reductionist (Sects. 6.4.3 and 6.4.4). This is why the dynamics of quantum systems is intrinsically an entrenchment of global and local aspects and why it is a step induced by the theory itself to describe the dynamics of the whole when that of the parts is known or even less known: quantum theory is able to provide such a maximally informative description of reality.[155] In other words, the inference regarding the environment and even the whole universe are not only justified but even *necessary* considered the structure of quantum theory. Thus, any widening thanks to purification (Sect. 6.4.6) naturally leads to consider *non-interpreted ontology*, or the ontology of quantum systems beyond strict controlled situations. Without this, the "notion of the universe as a whole" becomes "an embarrassing concept".[156] I also recall Einstein's word about the impossibility to ground physics on observable quantities only (Sects. 3.4.3 and 5.1.2).

There is however an important difference between interpreted and uninterpreted ontologies that was also the main concern of Bohr. In the context of the *interpreted ontology* and of the quantum theory we acquire information since we control the experimental operations on quantum systems. Things are not so for the 'external' world. We no longer control these dynamical processes and cannot follow what happens.

Role of Uninterpreted Ontology

If so, why we do not follow the Copenhagen interpretation and limit ourselves to what we know in controlled situations? The dissatisfaction with this interpretation derives from the fact that it is unable to deal with two fundamental gaps in our *theoretical* description of reality:

1. We can make experiments with single quantum systems but we do not know how they behave before letting them pass our procedures, and
2. We are able to describe many fundamental aspects of the world but, as stressed in the previous chapter, quantum theory needs to be integrated with something else to this purpose, as it happens for the standard model of particles, atomic and molecular physics, chemistry. Thus, QM appears as an abstract or even esoteric construction that is able to describe the 'external' world only with the help of additional theories that provide that construction of concrete, physical elements.

[154] D'Ariano et al. (2017, Sect. 6.1).

[155] D'Ariano et al. (2017, Chap. 5).

[156] I quote here Bell (1981).

So, they are sorts of input and output gaps. Of course, we expect from QM that it provides a description of those physical realities that represent the experimental contexts in which we perform our controlled operations. However, it is like moving in circle: quantum physics is able to explain the *arising* of the atomic–molecular structure of the macroscopic world but with some additional knowledge and this macroscopic world is in turn *presupposed* to be a condition (e.g. the experimental set-up) for dealing with quantum systems that are otherwise unknown. Then, the uninterpreted ontology that we need should be able to make *quantum-mechanically* reasonable both the experimental contexts and the results at our disposal: it should consist of those necessary conditions, whose existence is required for obtaining the experimental outcomes that follow the whole sequence of operations (preparation, premeasurement and detection). Thus, an uninterpreted ontology, thought of as a kind of dynamical entrenchment between features and events (Sect. 5.1.4), needs to be inferred from what the theory (theoretical entities plus algorithms) tells us according to the available experimental data, under the (Bayesian) hypothesis that such data are the consequences of both the dynamics of the external reality and the specific operations we perform. From this perspective, such an uninterpreted ontology as a kind of elusive reality that we cannot fully grasp or experience could be what makes our interpreted ontology (and our theory) meaningful.

Dealing with Information

Since the spontaneous operations occurring in Nature are uncontrolled, we need to cast our own controlled operations in a more general way that makes this analogy viable at all. The best way to do this is to cast these controlled operations in their most general form, namely as operations dealing with information: processing, sharing and acquiring (Sect. 6.3.4). Thus, information can bridge the gap between interpreted and uninterpreted ontology, between observation and reality or mind and matter. When we see that quantum systems give rise spontaneously to those determination processes that we get in our labs in a controlled way (by building, e.g. atoms and molecules: Sect. 7.3), we are justified in assuming that, analogously to our *controlled operations*, also in Nature there are *uncontrolled operations* that lead to those results. *Analogies* in fact cannot be established between concepts and their referents (i.e. possible entities) but can do between different concepts or classes (here, of operations or morphisms). The first option is an illusion against which already Einstein warned us.

Let us first consider the case that we know something about the world without control of the process apart from the final set-up that we choose: since we cannot control the details of such processes occurring spontaneously, what we can do is to look at the *results* or effects of some processes. Thus, at least a signal need to be received in order to say anything about what process occurred. In other words, we can judge about external reality when we deal with specific effects upon reception of some kind of signal. Thus, we cannot control what happens at the source (of that signal) apart from what the latter tells us. Nevertheless, in such conditions our

reception of the signal corresponds to experimentally controlled situations, as the delayed-choice experiment shows (Sect. 2.4.3).

Another issue is what we can say about systems that do not interact with us (and from which no signal is actually received). In such a case, although we are not able to acquire information from systems that we do not control or with which we do not, at least indirectly, interact, we can assume that such information is *acquirable*, or at least should be in principle receivable. In other words, we need to be able to say something also about such a situation.

Crisis of Representation

I have mentioned that all quantum systems of the world contain an amount of infinite information in themselves (Sect. 3.2.3). So, when they interact, any 'semantic' or referential aspect is totally irrelevant: such an information has only a syntactical or combinatorial character, and every system can be said to only 'represent' itself. For instance, they simply come to pair intrinsic codes when they become entangled (Sects. 3.2.2, 3.3.7, and 6.3.4). In other words, the only thread that they have with each other are the operations that they mutually perform on each other. And therefore, it is reasonably to assume that this is the basis of self-similarity. At least, everything that happens in our universe needs to be explained by satisfying the requirements imposed by physics, and in particular by QM. Therefore, I am suggesting that the main lesson of QM (in full agreement with the nature of symbolic systems: see Sect. 8.3.2) is to focus not on the way in which we view reality (and represent it) but on the way in which we *operate* on (and in) the world and the world *impacts* on us. I have already used this approach when I have shown that quantum systems are sensitive to the environment (Sects. 5.2.6, 6.1.1, and 7.2.1).

The schism of physics, mentioned in the Introduction, and the loss of the notion of reality are in fact the failure of the idea that our theories are *representations* of reality.[157] The gap that is opened between interpreted and uninterpreted ontology is rooted in the difference between our representations and our operations. As a matter of fact, the four great minds that I have taken as our Virgils (and also many other great minds) could not agree on a common *representation* of quantum reality (Sect. 5.3). They could not even fully appreciate their commonalities, as I believe to have shown in the previous chapters. Heisenberg went very near to the awareness that a correct operationalism is the solution, but, apart from his juvenile standpoints, he perhaps never fully discussed the possible nexus between operations in the world and operations as guided by a theory, that is, between ontological and epistemic aspects.

[157] Only in this specific sense I could agree with the dictum recalled in the Introduction, when the Nobel Prize winner Feynman says that QM cannot be explained or understood (Feynman et al. 1965, III, 1–1). Otherwise it would imply a renouncement to scientific research.

Representations and Operations

Why is there such a failure of representation? Because, as well understood by the subjectivists or idealists, any representation is the result of mental or neural processes that are *intrinsic* to each individual (and even to specimens of other species) and therefore cannot be shared as such (Sect. 8.3.1).[158] All organisms map external states into internal states in ways that, through the ladder of evolution, become more and more various and therefore individualised: they are kinds of "private" results of these mapping operations.

Thus, the problem with representations is that we never know what we share with others for certain, neither what we share with the external world, if not in the subsequent intercourse with the world and conspecifics. This is clear: representations (also in humans) are more sophisticated forms of perception (where different kinds of percepts are integrated) that serve the cause of the survival of the *single organism* (Sect. 8.3.1).[159] Although for each species, there are significant common traits (and limitations) in both perceiving and behaving, such representations are finally the result of the individual biography and are not hardwired, because they need to be plastically fine-tuned if they will be helpful at all. Our mind cannot change this biological fact (or constraint) on which our life depends. This is why our postmodern culture, after the schism of physics, has finally produced the idea that everything is *simulacrum*: representations that finally do not represent anything else if not themselves and play a role in the mass media spectacle.[160] This view stems ultimately from the traditional understanding by D. Hume of the mind as a screen on which different representations follows like on a theatre's scene.[161] In such a way, we would have an open succession of unrelated 'images', but would be unable to understand that, e.g. one is right and another wrong of the *same* object, since there is no other way to have access to the latter if not through another representation.[162]

Happily, as QM itself suggests (Sect. 2.2), the mind is not only a screen but is dynamically involved in the interaction with the world. And, different from representations, the operations themselves appear to be structurally similar across all levels of reality allowing us a tight thread with reality.[163] In the language of Peirce,[164] representations are icons, that is, structures that have been *endogenously generated* while the thread with reality is ensured by our ability to be factually and *operationally connected* with external world (especially through our motor system in the case of

[158]This is the so-called problem of the translation among different representations as pointed out in Quine (1960, Sect. 12), Quine (1969).

[159]As stressed in Auletta (2011a, Chap. 12), where further references can also be found.

[160]The philosophy of the postmodern world dominated by simulacra has been introduced in Baudrillard (1981).

[161]Hume (1739, p. 253).

[162]As pointed out in Auletta (2011a, Sect. 6.1).

[163]This has been the basic insight of Auletta (2011a).

[164]Peirce (1903b, 1907), Peirce CP (2019, 2.228, 2.247–8, 2.304), Peirce (1903c, 1.540). On this subject see also Auletta (2011a, Chaps. 8 and 19). Auletta, Gennaro.

animals, but in general through the actions that any organism undertakes towards external reality).[165] I recall that this aspect is called *indexical* by Peirce (see, e.g. Fig. 8.23, Sect. 8.3.1, and comments).

In order to better understand the meaning of these statements, we need to go over to evolutionary considerations. As a matter of fact, lower species show a perception of our world that is less complex than higher species. This contradicts the hypothesis that our perceptions are forged by templates on the external reality (Sects. 6.2.5 and 8.3.1), and show that 'objects' are rather the result of sophisticated construction processes.[166] For instance, at the start of the visual process, we only have single excitations of receptors. As the information is broadcast to higher areas, it will be elaborated and modulated also thanks to additional inputs coming top-down from high-level areas. In such a process, by integrating different sensorial inputs even coming from different sensory channels, we finally get the 'image' of an object. What is the biological significance of this? It means that objects' perception does matter, but does matter only *for advanced organisms that can interact with them in an appropriate manner*, and have thus coevolved with a niche-like environment in which such sophisticated perception is important for survival. However, this is a very late result of evolution. In other words, as anticipated, our sophisticated perception of the world (and the resulting representations) are *functional to the operations for controlling* as much as possible the external environment that are crucial for the survival of any species.[167] The basic components of these representations and their general schemes, especially regarding their species-specific survival value (i.e. the affordance: Sect. 8.3.2), have been selected in the course of evolution because they serve quite well this purpose and not because they are mirrors of an external reality.[168] In fact, the only result of natural selection is that they became associated to external objects and processes through the *operations* that the organism performs on that processes and objects.

I recall that it is only for humans that the issue of the faithfulness of our models of reality arises. And for this reason, we cannot affirm that representations of the world are not important or even that should be dismissed. Our statement that a controlling system needs models and therefore also representation of the controlled system goes into this direction (Sect. 8.3.1). Nevertheless, they remain provisional and approximate tools. Therefore, when we will judge the value of a given representation, we need finally to resort to the basic operations that have *produced* it or to the operations for which it is *functional*. This is the only way to correct and fine-tune our own representations (Sect. 8.3.2) and to be able to converge on a common and growing understanding of reality (Sect. 8.1.4).

Thus, when we deal with the external world, and especially with the information that we could acquire, we try to establish rational equivalence classes of objects

[165] This is the essence of motor cognition (Jeannerod 1988, 2006, 2009).

[166] Examination and literature in Auletta (2011a, Chaps. 4 and 13).

[167] It is not by chance that Poincaré said that organisms immobile (like plants) could not build a geometry, because they do not move and act in the space (Poincaré 1905, pp. 68–69).

[168] von Helmholtz (1867, pp. 586–589), Herbart (1824, Sect. 3).

and events. What we try to do is to make external realities adequate to our rational criterions (Sects. 7.3.3 and 8.3.2). The whole scientific enterprise is devoted to this and explains while it is intimately connected with technology, since this adequacy, being rooted in operations, involves also crucial pragmatic aspects. The world displays not only emergence of different levels of complexity (and this process has not yet terminated, as far as we can understand), which demand specific theories and methodologies, but also shows an 'irrational' element that is reproduced at all levels of complexity and resists to our rational attempts. This irrational element is crucial since it is the source of variety and therefore of any novelty in our universe. It is what justifies an uninterpreted ontology. Nevertheless, the work of reason is not ineffective. Quite the opposite: such an irrational element is the drive for performing better and better, for correcting more and more our inadequate explanations. As J. Wheeler remarks, "Life and mind: For how much can they be conceived to count in the scheme of existence? Nothing, say the billions of light years of space that lie around us. Everything, say the billions of years of time that lie ahead of us. It cannot matter that man in time to come will have been supplanted by, or will have evolved into, intelligent life of quite other forms. What counts […] is the rate of asking questions and obtaining answers by elementary quantum phenomena, acts of observer–participancy, exchanges of information."[169]

8.4 Summary

- Category theory represents the highest level of generalisation attained by formal sciences (mathematics, logic, information theory). It is centred on the notion of class, of which sets are a particular case.
- Categories have aspects and morphisms, which fulfil the composition requirement as well as associativity and identity laws.
- Monoids and groups are among the most important categories for physics.
- When categories have more objects, these are related through functors, as well as objects of different categories, while natural transformations connect different functors.
- Universality is expressed here through the notions of limit (and colimit), adjoint functors (that allow the treatment of both reduction and emergence), and representation (which goes together with the notion of aspect), allowing a relational view of aspects that is simultaneously objective (congruence or convergence of the different "views").
- Category theory when applied to QM allows the treatment of informational protocols (like teleportation) in logical terms.
- Boolean algebra can be treated in pure combinatorial terms thanks to logical IDs.
- Any nD Boolean algebra can be mapped to a vector space, which allows a connection between logic and structures of QM.

[169] Wheeler (1988).

- A careful analysis of logic shows that the only specificity of QM from a logical point of view is its probabilistic structure while the non-commutativity is logically irrelevant.
- Our operations in laboratories are not fundamentally different relative to the operations performed by all organisms. In all these cases, the notion of control is crucial.
- Our inferences (deduction, abduction, induction) are connected with these operations but are framed in symbolic systems that allow the treatment of aspects that are hidden to ordinary perception.
- In the whole process of inference, associative moments play an important role. The whole takes the form of Bayesian inferences.
- The progress of science goes through cycles of deduction–abduction–induction such that we move towards more and more general laws and more and more specific objects. These two processes can be considered in terms of right and left adjunctions.
- Interpreted ontology is integrated in the theory of basic QM, whose morphisms are reversible and thus constitute a group.
- This can be generalised to not-necessarily unitary morphisms that constitute a monoid.
- Crucial is also the notion of uninterpreted ontology.
- The crisis of representation does not cut our thread with reality but we need to ensure it in terms of operations.

References

Abramsky, S., and N. Tzevelekos. 2011. Introduction to Categories and Categorial logic. In [Coecke 2011, 3–94].

Abramsly, S., and B. Coecke. 2009. Categorial Quantum Mechanics. http://www.indiana.edu/~iulg/qliqc/HQL_fin.pdf

Adleman, Leonard M. 1994. Molecular Computation of Solutions to Combinatorial Problems. *Science* 266: 1021–1024. https://doi.org/10.1126/science.7973651.

Adleman, Leonard M. 1998. Computing with DNA. *Scientific American* 279 (2): 54–61.

Alberts, B., D. Bray, J. Lewis, M. Raff, K. Roberts, and J. D. Watson. 1983. *The Molecular Biology of the Cell*. New York: Garland P; 2nd ed. 1989; 3rd ed. 1994; Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P., 4th ed. 2002; 5th ed. 2008.

Aristotle. 1831. Opera, ed. Immanuel Bekker. Berlin: G. Reimerum.

Aristotle. 1949. Categorae, in [Aristotle Cat-Int.]; also in [Aristotle Op., 1–15].

Aristotle. 1964. *Analytica Priora et Posteriora*, ed. W.D. Ross and L. Minio-Paluello. Oxford.

Aristotle. 2019a. *Analytica Posteriora*. In [Aristotle AnPr–Po.]; also in [Aristotle Op., 71–100].

Aristotle. 2019b. *Analytica priora*, in [Aristotle AnPr–Po.]; also in [Aristotle Op., 24–70].

Aristotle. 2019c. *Categoriæ*. In [Aristotle Cat-Int.]; also in [Aristotle Op., 1–15].

Armstrong, David M. 1983. *What is a Law Of Nature*. Cambridge: Cambridge University Press.

Auletta, Gennaro. 2000. Foundations and Interpretation of Quantum Mechanics. In *The Light of a Critical-historical Analysis of the Problems and of a Synthesis of the Results*. Singapore: World Scientific; rev. ed. 2001.

Auletta, Gennaro. 2008a. Autarchy and Openness in Living Systems. In [Auletta 2008b, 17–25].
Auletta, Gennaro. (ed.). 2008b. *The Controversial Relations Between Science and Philosophy: New trtunities for a Fruitful Dialogue*. Libreria Editrice Vaticana: Vatican City.
Auletta, Gennaro. 2009. What About the Three Forms of Inference? *Acta Philosophica* 18 (1): 59–74.
Auletta, Gennaro. 2011a. *Cognitive Biology: Dealing with Information From Bacteria to Minds*. Oxford: University Press.
Auletta, Gennaro. 2011b. (in collaboration with I. Colagè, P. D'ambrosio, and L. Torcal), *Integrated Cognitive Strategies in a Changing World*. Rome: G and B Press.
Auletta, Gennaro. 2011c. Teleonomy: The Feedback Circuit Involving Information and Thermodynamic Processes. *Journal of Modern Physics* 2 (3): 136–145. https://doi.org/10.4236/jmp.2011.23021. http://www.scirp.org/journal/PaperInformation.aspx?paperID=4488.
Auletta, Gennaro. 2012. Causation Upside Down? *Revista Portuguesa de Filosofia* 68 (1-2): 9–32.
Auletta, Gennaro. 2013a. Information and Metabolism in Bacterial Chemotaxis. *Entropy* 15: 311–326. https://doi.org/10.3390/e15010311. http://www.mdpi.com/1099-4300/15/1/311.
Auletta, Gennaro. 2013b. *Mechanical Logic in Three-Dimensional Space*. Singapore: Pan Stanford Pub.
Auletta, Gennaro. 2013c. A New Way to Implement Quantum Computation. *Journal of Quantum Information Science* 3 (4): 127–137. https://doi.org/10.4236/jqis.2013.34017.
Auletta, Gennaro. 2013d. Aristotle's Syllogistic Revisited. *Epistemologia* 36: 233–261.
Auletta, Gennaro. 2015a. Emergence: Selection, Allowed Operations, and Conserved Quantities. *South-African Journal of Philosophy* 34: 93–105. https://doi.org/10.1080/02580136.2014.992156.
Auletta, Gennaro. 2015b. Les premiers humains modernes. Rapprochemententre la génétique, l'épigénétique et lessciences humaines. Le tornant cognitif–culturel qui a rendu lessystémes symboliques possibles. In *Sur le chemin de l'humanité*, ed. Henry de Lumley, 387–439. Paris, Pontifical Academyof Sciences and CNRS Ed.; Discussion: 440–444.
Auletta, Gennaro. 2015c. Irreducible Statements and Bases for Finite–Dimensional Logical Spaces. *Indian Ocean Review of Science and Technology* 2015, Article 2.
Auletta, Gennaro. 2016a. Networks and Causation Top-Down. *Revista Portuguesa de Filosofia* 72: 171–180. https://doi.org/10.17990/RPF/2016-72-1-0171.
Auletta, Gennaro. 2016b. From Peirce's Semiotics to Information-Sign-Symbol. *Biosemiotics* 2016: 1–16. https://doi.org/10.1007/s12304-016-9275-2.
Auletta, Gennaro. 2016c. Connection Between Scripts Embedding Motor Schemes and Decision Making. In *Processes of Believing: The Acquisition, Maintenance, and Change in Creditions*, ed. H.-F. Angel, L. Oviedo, and R. Seitz. Berlin: Springer.
Auletta, Gennaro. 2016d. Critical Examination of Peirce's Theory of Categories. *SCIO: Journal of Philosophy* 12: 23–49.
Auletta, Gennaro. 2017. A Critical Examination of Peirce's Theory of Natural Inferences. *Revista Portuguesa de Filosofia* 73 (3–4): 1053–1094. https://doi.org/10.17990/RPF/2017-73-3-1053.
Auletta. G., and M. Jeannerod. 2013. Introduction: Consciousness as a Top-Down Causal Agency. In [Auletta et al. 2013, 1–49].
Auletta, G., and L. Torcal. 2011. From Wave-Particle to Features-Event Complementarity. *International Journal of Theoretical Physics* 50: 3654–3668. https://doi.org/10.1007/s10773-011-0833-8.
Auletta, G., G. Ellis, and L. Jaeger. 2008. Top-Down Causation by Information Control: From a Philosophical Problem to a Scientific Research Program. *Journal of the Royal Society: Interface* 5: 1159–1172. https://doi.org/10.1098/rsif.2008.0018.
Auletta, G., I. Colagè, and P. D'Ambrosio. 2013. The Game of Life Implies Both Teleonomy and Teleology. In [Auletta/Pons 2013, 267–284].
Baez, J., and M. Stay. 2011. Physics, Topology, Logic and Computation: A Rosetta Stone. In [Coecke 2011, 95–172].

Barbieri, Marcello. 2003. *The Organic Codes: An Introduction to Semantic Biology*. Cambridge: University Press.
Barbieri, Marcello. 2015. *Code Biology: A New Science of Life*. Berlin: Springer.
Baudrillard, Jean. 1981. *Simulacres et simulation*. Paris.
Bell, John S. 1981. Quantum Mechanics for Cosmologists. In *Quantum Gravity 2*, ed. C.J. Isham, R. Penrose, and D. Sciama. Clarendon, Oxford; rep. in [Bell 1987, 117–138].
Bekenstein, Jacob D. 1987. *Speakable and Unspeakable in Quantum Mechanics*. Cambridge: University Press.
Beltrametti, E., and G. Cassinelli. 1981. *The Logic of Quantum Mechanics*. Redwood City: Addison-Wesley.
Berg, H.C., and D.A. Brown. 1972. Chemotaxis in Escherichia Coli Analysed by Three-dimensional Tracking. *Nature* 239: 500–504.
Birkhoff, G.D., and J. von Neumann. 1936. The Logic of Quantum Mechanics. *Annals of Mathematics* 37: 823–43; rep. In [Hooker 1975, 1–26].
Boole, George. 1854. *An Investigation of the Laws of Thought, on Which Are Founded the Mathematical Theories of Logic and Probabilities*. London: Walton and Maberly. New York: Dover.
Bourret, R.B., and A.M. Stock. 2002. Molecular Information Processing: Lessons From Bacterial Chemotaxis. *The Journal of Biological Chemistry* 277: 9625–9628.
Busch, P., M. Grabowski, and P.J. Lahti. 1995. *Operational Quantum Physics*. Berlin: Springer.
Byron Jr., F.W., and R.W. Fuller. 1969. *Mathematics of Classical and Quantum Physics*. New York: Dover.
Carnap, Rudolf. 1928. *Der logische Aufbau der Welt*. Meiner: Hamburg.
Chaitin, Gregory J. 1998. *The Limits of Mathematics: A Course on Information Theory and Limits of Formal Reasoning*, 2nd ed. Singapore: Springer.
Chaitin, G.J., N. da Costa, and F.A. Doria. 2011. *Gödel's Way: Exploits into an Undecidable World*. Milton Park: Taylor and Francis.
Chellas, Brian F. 1980. *Modal Logic. An Introduction*. Cambridge: University Press.
Chomsky, Noam. 2000. *New Horizons in the Study of Language and Mind*. Cambridge: University Press.
Clark, Andy, and Being There. 1997. *Putting Brain, Body, and World Together Again*. Cambridge, MA: MIT Press.
Coecke, Bob (ed.). 2011. *New Structures for Physics*. Berlin: Springer.
Coecke, B., and É.O. Paquette. 2011. Categories for the Practising Physicist. In [Coecke 2011, 173–286].
Conant, R.C., and R.W. Ashby. 1970. Every Good Regulator of a System Must Be a Model of That System. *International Journal of Systems Science* 1: 89–97.
Cosmides, L., and J. Tooby. 2000. Consider the Source: The Evolution of Adaptations for Decoupling and Metarepresentation. In [Sperber 2000, 53–115].
Cosmides, L., and J. Tooby. 2002. Unraveling the Enigma of Human Intelligence: Evolutionary Psychology and the Multimodular Mind. In [Sternberg and Kaufman 2002, 145–198].
Crile, George. 1941. *Intelligence, Power and Personality*. New York: McGraw-Hill.
D'Ariano, M., G. Chiribella, and P. Perinotti. 2017. *Quantum Theory From First Principles: An Informational Approach*. Cambridge: University Press.
Deacon, Terrence W. 1997. *The Symbolic Species: The Co-evolution of Language and the Brain*. New York: W. W. Norton.
Dehaene, Stanislas. 2014. *Consciousness and the Brain: Deciphering How the Brain Codes Our Thoughts*. New York: Viking.
Deutsch, David. 1997. *The Fabric of Reality: The Science of Parallel Universes and Its Implications*. London: Penguin Books.
Deutsch, David. 2011. *The Beginning of Infinity: Explanations that Transform the World*. London: Penguin Books.
Eddington. 1939. *Philosophy of Physical Science*. Cambridge: University Press.
Einstein, Albert. 1930. Zur Methodik der Theoretischen Physik, rep. in [Einstein 1993, 113–119].

Einstein, Albert. 1934. On the Method of Theoretical Physics. *Philosophy of Science* 1: 163–169; eng. tr. of [Einstein 1930].
Einstein, Albert. 1936. Physik und Realität. *Journal of Franklin Institute* 221: 349–382; eng. tr. in [Einstein 1956, 59–97].
Einstein, Albert. 1956. Out of My Later Years, Estate of A. Einstein: New York: Wings Books.
Epperson, M., and E. Zafiris. 2013. *Foundations of Relational Realism: A Topological Approach to Quantum Mechanics and the Philosophy of Nature*. Lexington Books.
Feynman, R., R.B. Leighton, and M. Sands. 1965. *Lectures on Physics*. Reading (Massachusetts): Addison-Wesley, I v.: 1963, 1977; II v.: 1964, 1977; III v.: 1965, 1966.
Fogassi, L., P.F. Ferrari, B. Gesierich, S., Rozzi, F., Chersi., and G. Rizzolatti. 2005. Parietal Lobe: From Action Organization to Intention Understanding *Science* 308: 662–667.
Fraenkel, A., Y. Bar-Hillel, and A. Levy. 1958. *Foundations of Set Theory*. Amsterdam: North-Holland.
Frege, Gottlob. 1884. *Die Grundlagen der Arithmetik: eine logisch-mathematische Untersuchung über den Begriff der Zahl*, Breslau, w. Koebner, 1884.
Frege, Gottlob. 1893. *Grundgesetze der Arithmetik*, Jena, H. Pohle; I v.; II v.: 1903; reprint: Hildesheim, Olms, 1998.
Friedman, Michael. 1983. *Foundations of Space–time Theories. Relativistic Physics and Philosophy of Science*. Princeton: University Press.
Friston, Karl J. 2005. A Theory of Cortical Responses. *Philosophical Transactions of the Royal Society, London B* 360: 815–836.
Friston, Karl J. 2012. A Free Energy Principle for Biological Systems. *Entropy* 14: 2100–2121.
Friston, Karl J. 2013. Life as We Know It. *Journal of the Royal Society Interface* 10: 20130475. https://doi.org/10.1098/rsif.2013.0475.
Friston, K.J., and S. Kiebel. 2009. Predictive Coding Under the Free-energy Principle. *Philosophical Transactions of the Royal Society* B364: 1211–1221.
Friston, K.J., and K.E. Stephan. 2007. Free-Energy and the Brain. *Synthese* 159: 417–458.
Friston, K.J., J. Kilner, and L. Harrison. 2006. A Free Energy Principle for the Brain. *Journal of Physiology (Paris)* 100: 70–87.
Friston, K.J., B. Sengupta, and G. Auletta. 2014. Cognitive Dynamics: From Attractors to Active Inference. *Proceedings of the IEEE* 102: 427–445. https://doi.org/10.1109/JPROC.2014.2306251.
Geroch, Robert. 1985. *Mathematical Physics*. Chicago: University of Chicago Press.
Gibson, James J. 1979. *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin, 1979; Hillsdale. NJ: L. Erlbaum Ass, 1986.
Gilbert, Scott F. 2006. *Developmental Biology*, 8th ed. Sunderland, MA: Sinauer.
Givant, S., and P. Halmos. 2009. *Introduction to Boolean Algebras*. Berlin: Springer.
Gödel, Kurt. 1931. Über formal unentscheidbar Sätze der. Principia Mathematica und verwandter Systeme. *I, Monatshefte für Mathematik und Physik* 38: 173–198.
Hauser, Marc D. 1996. *The Evolution of Communication*. Cambridge, MA: MIT Press.
Hauser, Marc D. 2009. The Possibility of Impossible Cultures. *Nature* 460: 190–196.
Herbart, Johannes F. 1824. *Psychologie als Wissenschaft*; rep. in [Herbart 1850, v. 5, 1–189].
Hilbert, David. 1903. *Grundlagen der Geometrie*. Leipzig: Teubner.
Hooker, Cliff A. (ed.). 1975. *The Logico-Algebraic Approach to Quantum Mechanics*. Reidel: Historical Evolution. Dordrecht.
Hume, David. 1739. *A Treatise of Human Nature*. London, Oxford: University Press.
Hume, David. 1748. *Enquiry Concerning Human Understanding*; 3rd ed. 1754; London.
Jaynes, Edwin T. 1967. Foundations of Probability Theory and Statistical Mechanics. In *The Delaware Seminar in the Foundations of Physics*, ed. M. Bunge, 77–101. New York: Springer.
Jeannerod, Marc. 1988. The Neuronal and Behavioral Organization of Goal-directed Movements. Oxford: University Press.
Jeannerod, Marc. 2006. *Motor Cognition: What Actions Tell the Self*. Oxford: Oxford University Press.

Jeannerod, Marc. 2009. *Le cerveau volontaire*. Paris: O. Jacob.
Jurica, M.S., and B.L. Stoddard. 1998. Mind your B's and R's: Bacterial Chemotaxis, Signal Transduction and Protein Recognition. *Structure* 6: 809–813.
Kandel, Eric R. 2006. *In Search of Memory: The Emergence of a New Science of Mind*. New York: W. W. Norton.
Kuhn, Thomas S. 1957. *The Copernican Revolution: Planetary Astronomy in the Development of Western Thought*. Cambridge, MA: Harvard University Press.
Kuhn, Thomas S. 1962. *The Structure of Scientific Revolutions*. Chicago: University Press; 2nd ed. 1970.
Lakatos, Imre. 1976. *Proofs and Refutations: The Logic of Mathematical Discovery*. Cambridge: University Press.
Landsman, Klaas. 2017. *Foundations of Quantum Theory: From Classical Concepts to Operator Algebras*. Berlin: Springer.
Laughlin, Robert B. 2005. *A Different Universe (Reinventing Physics From the Bottom Down)*. New York: Basic Books.
LeDoux, Joseph E. 2002. *Synaptic Self: How Our Brains Become Who We Are, Viking*. London: Penguin.
Leibniz, Gottfried W. 1666. *Dissertatio de Arte Combinatoria*. Leipzig; rep. In [Leibniz PS, IV, 27–104].
Leinster, Tom. 2014. *Basic Category Theory*. Cambridge: University Press.
Lemke, Jay L. 2000. Opening Up Closure: Semiotics Across Scales. *Annals of the New York Academy of Sciences* 901: 100–111.
Ludwig, Günther. 1978. *Die Grundstrukturen einer physikalischen Theorie*. Berlin: Springer; 2nd ed, 1990.
Mach, Ernst. 1905. *Erkenntnis und Irrtum*, 2nd ed. 1906, 5th ed. Leipzig, 1926, Darmstadt, WBG.
Margenau, Henry. 1950. *The Nature of Physical Reality: A Philosophy of Modern Physics*. New York: McGraw-Hill; Woodbridge. Conn.: Ox Bow Press, 1977.
Marlow, A.R. (ed.). 1978. *Mathematical Foundations of Quantum Theory.* New York, Academic.
Marshall-Pescini, S., and A. Whiten. 2008. Chimpanzees (Pan troglodytes) and the Question of Cumulative Culture: An Experimental Approach. *Animal Cognition* 11: 449–456.
Merleau-Ponty, Maurice. 1942. *La structure du comportment*. Paris: PUF.
Merleau-Ponty, Maurice. 1945. *Phénoménologie de la perception*. Paris: Gallimard.
Mink, J.W., R.J. Blumenschine, and D.B. Adams. 1981. Ratio of Central Nervous System to Body Metabolism in Vertebrates: Its Constancy and Functional Basis. *American Journal of Physiology* 241 (3): R203–212.
Nielsen, M.A., and I.L. Chuang. 2000. *Quantum Computation and Quantum Information*. Cambridge: University Press.
Peirce, Charles S. 1865. On the Logic of Science: Harvard Lectures. In [Peirce W, I, 161–302].
Peirce, Charles S. 1866. The Logic of Science or Induction and Hypothesis: Lowell Lectures. In [Peirce W, I, 357–504].
Peirce, Charles S. 1868a. Upon Logical Comprehension and Extension. *Proceedings of the American Academy of Arts and Sciences* 7: 416–432; [Peirce W, II, 70–86].
Peirce, Charles S. 1868b. Some Consequences of Four Incapacities. *Journal of Speculative Philosophy* 2: 140–157; In [Peirce W, II, 211–242].
Peirce, Charles S. 1878. Deduction, Induction, and Hypothesis. *Popular Science Monthly***13**: 470–482; in [Peirce W, III, 323–338].
Peirce, Charles S. 1884. Design and Chance. In [Peirce W, IV, 544–554].
Peirce, Charles S. 1887. *A Guess at the Riddle*. In [Peirce W, VI, 165–210].
Peirce, Charles S. 1891. The Architecture of Theories, *Monist* 1: 161–176. In [Peirce W, VIII, 98–110].
Peirce, Charles S. 1893. Evolutionary Love. *Monist* 3: 176–200; in [Peirce W, VIII, 184–205].
Peirce, Charles S. 1898. *Reasoning and the Logic of Things: The Cambridge Conferences Lectures of 1898*, Manuscript; Ed. K. L. Ketner. Cambridge, MA: Harvard University Press.

Peirce, Charles S. 1901. On the Logic of Drawing History From Ancient Documents, Especially From Testimonies. In [Peirce EP, II, 75–114].
Peirce, Charles S. 1903a. The Categories Defended. In [Peirce EP, II, 160–178].
Peirce, Charles S. 1903b. The Seven Systems of Metaphysics. In [Peirce EP, II, 179–95].
Peirce, Charles S. 1903c. Degenerate Cases. In [Peirce CP, 1.521–544].
Peirce, Charles S. 1907. Pragmatism. In [Peirce EP, II, 398–433].
Peirce, Charles S. 1982. *Writings*, vol. I. Bloomington: Indiana University Press.
Peirce, Charles S. 1998. *The Essential Peirce*, vols. I–II. Bloomington: Indiana University Press.
Peirce, Charles S. 2019. *The Collected Papers*, vols. I–VI, Ed. Charles Hartshorne and Paul Weiss. Cambridge, MA: Harvard University Press, 1931–1935; vols. VII–VIII (Ed. Arthur W. Burks). Cambridge, MA: Harvard University Press, 1958.
Poincaré, Henri. 1897a. Le rapport de l'analyse et de la physique mathématique. *Acta Mathematica* 21: 331–341; rep. in [Poincaré 1991, 17–30].
Poincaré, Henri. 1897b. Les idées de Hertz sur la mécanique. *Revue Générale des Sciences Pures et Appliqués* 8: 734–743; rep. In [Poincaré 1991, 62–86].
Poincaré, Henri. 1902. *La science et l'hypothèse*. Paris: Flammarion.
Poincaré, Henri. 1905. *La Valeur de la science*. Paris: Flammarion.
Poincaré, Henri. 1909. La logique de l'infini. *Revue de Métaphysique et de Morale* 17: 461–482; rep. in [?, 7–31].
Popper, Karl R. 1934. *Logik der Forschung*. Wien: Springer; 8th ed. Tübingen: Mohr, 1984.
Quine, Willard, and van Orman. 1960. *Word and Object, 1989*. Cambridge, MA: MIT University Press.
Quine, Willard, and van Orman. 1969. *Ontological Relativity and Other Essays*. New York: Columbia University Press.
Raichle, M.E., and D.A. Gusnard. 2002. Appraising the Brain's Energy Budget. *Proceedings of the National Academy of Sciences USA* 99: 10237–10239.
Ramon Fuentes, Miguel. 2016. *The Concept of Function in Molecular Biology: A Theoretical Framework and a Case Study*. Rome: G and B Press.
Rizzolatti, G., M. Gentilucci, R.M. Camarda, V. Callese, G. Luppino, M. Matelli, and L. Fogassi. 1990. Neurons Related to Reaching-grasping Arm Movements in The Rostral Part of Area 6 (area 6aβ). *Experimental Brain Research* 82: 337–350.
Russell, Bertrand. 1902. Letter to Frege. In *From Frege to Gödel*, ed. J. van Heijenoort, 124–125. Cambridge (MA).
Russell, Bertrand. 1903. *Principles of Mathematics*, 2nd ed. 1937. London: Routledge.
Schrödinger, Erwin. 1958. *Mind and Matter*. Cambridge: University Press; rep. in [Schrödinger 1992, 91–164].
Sengupta, B., M.B. Stemmler, and K.J. Friston. 2013. Information and Efficiency in the Nervous System-a Synthesis. *PLOS Computational Biology* 9 (7): e1003157.
Shimizu, T.S., Y. Tu, and H.C. Berg. 2010. A Modular Gradient–sensing Network for Chemotaxis in *Escherichia Coli* Revealed by Responses to Time–varying Stimuli., *Molecular Systems Biology* 6 (382): 1–14.
Shimony, Abner. 1970. Scientific Inference. In *The Nature and Function of Scientific Theories*, ed. R.G. Colodny, Pittsburgh: 79–172; rep. in [Shimony 1993, I, 183–273].
Smolin, Lee. 2013. *Time Reborn: From the Crisis in Physics to the Future of the Universe*. Boston: Houghton Mifflin Harcourt.
Sperber, Dan (ed.). 2000. *Metarepresentations: A multidisciplinary Perspective*. Oxford: University Press.
Spivak, David I. 2013. *Category Theory for Scientists*. Cambridge, MA: MIT Press.
Still, S., D.A. Sivak, and A.J., Bell, and G.E. Crooks. 2012. Thermodynamics of Prediction. *Physical Review Letters* 109: 120604.
Stout, D., and T. Chaminade. 2012. Stone Tools, Language and the Brain in Human Evolution. *Philosophical Transactions of the Royal Society* B367: 75–87.
Suppes, Patrick. 1960. *Axiomatic Set Theory*. Toronto: Van Nostrand.

Uhlen, M., and F. Ponten. 2005. Antibody-Based Proteomics for Human Tissue Profiling. *Molecular and Cellular Proteomics* 4 (4): 384–393. https://doi.org/10.1074/mcp.R500009-MCP200.

von Helmholtz, Hermann L.F. 1867. *Handbuch der physiologischen Optik*. 2nd ed. L. Voss: Hamburg and Leipzig.

von Helmholtz, Hermann L.F. 1883. *Wissenschaftliche Abhandlungen*, vol. 2. Hamburg: J. A. Barth.

West-Eberhard, Mary Jane. 2003. *Developmental Tlasticity and Evolution*. Oxford: University Press.

Wheeler, John A. 1983. Law Without Law. In [Marlow 1978, 9–48].

Wheeler, John A. 1988. World as System Self-synthesized by Quantum Networking. *IBM Journal of Research and Development* 32: 4–15. https://doi.org/10.1147/rd.321.0004.

Wheeler, John A. 1990. Information, Physics, Quantum: The Search for Links. In [Zurek 1990, 3–28].

Zermelo, Ernst. 1908. Untersuchungen über die Grundlagen der Mengenlehre I. *Mathematische Annalen* 65: 261–281. https://doi.org/10.1007/BF01449999.

Zermelo, Ernst. 1930. Über Grenzzahlen und Mengenbereiche. *Fundamenta Mathematicae* 16: 29–47.

Zurek, Wojciech H. (ed.). 1990. *Complexity, Entropy and the Physics of Information*. Redwood City: Addison-Wesley.

Bibliography

Allcock, G.R. 1969. The Time of Arrival in Quantum Mechanics. *I-III. Annals of Physics* 53: 253–348.

Auletta, G. 2003. Some Lessons of Quantum Mechanics for Cognitive Science. *Intellectica* 36–37: 293–317.

Auletta, G. 2004. Decoherence and Triorthogonal Decomposition. *International Journal of Theoretical Physics* 43: 2263–2274.

Auletta, G. 2008. Biological Systems Integrating Information and Entropic Fluxes in [Auletta 2008c, 27–42].

Auletta, G. 2010. A Paradigm Shift in Biology? *Information* 1: 28–59. https://doi.org/10.3390/info1010028. http://www.mdpi.com/2078-2489/1/1/28/.

Auletta, G. 2012. Inferences with Information. *Universal Journal of Applied computer Science and Technology* 2.2: 216–221. http://uniascit.in/files/documents/2012_020201.pdf.

Auletta, G., G. Tarozzi. 2006. Premeasurement Versus Measurement: A Basic Form of Complementarity. In *The Foundations of Quantum Mechanics: Historical Analysis and Open Questions*, ed. C. Garola, A. Rossi, and S. Sozzo, 40–47. Singapore.

Barnett, S.M. 2009. *Quantum Information*. Oxford: Oxford University Press.

Bell, J.S. 1973. Subject and Object. In [Mehra 1973]; Report in [Bell 1987, 40–44].

Bell, J.S. 1987. *Speakable and Unspeakable in Quantum Mechanics*. Cambridge: Cambridge University Press.

Bennett, C.H. 1986. On the Nature and Origin of Complexity in Discrete, Homogeneous, Locally-Interacting Systems. *Foundations of Physics* 16: 585–592.

Bennett, C.H. 2003. Quantum Information: Qubits and Quantum Error Correction. *International Journal of Theoretical Physics* 42: 153–176. https://doi.org/10.1023/a:1024439131297.

Bennett, C.H., G. Brassard, and A. Ekert. 1992. Quantum Cryptography. *Scientific American* 267 (4): 26–33.

Bohr, N. 1973. Causality and Complementarity. *Philosophy of Science* 4: 289–298; Also report in [BOHR CW X, 37-48].

Bohr, C.W. 1985. *Collected Works*. Amsterdam: North Holland.

Chaitin, G.J. 1987. *Algorithmic Information Theory*. Cambridge: Cambridge University Press.

Chaitin, G.J. 1999. *The Unknowable*. Singapore: Springer.

Chiribella, G., G.M. D'Ariano, and P. Perinotti. 2011. Informational Derivation of Quantum Theory. *Physical Review A* 84: 012311. https://doi.org/10.1103/PhysRevA.84.012311.

© Springer Nature Switzerland AG 2019

G. Auletta, *The Quantum Mechanics Conundrum*,
https://doi.org/10.1007/978-3-030-16649-6

Cushing, J., and E. Mcmullin (eds.). 1989. *Philosophical Consequences in Quantum Theory: Reflections on Bell's Theorem*. USA: Notre Dame University Press.
de Muynck, W.M. 2000. Preparation and Measurement: Two Independent Sources of Uncertainty in Quantum Mechanics. *Foundations of Physics* 30: 205–225.
Ekert, A.K. 1991. Quantum Cryptography Based on Bell's Theorem. *Physical Review Letters* 67: 661–663.
Fano, U. 1957. Description of States in Quantum Mechanics by Density Matrix and Operator Techniques. *Review of Modern Physics* 29: 74–93. https://doi.org/10.1103/RevModPhys.29.74.
Grot, N., C. Rovelli, and R.S. Tate. 1996. Time-of-Arrival in Quantum Mechanics. *Physical Review* A54: 4676–4690.
Hartle, J.B. 2010. Quasiclassical Realism. in [Saunders et al. 2010, 73–98].
Heisenberg, W. 1971. *Physics and Beyond: Encounters and Conversations* New York: Harper and Row; Eng. Tr. of [Heisenberg 1969].
Leibniz, G.W. 1678. Ad Ethicam B. d. Sp. in [Leibniz PS, I, 139–152].
Leibniz, P.S. 1978. *Philosophische Schriften* (Ed. Gerhardt), Halle, 1875. rep. Olms: Hildesheim.
Leibniz, G.W. 1988. Opuscules et Fragments inédits (Ed. Couturat). Paris: Alcan, 1903; Report Hildesheim, Olms.
Lewontin, R.C. 1970. The Units of Selection. *Annual Review of Ecology and Systematics* 1: 1–18.
Mehra, J. (ed.). 1973. *The Physicist's Conception of Nature*. Dordrecht: Reidel.
Peirce, C.S. 2019. *The New Elements of Mathematics* (Ed. Carolyn Eisele). Mouton: The Hague.
Poincaré, H. 1913. *Dernières Pensées*. Paris: Flammarion.
Putnam, H. 1988. *Representation and Reality*. Cambridge: MIT Press.
Schumacher, B.W. 1995. Quantum Coding. *Physical Review* A51: 2738–2747.
Saunders, S., J. Barrett, A. Kent, and D. Wallace (eds.). 2010. *Many Worlds?*. Oxford: Oxford University Press.
Toffoli, T. 1988. Information Transport Obeying the Continuity Equation. *IBM Journal of Research and Development* 32: 29–36. https://doi.org/10.1147/rd.321.0029.
Turing, A.M. 1950. Computing Machinery and Intelligence. *Mind* 59: 433–460.
Vedral, V. 2010. *Decoding Reality: The Universe as Quantum Information*. Oxford: Oxford University Press.
von Helmholtz, H.L.F. 1882. *Wissenschaftliche Abhandlungen*, vol. 1. Hamburg: J. A. Barth.
von Helmholtz, H.L.F. 1895. *Wissenschaftliche Abhandlungen*, vol. 3. Hamburg: J. A. Barth.
von Neumann, J. 1975. *Evolution of an Evolutionist*. New York: Cornell University Press.
Worrall, J. 1989. Structural Realism: The Best of Both Worlds? *Dialectica* 43: 99–124. https://doi.org/10.1111/j.1746-8361.1989.tb00933.x.
Zeilinger, A. 2004. Why the Quantum?It From Bit? A Participatory Universe?. Three Far–reaching Visionary Questions From John Archibald Wheeler and How They Inspired a Quantum Experimentalist [BARROW et al. 201–220].

Author Index

A

Abel, Niels H., 184
Abramsky, Samson, 315, 320, 717, 718, 752, 754
Acin, A., 317, 327, 333
Adami, C., 240, 514
Adami, Christoph, 317
Adams, D. B., 799
Adleman, Leonard M., 796
Adler, Stephen L., 226
Afriat, Alessandro, 503
Agrawal, Govind P., 708
Aguirre, Anthony, 234, 249, 634, 649
Aharonov, Yakir, 290, 291, 428
Alai, Mario, 166
Alberts, Bruce, 790, 792
Alley, C. O., 306
Almeida, M. L., 317, 333
Anagnostopoulos, Georgios, 341
Anscombe, G. Elizabeth M., 8, 199
Apostol, Tom M., 18
Araki, Huzihiro, 258
Aravind, P. K., 338
Aristotle, 9, 339, 341, 343, 345, 347, 413, 717, 805, 809, 811
Armstrong, David M., 197, 339, 460, 498, 780
Arnold, Vladimir I., 88
Arthur, J., 286
Ashby, Ross W., 791
Aspect, Alain, 25, 305, 306
Atkins, Peter W., 411, 687, 699, 700, 707, 709
Atwood, D. K., 286
Auletta, Gennaro, *passim*

B

Baez, John, 752, 753
Ballentine, Leslie E., 226, 419
Barabási, Albert–László, 343
Barbieri, Marcello, 791, 792
Barbour, Ian, 656
Barenco, A., 376
Bar–Hillel, Yehoshua, 718
Barnett, Stephen M., 240, 254, 255
Barnum, Howard, 501
Barrat, Alain, 343
Barrett, Jonathan, 232, 426
Barrow–Green, June, 690
Barrow, John D., 234
Barthélemy, Marc, 343
Battail, Gérard, 272, 319, 517, 551, 580
Batterman, Robert W., 674
Batygin, Konstantin, 708
Baudrillard, Jean, 826
Bauer, Edmond, 216, 219, 220, 225, 227, 242, 251
Baumgarten, Alexander Gottlieb, 25, 347
Bayes, Thomas, 554
Bedau, M. A., 674
Bekenstein, Jacob D., 259, 666–668
Bell, A. J., 788
Bellan, Paul M., 708
Bell, John S., 161, 231, 242, 261, 292, 300, 301, 303, 336, 348, 823
Beltrametti, Enrico G., 515, 782
Bénard, Henri, 414
Bennett, Charles H., 270, 311, 317, 320, 368, 440, 447, 513, 557
Berg, Howard C., 796
Berkeley, George, 222
Bernoulli, Jacob, 141
Berry, Michael V., 108, 290, 698, 699

© Springer Nature Switzerland AG 2019
G. Auletta, *The Quantum Mechanics Conundrum*,
https://doi.org/10.1007/978-3-030-16649-6

Bessel, Friedrich Wilhelm, 56
Bialynicki–Birula, Iwo, 286
Bilicki, M., 670
Bird, Alexander, 197, 280
Birkhoff, Garrett D., 782
Bjorken, James D., 538, 632
Blake, C., 670
Bloch, Felix, 469
Blumenschine, R. J., 799
Bohm, David, 168, 286–291, 311, 333, 417, 428, 434, 435, 444, 488, 573
Bohr, Niels, 5–7, 71, 108, 109, 151, 155, 159, 191, 193, 195, 196, 198, 201, 204, 221, 246, 260, 275, 278, 279, 281, 282, 285, 292, 299, 301, 348, 350, 406, 410, 420, 426, 435, 440, 489–491, 555, 620, 647, 648, 812, 823
Bokulich, Alisa, 106, 108, 137, 421
Boltzmann, Ludwig, 2, 107, 341, 552
Boole, George, 757, 758
Borel, Emile, 412
Born, Max, 5, 8, 14, 26, 31, 194, 197, 341, 346, 351, 421, 549
Boscovich, Rogerius J., 7, 192, 600
Bose, Satyendra N., 102
Bothe, Walther, 290
Bourret, Robert B., 796
Bouwmeester, Dik, 316
Braginsky, Vladimir B., 279
Brassard, Gilles, 311, 317
Braunstein, Samuel L., 304, 309, 316, 501
Bray, Dennis, 790, 792
Bricmont, Jean, 288
Bridgman, Percy W., 349, 442, 443
Brillouin, Léon, 271, 551
Brough, S., 670
Brout, Robert, 631
Brouwer, M. M., 670
Brown, Douglas A., 796
Brown, Harvey R., 656, 672
Brown, R., 422
Brown, Robert, 710
Brukner, Caslav, 250
Brune, M., 245
Brunner, N., 503
Bub, Jeffrey, 241, 242, 299, 508
Buddelmeijer, H., 670
Busch, Paul, 461, 479, 480, 783
Bužek, Vladimir, 273
Byron, Frederick W. Jr., 18, 29, 32, 37, 38, 42, 44, 46, 52, 56, 63, 66–68, 78, 79, 90, 111, 116, 537, 600, 601, 607, 771, 774, 776

C
Callese, V., 801
Camarda, R., 801
Carnap, Rudolf, 223, 339, 410, 489, 779
Carroll, John W., 197
Cassinelli, Gianni, 782
Cassirer, Ernst, 198
Cauchy, Augustin–Louis, 82
Cavalcanti, D., 317, 333
Caves, Carlton M., 501, 525, 554
Cerf, Nicolas J., 240, 317, 514
Chaitin, Gregory J., 165, 367, 718
Chalmers, David J., 235
Chaminade, T., 800
Chau, H. F., 508
Chellas, Brian F., 780
Chersi, Fabian, 801
Chibbaro, Sergio, 551, 674, 698, 699, 710
Chiribella, Giulio, 122, 132, 151, 193, 194, 267, 345, 447, 464, 479, 486, 507, 518, 526, 536, 558, 560, 562, 567, 574, 580, 582–584, 586–588, 590, 591, 655, 820, 823
Chomsky, Noam, 803
Chuang, Isaac L., 122, 255, 265, 267, 272, 333, 368, 372, 375, 379, 389, 391, 392, 395, 397, 440, 447, 449, 465, 468, 473, 504, 506, 799
Chuaqi, Rolando, 718
Church, Alonzo, 367, 372
Clark, Andy, 799
Clauser, John F., 186, 303
Claverie, P., 703
Clayden, Jonathan, 551, 700
Clayton, Philip, 674
Clebsch, Alfred, 101
Clifford, William K., 393
Clifton, R., 508
Coecke, Bob, 315, 320, 486, 752, 754, 782
Cohen–Tannoudji, Claude, 85, 271, 549, 556, 557, 635
Colagè, Ivan, 798
Collett, Matthew J., 458
Collins, Harry, 340, 498
Compton, Arthur H., 7, 10
Conant, R. C., 791
Conway, J., 155, 161, 268, 290, 300
Cook, Stephen A., 368
Corrêa, R., 480
Cosmides, L., 801
Cottingham, W. Noel, 613, 622
Cramer, John G., 655
Crepeau, Claude, 311

Crile, George, 800
Croca, J. R., 286
Crooks, G. E., 788

D
Da Costa, Newton, 165, 367, 718
D'Alembert, Jean Baptiste Le Rond, 178
Dalibard, Jean, 305
Dall, R. G., 202
D'Ambrosio, Paolo, 798
D'Ariano, G. Mauro, 122, 132, 151, 193, 194, 267, 345, 408, 419, 447, 464, 479, 486, 500, 507, 518, 526, 536, 558, 560, 562, 567, 574, 580, 582–584, 586–588, 590, 591, 650, 655, 820, 823
Darwin, Charles R., 196
Davies, E. B., 461, 481
Davies, Paul C. W., 674
Davisson, Clinton J., 12
Deacon, Terrence W., 796, 801
De Broglie, Louis, 10–12, 15, 26, 286, 417
De Coulomb, Charles-Augustin, 600, 602, 606, 681
Dehaene, Stanislas, 812
De La Forge, Louis, 188
Democritus, 675
De Morgan, Augustus, 295
De Muynck, Willem M., 246, 461, 474
Denkmayr, Tobias, 480
De Paula, Julio, 411, 687, 699, 700, 707, 709
Descartes, René, 9, 160, 188
D'Espagnat, Bernard, 86, 280
Deutsch, David, 130, 223, 227, 232–235, 255, 264, 279, 352, 365, 369, 372, 376, 382, 385, 413, 423, 497, 498, 510, 511, 696, 798, 814, 816
DeWitt, Bryce S., 228, 663
Dieks, Dennis, 31, 132, 348, 499
Dirac, Paul A. M., 13, 14, 24, 25, 46, 95, 102, 106, 421, 538, 606, 612, 622
Dolev, Shahar, 656
Doria, Francisco A., 165, 367, 718
Drell, Sidney D., 538, 632
Dreyer, J., 245
Dvornik, A., 670

E
Earley, Joseph E., 709
Earman, John, 172, 189
Eberhard, Philippe H., 320, 321, 323
Eddington, Arthur Stanley, 13, 166, 221, 786
Edelman, Gerald M., 343
Eibl, Manfred, 316
Einstein, Albert, 3, 4, 7, 13, 102, 106, 109, 164–168, 170–172, 174, 187, 189–191, 193, 222–224, 226, 227, 248, 275, 277, 278, 280–282, 284–286, 299, 301, 342, 345, 346, 348–350, 407–409, 419, 420, 489–491, 550, 556, 584, 599, 612, 637, 646, 648, 654, 655, 670, 710, 711, 799, 805, 807, 811, 819, 823, 824
Ekert, Artur, 376
Ekert, Artur K., 308
Elby, Andrew, 241, 242
Elitzur, Avshalom C., 145, 146, 412, 549, 552, 557, 656
Ellis, George F. R., 339, 341, 696, 791
Englert, Berthold–Georg, 109, 110, 231, 290, 439, 454, 455, 458
Englert, François, 631
Epicurus, 675
Epperson, Michael, 231, 348, 487, 782
Erben, T., 670
Euler, Leonhard, 21, 57, 381
Everett, Hugh III, 226–228, 231

F
Fano, Vincenzo, 286
Faraday, Michael, 600, 603, 659
Fedrizzi, Alessandro, 440, 652, 653
Fermi, Enrico, 102
Ferrari, Pier Francesco, 801
Ferrero, Miguel, 186, 308
Feynman, Richard P., 161, 365, 369, 497, 539, 540, 612, 632, 655, 825
Fine, Arthur, 277
Finkelstein, David R., 465
Fock, Vladimir, 522, 611
Fogassi, Leonardo, 801
Fortunato, Mauro, *passim*
Fourier, J.-B. Joseph, 56
Fraenkel, Abraham, 718
Franson, J. D., 25
Fredkin, Edward, 370
Frege, Gottlob, 718
French, A. P., 172, 175, 176
Friedman, Michael, 172, 637, 814
Friedman, Ronald, 411, 687
Friston, Karl J., 788, 794, 801
Fuchs, Christopher A., 316, 501, 526, 554
Fuchs, Jürgen, 90, 91, 605
Fuller, Robert W., 18, 29, 32, 37, 38, 42, 44, 46, 52, 56, 63, 66–68, 78, 79, 90, 111,

116, 537, 600, 601, 607, 771, 774, 776
Furusawa, A., 316

G
Galilei, Galileo, 173
Gally, Joseph A., 343
Galois, Évariste, 91
Garuccio, Augusto, 286, 308
Gassed, Pierre, 188
Gauss, Carl Friedrich, 55, 669
Gell–Mann, Murray, 627
Gentilucci, Maurizio, 801
Geppert, Hermann, 480
Gerlach, Walther, 95
Germer, Lester H., 12
Geroch, Robert, 178, 222, 406, 407, 441, 581, 656, 733
Gesierich, Benno, 801
Geulincx, Arnold, 188
Gibbs, Josiah Willard, 105
Gibson, James J., 802
Gilbert, Scott F., 793
Gisin, Nicolas, 327, 441
Giulini, D., 237
Givant, S., 757, 761, 763, 765
Gleason, Andrew M., 292
Gnedenko, Boris V., 140, 531
Gödel, Kurt, 205, 779
Gordan, Paul, 101
Gordon, Walter, 616
Gottesman, Daniel, 392, 395, 397, 440
Grabowski, Marian, 461, 479, 480, 783
Graham, Neill, 232
Grangier, Philippe, 25
Grayson, T. P., 286
Greenberger, Daniel M., 206, 334, 436, 439
Green, George, 537, 600
Greenwood, Derek A., 613, 622
Greeves, Nick, 551, 700
Greiner, W., 627
Griffiths, Robert B., 239
Grover, Lov K., 389
Grünbaum, Adolf, 441, 442
Guralnik, Gerald, 631
Gusnard, D. A., 799

H
Habib, Salman, 243
Hadamard, Jacques, 143
Hagen, C. Richard, 631
Hagley, E., 245
Halliwell, Jonathan J., 239, 658
Halmos, Paul R., 757, 761, 763, 765
Halvorson, H., 508
Hamilton, William Rowan, 15, 18, 288
Hardy, Lucien, 286
Haroche, Serge, 245
Harrigan, Nicholas, 277
Harrison, Lee, 801
Hartle, James B., 239
Hasegawa, Yuji, 480
Hauser, Marc D., 802, 803
Hawking, Stephen W., 239
Hayat, Axel, 264, 273
Hayden, Patrick, 510, 511
Healey, Richard, 427
Hegel, George W. F., 162, 222
Hegerfeldt, Gerhard C., 409
Heisenberg, Werner, 14, 31, 62, 80, 83, 85, 86, 130, 156, 161, 198, 215, 224, 246, 248, 278, 282, 348–351, 417, 419, 420, 445, 489–491, 552, 812
Hellmuth, T., 202
Helstrom, Carl W., 553
Hempel, Carl G., 280, 347, 674
Herbart, Johannes F., 827
Herbst, Thomas, 440, 652, 653
Hermite, Charles, 37
Heymans, C., 670
Hibbs, A. R., 161, 369, 539, 540
Higgs, Peter, 631
Hilbert, David, 13, 718
Hildebrandt, H., 670
Hiley, Basil J., 289, 290, 573
Hoekstra, H,, 670
Holevo, Alexander S., 265
Holland, Peter R., 288, 289
Holt, Richard A., 186, 303
Holwerda, B. W., 670
Home, Dipankar, 7, 106, 108, 190, 221, 248, 275, 279, 285, 290, 350
Hopkins, A. M., 670
Horne, Michael A., 186, 286, 303, 308, 334
Horodecki, Michael, 317
Howard, Don, 186, 277, 409
Huang, Kerson, 250, 705, 706
Hume, David, 194, 809, 826
Humphreys, P., 674
Husserl, Edmund, 222
Huygens, Christiaan, 537

I
Imamoḡlu, Atac, 520
Ionicioiu, R., 503

J
Jackson, John D., 602, 603, 609
Jacobi, Carl G. J., 18, 65, 288
Jaeger, Luc, 339, 341, 696, 791
Jammer, Max, 2, 15, 26, 31, 62, 80, 95, 106, 108, 171, 191, 226, 277, 285, 287, 419
Jaynes, Edwin T., 816
Jeannerod, Marc, 791, 827
Jeans, James H., 107
Jennewein, Thomas, 440, 652, 653
Johnson, Alexander, 790, 792
Jona–Lasinio, G., 703
Joos, Erich, 237, 411
Jordan, Pascual, 14, 31, 607
Joshi, Shiv D., 518
Jozsa, Richard, 311, 376, 501
Jurica, Melissa S., 796

K
Kandel, Eric R., 804
Kant, Immanuel, 25, 86, 192, 222, 347, 420, 555
Kastner, Ruth E., 19, 441, 487, 655
Kaszlikowski, D., 328
Kent, Adrian, 232
Kepler, Johannes, 95
Khakimov, R. I., 202
Khalili, Farid Y., 279
Khinchin, A. I., 250
Kibble, Tom, 631
Kiebel, S., 801
Kiefer, Claus, 237, 649
Killing, Wilhelm, 669
Kilner, James, 801
Kimble, H. Jeffrey, 316
Kim, Yoon–Ho, 458
King, B. E., 245
Kircher, Athanasius, 600
Kistler, Max, 193, 198, 205
Kitagawa, Masahiro, 520
Klaes, D., 670
Klauber, Robert D., 613
Klein, Oskar B., 607, 616
Knight, Peter L., 254
Kochen, Simon, 155, 161, 268, 290, 295, 300
Kofler, Johannes, 440, 652, 653
Kolmogorov, Andrei N., 140
Koyré, Alexandre, 189, 190
Kraus, Karl, 461, 463, 480, 481, 485, 486, 584
Krenn, Günter, 337
Kronecker, Leopold, 29
Kuhlmann, Meinard, 416, 558
Kuhn, Thomas S., 2, 5, 107, 406, 810, 816, 817
Kuijken, K., 670
Kulik, Sergei P., 458
Kupsch, J., 237
Kwiat, Paul, 447

L
Lagrange, Joseph–Louis, 18, 19, 21
Lahti, Pekka J., 461, 479, 480, 783
Lakatos, Imre, 817
Lal, Raymond, 486
Landauer, Rolf, 368, 370, 513, 557
Landau, Lev D., 19, 22, 65, 111, 138, 710
Landsberg, Peter T., 698
Landsman, Klaas, 65, 90, 113, 140, 155, 292, 295, 782
Langevin, Paul, 710
Laplace, Pierre–Simon de, 52, 193, 201, 272, 347
Laughlin, Robert B., 234, 413, 631, 674, 699, 781
LeDoux, Joseph E., 795
Legendre, Adrien–Marie, 78
Leibniz, Gottfried W., 7, 9, 18, 62, 187–193, 223, 232, 233, 311, 406, 444, 556, 698, 762, 818
Leighton, R. B., 825
Leinster, Tom, 718, 719, 739, 743
Lemaître, Georges, 671
Lemke, Jay L., 796
Lemmel, Hartmut, 480
Leon–Garcia, Alberto, 407
Levi–Civita, Tullio, 70
Levitin, Lev B., 401
Levy, Azriel, 718
Lewes, George H., 674
Lewis, David, 205, 232
Lewis, Julian, 790, 792
Lewontin, Richard C., 421
Lieb, Elliott H., 252, 258, 271
Lie, Sophus, 63, 91
Liftshitz, Evgenij M., 19, 22, 65, 138
Lindblad, Göran, 132, 249, 467
Ling, San, 267
Liouville, Joseph, 138
Liske, J., 670
Lloyd Morgan, Conwy, 674
Lloyd, Seth, 368, 369, 550, 557
Locke, John, 192

Lockwood, Michael, 235, 443
Lo, Hoi–Kwong, 508
London, Fritz, 216, 219, 220, 225, 227, 242, 251
Lorentz, Hendrik A., 175
Loveday, J., 670
Lüders, Gerhart, 219
Ludwig, Günther, 130, 151, 152, 479–481, 821
Luppino, G., 801

M
Maître, X., 245
Maali, A., 245
Mabuchi, Hideo, 524
Mach, Ernst, 190, 198, 798, 802
Mach, Ludwig, 142
Maczyński, Maciej J., 515
Mahler, Dylan H., 264, 273
Malament, David B., 178, 205, 409, 636, 657, 670
Malebranche, Nicolas, 188
Mana, Piero G. L., 250
Mandel, Leonard, 25, 231, 286, 287, 306, 452, 650
Mann, A., 304, 309
Manning, A. G., 202
Margenau, Henry, 8, 130, 150, 166, 199, 223, 278, 280, 339, 350, 410, 418, 804, 806, 814
Margolus, Norman, 401
Markov, Andrej A., 548, 793
Marshall–Pescini, S., 801
Marshall, Trevor W., 308
Martens, Hans, 474
Masanes, L., 327
Matelli, M., 801
Mattle, Klaus, 316, 447
Matzkin, Alexandre, 480
Maudlin, Tim, 172, 179, 205, 232, 280, 311, 440
Ma, Xiao–Song, 440, 652, 653
Maxwell, James C., 4, 271, 600, 602
Mayers, Dominic, 508
Maynard Smith, John, 697
McFarland, J., 670
McKann, Edwin, 192
Meekhof, D. M., 245
Mehra, Jagdish, 2, 162, 194, 197, 237, 248, 284
Mensky, Michael B., 524
Merleau–Ponty, Maurice, 799
Michelson, Albert A., 173
Mill, John Stuart, 674
Mink, J. W., 799
Minkowski, Hermann, 178
Misner, Charles W., 54, 179, 405, 407, 428, 634–636, 639, 641–646
Mitchell, Sandra D., 674
Mittelstaedt, Peter, 246
Monken, C. H., 480
Monroe, C., 245
Morley, Edward W., 173
Munowitz, Michael, 223
Mycielski, Jerzy, 286

N
Nagel, Ernest, 280, 347
Nakajima, R., 670
Naqsh, F., 250
Newton, Isaac, 4, 10, 21, 173, 190
Ng, Y. J., 369, 550
Nielsen, Michael A., 122, 255, 265, 267, 272, 333, 368, 372, 375, 379, 389, 391, 392, 447, 449, 465, 468, 473, 504, 506, 525, 799
Noether, A. Emmy, 88
Norsen, Travis, 232
Norton, John D., 195

O
O'Brien, J. L., 503
Ohya, M., 250
Ollivier, Harold, 429, 431
Oppenheim, Jonathan, 317, 328
Oppenheim, Paul, 674
Ott, Edward, 271
Ou, Z. Y., 306

P
Pan, Jian–Wei, 316
Paquette, É. O., 782
Parisi, Giorgio, *passim*
Parseval des Chênes, Marc–Antoine, 56
Partovi, Hossein M., 257
Pasnau, Robert, 342
Paterek, T., 328
Pati, Arun K., 501
Patney, Rakesh K., 518
Pauli, Wolfgang, 95, 97, 105, 350, 527, 620
Pawłowski, Marcin, 326, 328, 329
Paz, Juan P., 243
Pearl, Judea, 279

Peirce, Charles S., 25, 185, 189, 224, 247, 343, 418, 421, 487, 698, 779, 781, 796, 797, 804–808, 814, 816, 820, 826
Penrose, Roger, 62, 91, 104, 178, 393, 428, 623, 638, 639, 649, 659, 670
Peres, Asher, 121, 226, 311, 526
Perinotti, Paolo, 122, 132, 151, 193, 194, 267, 345, 447, 464, 479, 486, 507, 518, 526, 536, 558, 560, 562, 567, 574, 580, 582–584, 586–588, 590, 591, 650, 655, 820, 823
Peruzzo, A., 503
Peshkin, M., 290
Petz, D., 250
Pfister, Corsin, 526
Pfleegor, R. L., 25
Phoenix, Simon J. D., 240, 254, 255
Pitowsky, Itamar, 295
Planck, Max, 1–6, 30, 106, 193, 245, 248, 262, 812
Plato, 161
Plenio, M. B., 254
Plesch, Martin, 273
Podolsky, Boris, 109, 164–168, 170, 275, 277, 280–282, 284, 285, 299, 301, 811
Poincaré, Henri, 8, 18, 182, 184, 186, 194, 200, 263, 352, 412, 413, 418, 422, 552, 584, 599, 690, 764, 800, 809, 813, 814, 817, 818, 827
Poisson, Siméon Denis, 18, 65
Polzik, E. S., 316
Pons Domenech, Santiago, 196
Ponten, F., 790
Popescu, Sandu, 186, 324, 503
Popper, Karl R., 165, 190, 224, 814
Porter, T., 422
Potocki, K. A., 25
Prieur, A., 246
Proca, Alexandru, 619
Prugovečki, Eduard, 37, 527, 538
Pusey, Matthew F., 426

Q

Qarry, Angie, 440, 652, 653
Quine, Willard van Orman, 826

R

Raff, Martin, 790, 792
Raichle, Marcus E., 799
Raimond, J. M., 245
Ramelow, Sven, 440, 652, 653
Ramon Fuentes, Miguel, 790
Rapisarda, V., 286
Ratschbacher, Lothar, 440, 652, 653
Rayleigh, John W. S., 107
Rechenberg, H., 2, 162, 194, 197, 237, 248, 284
Rees, Martin J., 234, 249, 634, 649
Reid, Thomas, 224
Renninger, Mauritius K., 87, 144, 225, 352
Revzen, M., 304, 309
Riemann, Bernhard, 263, 540, 644
Rindler, Wolfgang, 8, 173, 174, 176, 177, 280, 407, 410, 619, 635–637, 640
Rippin, M. A., 254
Rizzolatti, Giacomo, 801
Roberts, Keith, 790, 792
Robertson, H. P., 82
Roger, Gérard, 25, 305
Rohrlich, David, 186, 324
Rondoni, Lamberto, 551, 674, 698, 699, 710
Rosen, Nathan, 109, 164–168, 170, 275, 277, 280–282, 284, 285, 299, 301, 811
Rovelli, Carlo, 270, 442, 480, 487, 649, 659–662, 664
Royer, Antoine, 482, 532
Rozema, Lee A., 264, 273
Rozzi, Stefano, 801
Rudolph, Terry, 426
Ruelle, David, 186
Russell, Bertrand, 195, 736
Rutherford, Ernest, 6

S

Safinejad, F., 250
Saha, Kaushik, 518
Saldanha, P. L., 480
Sánchez–Gómez, J. L., 186
Sands, M., 825
Santos, Emilio, 306, 308
Santos, M. F., 480
Saunders, Simon, 232
Scarani, Valerio, 317, 326, 328, 329, 333
Schack, Rüdiger, 554
Scheidl, Thomas, 440, 652, 653
Schelling, Friedrich W. J., 162
Schieder, R., 246
Schleich, Wolfgang P., 202
Schlosshauer, Maximilian, 26, 235, 237, 238, 244, 418, 467, 564
Schmidt, Erhard, 121
Schramm. S., 627

Schrödinger, Erwin, 14, 15, 26, 162, 170, 191, 226, 227, 235, 237, 244, 247, 248, 260, 271, 275, 282, 283, 285, 292, 301, 308, 417, 444, 489–491, 590, 797, 799, 812
Schumacher, Benjamin W., 264, 501, 515, 516
Schuster, Heinz Georg, 271
Schwartz, Matthew D., 106, 184, 611, 613, 617, 620, 622–624, 626
Schwarz, Hermann A., 82
Schweigert, Cristoph, 90, 91, 605
Scully, Marlan O., 109, 110, 231, 290, 454, 455, 458, 650
Selleri, Franco, 286, 308, 341
Sengupta, Biswa, 788, 794
Serjeant, Stephen, 250, 634, 666
Shadbolt, P., 503
Shafiee, A., 250
Shannon, Claude E., 250, 271, 272, 580
Shih, Yanhua H., 306, 458
Shimizu, Thomas S., 796
Shimony, Abner, 186, 201, 235, 260, 279, 303, 311, 334, 348, 350, 814
Shor, Peter W., 387, 440
Shull, C. G., 286
Sifón, C., 670
Singh, Pushpendra, 518
Sivak, D. A., 788
Smart, J. J. C., 443
Smolin, Lee, 235, 781
Smoluchowski, Marian, 710, 711
Søensen, J. L., 316
Sommerfeld, Arnold, 634
Specker, Ernst, 295
Spekkens, Robert W., 277
Spinoza, Baruch, 191, 227, 413
Spivak, David I., 305, 719, 722, 724–728, 732, 742, 745, 747, 760
Sponar, Stephan, 480
Stachel, John, 221, 282
Stamatescu, I.–O., 237
Stanzione, Massimo, 343
Stapp, Henry P., 221
Stay, M., 752, 753
Stefanov, A., 441
Steinberg, Aephraim M., 264, 273
Stein, E., 627
Stemmler. M. B., 788
Stephan, Klaas E., 801
Stern, Otto, 95
Still, S., 788
Stock, Ann M., 796
Stoddard, Barry L., 796
Stoffels, W. W., 474
Stokes, George G., 601, 710
Stone, Marshall H., 90
Storey, E. Pippa, 458
Stout, D., 800
Suarez, Antoine, 440, 441
Suppes, Patrick, 718
Susskind, Leonard, 633, 666
Svozil, Karl, 206
Synge, John Lighton, 637, 673
Szathmáry, Eörs, 697

T

Tan, Sze M., 458
Tarozzi, Gino, 159, 166, 286, 350, 417, 423, 424, 503
Taylor, Brook, 18
Taylor, E. N., 670
Tegmark, Max, 234, 249, 634, 649
Terno, D. R., 503
Tetik, Nuray, 440, 652, 653
Thomson, George P., 12
'T Hooft, Gerard, 104, 109, 131, 329, 340, 410, 421, 657, 666
Thorne, Kip S., 54, 179, 405, 407, 428, 634–636, 639, 641–646
Timpson, Christopher G., 262, 320, 339, 369, 511, 512, 580
Tinkham, Michael, 90
Tipler, Frank J., 234
Toffoli, Tommaso, 370
Tollaksen, Jeff, 480
Tonomura, A., 290
Tooby, J., 801
Torcal, Lluc, 342, 344, 418, 436, 486, 821
Tosini, Alessandro, 408, 584
Truscott, A. G., 202
Tryon, Edward P., 200
Tsirelson, Boris S., 326
Turing, Alan M., 366, 367, 372
Turner, Peter S., 264, 273
Tu, Yuhai, 796
Tzevelekos, N., 717, 718

U

Ueda, Masahito, 520
Uhlen, M., 790
Unruh, William, 667
Ursin, Rupert, 440, 652, 653

V
Vaidman, Lev, 145, 146
Valentijn, E. A., 670
Van Fraassen, Bas C., 31, 198, 351
Vedral, Vlatko, 254
Venn, John, 430
Verlinde, Erik, 441, 666, 667, 670
Vespignani, Alessandro, 343
Vigier, Jean-Pierre, 286
Visser, M. R., 670
Von Bertalanffy, Ludwig, 489
Von Helmholtz, Hermann L. F., 601, 795, 804, 827
Von Laue, Max T. F., 248
Von Neumann, John, 90, 138, 156, 159, 160, 222, 242, 291, 481, 504, 782
Vulpiani, Angelo, 551, 674, 698, 699, 710

W
Wallace, David, 130, 231, 232, 234, 239, 243
Walls, Daniel F., 458
Walter, Peter, 790, 792
Walther, H., 231
Walther, Herbert, 109, 202, 454, 455, 458
Wang, L. J., 231, 286, 452
Wang, Shang-Yung, *passim*
Warren, Stuart, 551, 700
Watson, Andrew, 627, 675
Watson, James D., 790, 792
Wehner, Stephanie, 328, 526
Wehrl, Alfred, 251, 258
Weinfurter, Harald, 316, 447
West–Eberhard, Mary Jane, 793
Weyl, Hermann, 190
Wheeler, John A., 8, 54, 162, 179, 202, 203, 221, 261, 369, 405, 407, 410, 411, 418, 422, 425, 428, 441, 443, 444, 497, 547, 550, 552, 555, 586, 634–636, 639, 641–646, 655, 663, 780, 781, 828
Whitaker, M. Andrew B., 7, 106, 108, 190, 221, 248, 275, 279, 285, 290, 350
Whitehead, Alfred N., 201, 299, 348, 450
Whiten, Andrew, 801
Wiener, Norbert, 552
Wien, Wilhelm C. W., 556
Wiesner, Stephen J., 311, 317, 330, 447
Wigner, Eugene P., 93, 221–223, 225, 248, 260, 351, 409, 426, 528, 532, 607
Wilczek, Frank, 105, 234, 249, 629, 631, 634, 649, 675, 676
Wineland, D. J., 245
Winter, Andreas, 317, 328
Withaker, M. Andrew B., 280
Wootters, William K., 132, 246, 311, 499
Wothers, Peter, 551, 700
Wu, Guozhen, 704
Wunderlich, C., 245

X
Xing, Chaoping, 267

Y
Yasin, Allaine, 436, 439
Young, Thomas, 9
Yuen, Horace P., 132, 419, 500, 528
Yu, Rong, 458

Z
Zafiris, Elias, 231, 348, 487, 782
Zajonc, A. G., 202
Zbinden, H., 441
Zeh, H. Dieter, 235, 237, 411
Zehnder, Ludwig L. A., 142
Zeilinger, Anton, 196, 222, 250, 261, 308, 316, 334, 337, 440, 447, 652, 653
Zermelo, Ernst, 718
Zoller, Peter, 524
Zou, X. Y., 231, 286, 452
Zukowski, Marek Z., 308, 328
Zurek, Wojciech H., 132, 162, 226, 228, 231, 235, 236, 240–243, 246, 268, 348, 429, 431–433, 435, 499, 696
Zweig, George, 627
Zwiebach, Barton, 650
Zwolak, Michael, 242, 432, 433, 435

Subject Index

A
Abduction, 5, 157, 285, 807–812
— as inverse Bayesian inference, 812
— as retrograde inference, 816
Abelian group, 184, 628
Absorber, 655
Action, 19, 538, 541, 542
– at–a–distance, 233
No —, 549
– at-a-distance, 189
– reaction principle, 173
Principle of least —, 19, 543
Quantum of —, 193, 556
Active transformation, 133
Actuality, 340, 413
Adaptation, 797, 798
Knowledge as —, 798
Additivity, 16
Adjoint, 14, 16, 753
— matrix, 33, 34
— vector, 14
Aether, 139, 173, 289
Affordance, 802
Aharonov–Bohm effect, 290, 291, 428, 488
Algorithm, 367, 368
Deutsch's —, 384
Efficient —, 367
Analogy, 421, 422
Analyticity, 18
Ancilla, 312
Angular momentum
— eigenfunction, 75, 77, 78
— eigenvalues, 75
Orbital —, 65, 66, 68, 70, 678, 679
Reduced —, 68
Quantisation of —, 71
Total —, 89, 100
Annihilation operator, 521–523, 610
Antibaryon, 630
— number, 634
Antibonding, 701, 702
Anticolour, 628
Antiparticle, 103, 620
Antiquark, 627, 628, 630
Anti-unitary transformation, 92
Apparatus, 152, 158, 219
— as a regulator, 578, 579
— as coupling device, 460
— as macroscopic, 246
—'s basis, 157
Setting of —, 320, 321
Araki–Lieb inequality, 258, 259, 434
Aspect, 722, 723, 744, 802
Association, 809
Metaphoric —, 810
Metonymic —, 810
Associativity, 90
Atom, 6
Bohr's atomic model, 6
Rutherford's atomic model, 6
Atomicity of the composition, 586, 588
Axiomatic system, 718

B
Baryon, 628–630
— number, 628–630, 634
Basis
Bases degeneracy, 228
Change of —, 39–43, 134
Computational —, 267
Orthonormal —, 29, 30
Preferred —, 240, 241
Bayes theorem, 554

© Springer Nature Switzerland AG 2019

G. Auletta, *The Quantum Mechanics Conundrum*,
https://doi.org/10.1007/978-3-030-16649-6

BB84 protocol, 318, 508, 509
Beable, 348
Beam splitter, 104, 142
– as unitary operator, 143
Polarisation —, 652
Bekenstein bound, 667
Bell inequalities
Quadrilateral information–distance, 515
Bell inequality, 300, 303, 515
Bell measurement, 310, 312
Bell operator, 304, 310, 325, 326
Bell states, 309, 310, 377, 378, 398
Bénard cell, 414, 415
Bernoulli trial, 141, 671
Bessel–Parseval relation, 56
Big Bang, 200, 656, 657
Binomial coefficient, 104
Bit commitment, 507, 508
No —, 508–510, 512
Black body, 1–3, 5, 30, 107
—'s spectrum, 3, 107
Black hole, 250, 633, 666
Entropy of —, 666
Bloch sphere, 470, 471, 473, 474, 777
Bloch vector, 469–471
Block universe, 190, 200, 442, 443
Bohm–Jacobi equation, 289
Bohr magneton, 620
Bohr radius, 682
Boltzmann constant, 107, 556, 557
Bonding, 701, 702
Boolean algebra, 757–759, 761–766
Complete —, 761
Boolean function, 372
Balanced —, 384
Constant —, 384
Boost
—s generator, 183, 184, 620
Borromean ring, 338
Bose–Einstein statistics, 102, 105, 522, 610
Boson, 102–104
—s as indistinguishable, 104, 105, 131
Higgs —, 631
W —, 631, 632
Z —, 631, 632
Boundary conditions, 198–200
Bra, 14, 24, 754
D-shaped —, 483
Brillouin constant, 557
Brownian motion, 710, 711
Bulk motion, 415

C
Canonical conjugate, 57
Cartesian coordinates, *see* Coordinates
Category, 727, 728
— theory, 421, 717, 719
2 —, 751
Dagger —, 752
Cauchy–Schwarz inequality, 81
Causal constraint, 341, 343, 549, 550, 585
Causality, 191, 193, 583
– according to SR, 171, 172
– as Nature's laws, 194, 195
– as different from Nature's laws, 197, 199, 205
Causal mechanism, 195
Mechanical —, 193, 343
Principle of —, 582–586
Generalised —, 655
Principle of causal closure, 160, 188, 249, 345
Top-down —, 791
Cavity, 607
Chain rule of derivatives, 78
Channel, 316, 518
Classical —, 315, 316
Noisy —, 519
Quantum —, 316, 519
Characteristic function
Classical —, 531
Quantum —, 531
Chirality, 631
Chiral representation, 622
Chromodynamics, 627
CHSH inequality, 186, 303–306, 324
Church–Turing thesis, 367, 372
Circular polarisation, 39
Class, 718, 719
Scheme of —, 779
Classical limit, 108, 288, 539, 543
Clebsch–Gordan coefficients, 101, 102
Clifford group, 393, 623
Cloning
No — theorem, 132, 317, 499–501, 512
Violation of unitarity, 499–501
CNOT gate, 376, 377
Cobordism, 752, 753
Code, 267, 791
— alphabet, 267
Plurality of —s, 267, 500
— coupling, 330, 331, 451, 500
—words, 267
Linear decoding, 517, 518
Linear encoding, 267, 270, 517, 518

Efficient —, 518
Ideal —, 518
Superdense coding, 447
Codification
— as a local procedure, 330, 451, 500
Coding
Superdense —, 447
Coequaliser, 726
Colimit, 739–742
Collapse determinism
No —, 549, 573, 793
Colour, 627–630
— as unobservable, 628
— hypercharge, 628
Commutation relation, 62–64
Angular momentum —, 70
Annihilation–creation operators —, 523
Dirac matrices —, 623
Field–momentum —-, 614
Position–momentum —, 62–64
Quadratures —, 528
Spin —, 97
Total angular momentum —, 100
Commutativity, 61, 62
Commutator, 63
—'s properties, 64, 65
Complementarity, 108
— experiment, 201, 425, 454, 455
— localisation–causality, 195
— principle, 13, 108–110, 159, 199, 246
— quantum discord–Holevo quantity, 433, 435
Wave–particle —, 109, 435, 454
Which–path/visibility —, 437–440, 454–456
Completeness relation, 36, 37
— for momentum eigenstates, 49
— for position eigenstates, 44
Complexity, 696
Compton effect, 7, 9, 11
Computation time, 401, 402
Concept, 802
Configuration space, 26
Confluent hypergeometric function, 683
Conjunction, 276, 758
Continuity, 7
— principle, 7, 8, 151, 192
Continuity equation, 289
Continuum, 8
Physical —, 8
Contravariant quadrivector, 181
Control, 788
— needs a model, 791
Convection, 415
Convexity, 113
Cook's theorem, 368
Coordinates
Cartesian —, 28
Spherical —, 76, 77
Spheroidal —, 417, 692
Copenhagen interpretation, 109, 221, 224, 246, 248, 278, 279, 286, 352, 417, 422
Coproduct, 724
Corpuscular behaviour, 12, 109
Correlation, 185, 650
—s as independent of space-time, 441, 654
–s as static relations, 185, 186
Classical —, 414, 415
— as fragile, 551
Correspondence principle, 105, 106, 110, 246, 349
Cosmological constant, 646
Cospan, 724
Coulomb law, 602, 659
Counterfactual statement, 227
Coupling, 331, 378
— function, 217
Covalent bond, 701
Covariance, 184
Covariant quadrivector, 181
CPT theorem, 93, 103, 198, 632
Creation operator, 521–523, 610
Cross product, 66
Properties of —, 67
Curl, 601

D

Dalembertian, 178
Dark matter, 670
Darwinian selection, 235, 797
Darwinism, 196, 795
Data, 554
— state, 504
Decoherence, 235–237, 239–241, 258, 565
— time, 243, 244
Small scale of —, 259
Spontaneous —, 237, 371, 411
Decoherent histories, 239
Deduction, 284, 806, 807, 811, 812
— as anterograde inference, 816
Delayed choice experiment, 202–204, 422, 423, 425
Deleting

No–deleting theorem, 501–503, 512
De Morgan's law, 295
Density matrix, 110
Off-diagonal elements of –, 158
Properties of —, 115
Reduced —, 120
Sphere of density matrices, 115
Detection, 152, 161, 265, 266, 378
– as not unitary, 154
— as an outcome selection, 479, 486, 488
— as answer to interrogation, 460
— as not linear, 241
— as not unitary, 152, 153, 220
— outcome, 153
— as ultimate fact, 581
Detector, 152
— as decider, 578, 579
— as even–selection device, 460
Determinism, 193
Classical —, 154, 194, 227, 285, 343, 346
Quantum —, 154, 196, 197, 507
Deutsch algorithm, 385
Deutsch problem, 385
Diffeomorphism, 660
Differential, 19, 45
Diffusion equation, 49
Dipole, 700
Dirac algebra, 623
Dirac delta function, 46, 58
Dirac equation, 622, 623
Dirac matrices, 623
Dirac picture, 139
Dirac's notation, 13, 14
Direct product, 94, 116
Direct sum, 94
Discriminability
Local —, 587–589
Perfect —, 587
Discrimination, 799
Disjunction
Exclusive —, 276
Inclusive —, 276, 758
Displacement operator, 530, 531, 533
Divergence, 52, 600
Dot product, 27
D-shaped circuit, 558, 559
d-state, 80, 677
Dynamical variables, 3

E
Eberhard's theorem, 321–324
Ebit, 316, 317, 344
Effect, 150, 461, 481
— operator, 462–464
Ehrenfest theorem, 108
Eigenfunction
Non–normalisable —, 57, 58
Eigenvalue
Degenerate —, 93, 94
Eigenvalue equation, 43
Einstein's box, 646, 647, 649
Einstein's clock–shutter device, 277, 646, 649
Einstein's field equations, 646
Einstein's locality, 324
Einstein's summation convention, 180, 181
Electric field
— operator, 608
— vector, 602, 603
Dimensionless —, 528
Electric permittivity in vacuum, 603, 681
Electromagnetic field
Classical —, 4
Quantisation of —, 610
Electromagnetic force, 103
Electron, 103, 679
— positron annihilation, 632
— positron generation, 632
Electronegativity, 701, 703
Embodied cognition, 799, 802
Emergence, 248, 674, 675, 695–697, 742, 743, 787
Emergent monism, 696
Empty wave, 287
Energy, 2, 557
— conservation, 198
— density, 614, 616, 619
— efficiency, 787
— eigenvalue equation, 59
— spectrum, 60
Bound state, 61, 676, 677
Ground state of —, 61, 682
Kinetic —, 16, 19
Potential —, 16
Relativistic —, 177
Rest —, 177
Total —, 15
Unbound state, 61
Energy–momentum quadrivector, 181
Entanglement, 157, 253, 282, 338
— as a reduction of possible outcomes, 217, 260
— as mutual information, 253
— as quantum–classical correlation, 415, 416, 431, 519

— of outcomes, 330
— swapping, 308, 572–574, 576
— exchanges coupling and detection, 576
— without signals, 283, 308, 311
Conditional —, 337, 338
Degrees of —, 254
Entropic quantum gravity, 666–670, 672
Entropy
— as relative, 513
— gradient, 668
Bekenstein–Hawking —, 667, 670
Boltzmann —, 259, 707–709
Conditional —, 430
Negative —, 317, 515
Quantum —, 431
Local increase of —, 242, 257, 552
Local shift of —, 257, 266, 513
Shannon —, 250, 251
Sub–additive property of —, 252, 515
Strong —, 257
von Neumann —, 250, 251
— of observables, 254, 255, 514
Environment, 235, 236, 465–467
Umwelt, 242
EPR argument, 165, 275
Logical restatement of —, 275–277
EPR experiment, 166, 167
EPR state, 312, 425
EPRB experiment, 167–170, 301
Equaliser, 305, 726
Equivalence
— class, 318, 319
Functional —, 791, 792
Logical —, 759
Material —, 721, 759
Path —, 722, 735
Equivalence principle
Strong —, 637
Reformulation of the —, 673
Weak —, 637
Equivocation, 216, 430
Error, 520
— correction, 520, 560–564, 566, 567, 813, 817
Euclidean space, 13
Euler angles, 68
Euler formulas, 57
Euler identity, 381
Euler–Lagrange equation, 21, 615
Event, 140
— as dynamical cross, 488
— as outcome, 412
— as source of disturbance, 524
— as source of effects, 413
— as source of signals, 409, 548, 550
— coincidence, 407, 408
— generating a stationary state, 658
— horizon, 667
—'s randomness, 408, 409, 412
—s as connected by Lorentz transformations, 654
—s as local, 247
—s as tokens, 340
Controlled —, 409
Detection –, 196, 206
Detection —, 247
Observed —, 204, 221
Probabilistic —, 407, 408
Quantum —, 408, 411–413, 436, 442, 448, 548, 651
Relativistic —, 407, 408, 442, 549
Uncontrolled, 411
Experience, 406
— as choice, 221
Controlled —, 406
Direct —, 418
Indirect —, 420–422
Ordinary —, 406
Explanation, 130
Exponential function, 17, 18
Extension, 806

F

Factorial, 17
Factorisation, 387
FAPP explanations, 242
Faraday's law, 603
Feature, *see* Quantum feature
Fermi–Dirac statistics, 102, 105
Fermion, 102
Feynman's diagram, 632, 633, 655, 661, 752
Feynman's notation, 624
Fibre
— bundle, 178, 179
— product, 723
— sum, 725
Fidelity, 564
Filter, 64
Fine structure constant, 634
Flavour, 627, 628, 630
Fock space, 611
Force carrier, 103
Form, 338
Formal cause, 341

Formal system, 367
Fourier's analysis, 57
Fourier transform, 56, 57
  Inverse —, 56, 57
Free falling body, 637
Frequency, 3
  Angular —, 57, 468
f–state, 80, 677
Functor, 736–738
  Adjoint —, 742, 743
    Forgetful —, 742, 743, 817
    Free —, 742, 743, 817
  Representable —, 743, 744

G
Galilei transformations, 172, 173
Gauge condition, 605
  Coulomb —, 606, 608
  Lorentz —, 618
Gauge symmetry, 604, 605
Gauge transformation, 605
Gaussian curvature, 639, 640
Geodesic, 639–641
  — as intrinsic, 639, 640
  — deviation, 639
  — plane, 640
  Null —, 638
  Timelike —, 638
Geometric phase, 428, 429
Gestalt school, 221
GHSZ state, 334, 336–338, 378, 565
Global–local entrenchment, 235, 244, 247, 498, 587
Gluon, 628–630
Gradient, 51
Graph, 734, 735
Gravitation, 189
Gravitational constant, 249, 635
Gravitational contraction, 638
Gravitational coupling constant, 649
Gravitational waves, 646
Greenberger–Yasin equality, 439, 440
Green's function, 537–539, 660
  Free —, 538
Green's theorem, 600
Group, 90, 91, 733
  Orthogonal —, 91
  Permutation —, 91
  Rotation —, 91
  Unitary —, 91
    Special —, 91
Group velocity, 83
Grover's algorithm, 389–392
Grover's operator, 390

H
Hadamard gate, 373, 374
Hadamard operator, 143
Hadron, 628
Halting problem, 367
Hamiltonian, 15
  — of the electromagnetic field, 610
  — operator, 15
  1D —, 15, 16
  4D —, 663
  Classical —, 22
  Dirac —, 622
  Displaced —, 535
  Free —, 139
  Interaction —, 139, 217, 459, 650
Hamilton–Jacobi equation, 288, 539
Hamilton's equations, 23
Harmonic oscillator, 521, 522, 610
Hawking radiation, 671
Heisenberg equation, 136
Heisenberg picture, 133–137
  — for computation, 392–395
  Generalised —, 486
Heisenberg's experiment, 85, 86
Helmholtz decomposition, 601
Helmholtz's theorem, 601
Hidden variables, 171, 286, 291–294, 300, 301
  — theory, 285, 286, 292, 300
  Contextual —, 286
Hilbert space, 13, 26, 31
  Complex —, 13, 588
  Real —, 588
Holevo bound, 265, 328
Holevo quantity, 266, 431–433, 435
Holevo theorem, 265, 266, 270
Holographic principle, 666, 671
Holonomy, 642, 659
Homogeneity, 16
Homonuclear molecule, 693
Hopf ring, 338
Huygens' principle, 537, 550
Hydrogen atom, 675–687
  Eigenfunctions of —, 684–687
  Energy eigenvalues of —, 684
  Ground state of —, 684
Hydrogen molecular ion, 692
Hydrogenoid atom, 684, 686, 687
Hyperbolic trigonometric functions, 183
Hyper-correlations, 328

I
Idealism
Objective —, 161
Strong —, 222, 225, 411
Subjective —, 161
Weak —, 223
Idempotency, 35
Identity operator, 16, 36
Imaginary part, 14
Implicate order, 434, 435
Implication
Logical —, 275, 759
Material —, 759
Improvisional intelligence, 801, 802
Incertitude
Objective, 412
Subjective, 412
Incompleteness
— of QM, 165
Induction, 157, 285, 807, 808, 811, 812
— as Bayesian inference, 812
— as retrograde inference, 816
Inertial motion, 173
Inference, 418, 421, 555, 579, 804, 805, 812
Bayesian —, 554, 801, 805
Inverse —, 420, 554
Circle of —s, 814–816
Metaphoric —, 421, 422
Metonymic —, 421, 422
Natural —, 805
Scientific —, 805
Information, 806
— about the context, 290
— acquiring, 249, 266, 270, 332, 368, 554, 578
— as irreversible, 459
— as selecting already actionable information, 577
Information acquisition principle, 555
No — without disturbance, 520, 524–526
— as a second–order equivalence class, 575
— as a type, 318
— as combinatorics, 261
— as connection among events, 261
— as dimensionless, 249
— as formal, 249, 333, 513
— as independent of the physical context, 270
— as intrinsic, 498, 513
— as objective, 261, 270
— as self–similar, 503
— causality principle, 328, 329, 333, 447
— compression, 271–273, 275, 367
Ideal —, 588, 589
— deficit, 432
— extrapolation, 163, 227, 235, 247, 553
— is conserved, 503
— processing, 370, 574, 575
— selection, 370, 434, 553, 574, 586
— as reduction of incertitude, 580, 581
— costs energy, 557
— swap, 258, 459, 499, 502, 503, 513
— triad, 579
—'s causal power, 339
Actionable —, 348, 423
Active —, 289, 311, 488
Informational completeness, 527, 528
Linear transfer of —, 240, 241, 330, 575, 578
Locally accessible —, 242, 257, 266, 268, 499, 510
Holevo bound on —, 265, 272
Locally inaccessible —, 269, 510
Mutual —, 240, 250, 252–254, 370, 429, 430, 574, 575
Conditional —, 514
Independence of — from signalling, 253
Potential —, 266, 268, 317, 454
— as infinite, 263, 577
Quantum of —, 556
Initial conditions, 199, 200
Insight, 809
Instantiation principle, 780, 781
Instrumentalism, 224, 348
Integration by parts, 20
Intension, 806
Interaction
Quantum of —, 556
Interaction time, 218
Interactionism, 86, 282, 348, 489
Interference, 9, 11, 416
— terms, 147, 149, 156, 199, 216, 219, 236, 269
— as intrinsic, 226
— visibility, 438, 439, 453
— observable, 475–477
Constructive —, 11
Destruction of —, 459
Destructive —, 11
Fourth-order —, 452
Self–interference, 25, 415

Interferometer
Gravitational–waves —, 646, 647
Mach–Zehnder —, 142, 436
Michelson–Morley –, 174
Internuclear axis, 700
Invalidating feedback, 797
Invariance, 90, 184
Ionic bond, 701
Isomer molecules, 704
Isospin, 628
Itinerant dynamics, 697

J
Jacobi identity, 65

K
Kantian epistemology, 221, 222
Kepler's problem, 95
Ket, 13, 24, 754
D–shaped —, 483
Killing vector, 669, 670
Kinematic variables, 3
Klein–Gordon equation, 616, 617, 624
Spin–0 particle —, 617
Spin–1 particle —, 618
Knowing/recognising, 799
Kochen–Specker graph
10-node —, 296
117-node —, 297
Kochen–Specker theorem, 295–299, 480
Kraus condition, 584
Kraus theorem, 485, 486
Kronecker delta, 29, 34

L
Lagrangian
— density, 614
Classical —, 19, 539
Proca —, 619
Legendre polynomials, 78
Associated —, 79
Leibniz's product rule, 62
Lepton, 103
Levi–Civita tensor, 69, 70
Lie algebra, 91
Lie bracket, *see* Commutator
Lie derivative, 670
Light, 4
— propagation, 609
— as generation of the space-time manifold, 656
Light cone, 179, 180, 654
Light like interval, 179
Limit, 739, 740
Lindblad superoperator, 467, 468, 544
Linear Combination of Atomic Orbitals (LCAO), 690–693, 700
Linearity, 16, 32
Liouville equation, 138, 139
Liouville theorem, 139, 467
List, 729, 730
Local inertial frame, 637
Localisation, 195, 198, 408, 409
Locality, 171, 586
Einstein's —, 174, 185, 186, 312, 317
LOCC protocol, 333
Logarithm, 251
Logarithm function
Natural —, 17
Logical coherence, 800
Logical ID, 762
Logic gate, 372
Loop, 659
— state, 659
Loop Quantum Gravity, 658–663, 665, 666, 670, 672
Lorentz force law, 645
Lorentz group, 183, 184, 620
Lorentz invariance, 182, 612, 613
Lorentz transformations, 174–176, 180–184, 612
Boosts, 183
Rotations, 182
Lowering operator, 468, 521

M
Macroscopic world
— as emergent, 246
Magnetic field
— operator, 608
— vector, 602
Dimensionless —, 528
Magnetic induction, 602
Magnetic permeability, 602, 603
Malament's theorem, 657
Manifold, 178, 660
Differentiable —, 179
Many-World interpretation, 228, 230–235, 417, 487, 498, 781
Map, 374, 719, 720
— composition, 720
Horizontal —, 749, 750
Vertical —, 749, 750

Bijective —, 720
Injective —, 560, 720
Surjective —, 720
Markov blanket, 793, 794
Markovian sequence, 548
Mass
— as energy store, 673
— defect, 635, 681
Equivalence — energy, 8
Gravitational —, 636, 673
Inertial —, 636, 673
Reduced —, 681
Rest —, 177
Master equation, 468–470, 473, 474
Lindblad —, 467–469, 544
Matrix, 32
— product, 32
— sum, 32
Matter, 8
— divisibility, 8, 9
— energy, 177
— waves, 10, 11, 15
Maxwell demon, 271, 513
Maxwell equations, 177, 602
First Maxwell equation, 602
Fourth Maxwell equation, 604
Second Maxwell equation, 602
Third Maxwell equation, 604
Maxwell tensor, 618, 645
Measurement, 36, 152, 464, 576, 812, 813
— error, 7, 269
— outcome, 158
—'s channels, 573
Ideal —, 525
Indirect —, 279, 561
Interaction-free —, 87, 144–147, 279, 348
Interaction-free—, 87
Joint measurability of observables, 63, 150, 154
Unsharp —, 463
Meson, 628–630
Metamorphosis, 435
Metre, 152, 217, 219
Metric tensor, 181
Michelson–Morley experiment, 173, 174
Microcanonical ensemble, 705
Mind, 188, 221, 222
— body relation, 188
Universal —, 235
Minkowskian space-time, 178
Minkowski metric, 181
Minkowski signature, 181
Mirror neuron, 801
Modal interpretation, 31
Modulo–2 addition, 325
Molecular polarity, 700
Momentum, 5, 46
— eigenvalue equation, 49
— representation, 49
1D — eigenfunction, 49
1D — observable, 48
3D — eigenfunction, 53
3D — eigenstate, 52
Relativistic –, 177
Monad, 9, 188, 190, 192, 556, 671
Monoid, 729
Free —, 730
Morphism, 719
Epimorphism, 721
Homomorphism, 732, 733
Isomorphism, 721, 759
Multipath dynamics, 154, 196, 546, 654
Multisimultaneity, 440
Muon, 103

N

Nahewirkung, 187, 345, 346, 612
Natural transformation, 745–747
Nature's laws, 130
— as based on symmetries, 198, 551, 780
— as probabilistic, 196, 197
— as relations among observables, 198
Neg-reversal operation, 763
Neutrino, 103
Neutron, 105, 630, 675, 676
Newton's second law, 21
Noether's theorem, 88, 95
Non-locality, 131, 171, 321
Normal ordering, 523
Normalisation
— condition, 26
No-signalling requirement, 328, 511
Number operator, 521, 522, 610
Number states, 522

O

Object
Ontological meaning, 130, 497
Objectivism, 130, 222
Quantum —, 248, 264, 265, 270, 281, 284, 411, 417, 487, 491
Objectivist interpretation, 159–161, 226, 227, 248
Observable, 31, 37, 150, 292

— as equivalence class of premeasurements, 459, 460, 575
— as type–token, 460
— change, 132
—'s expectation value, 38, 116, 465
—'s selection, 230, 241
— as a local choice, 242
—s Eigenvalue, 37
Complete —, 664
Dispersion-free —, 291, 292
Dynamical —, 88
Kinematic —, 88
Partial —, 661, 664, 666
Sharp —, 527
Unsharp —, 474, 477–479, 527, 528
Observer, 156, 161, 220, 227, 407, 409
—'s mind, 159, 160, 227, 443
Occasionalism, 160, 188
Occupation number, 610
Olog, 722
Omnimoda determinatio, 25, 87, 149, 165, 193, 197, 277, 285, 291, 346, 549
Omnimoda determination, 26
1-form, 53, 54
Ontological status of physical entities, 131
Ontology
Interpreted —, 819–822
Uninterpreted —, 822–824
Operation, 481, 484, 485
Identical —s, 560
Trace–preserving —, 482, 562
Operationalism, 352
Process —, 486
Operationalist epistemology, 349
Operator, 31
Bounded —, 38, 60
Hermitian —, 37, 90
Laplacian —, 52
— in spherical coordinates, 678, 679
Nabla —, 51
Positive semidefinite —, 34
Super —, 462, 482, 483
Symmetric —, 38, 60
Oracle, 389
Overlapping integral, 691
Oxygen, 700

P

Parallel transport, 428, 429, 641, 642
Parametric–down–conversion, 306, 307
Parity, 93, 183
Partial derivative, 20
Particle
Antisymmetric —, 103–105
Extended —, 13
Symmetric —, 102–104
Test —, 636
Virtual —, 626, 633
Partition function
Classical —, 706
Quantum —, 709
Passion–at–a–distance, 311, 348
Passive transformation, 133, 134
Path integrals, 539–542, 547, 660, 661, 821
Path predictability, 438
Pauli matrices, 97, 100, 314, 374
Pauli problem, 527
Pauli's exclusion principle, 95, 105, 110
Pauli-X gate, 374
Pauli-Y gate, 374, 381
Pauli-Z gate, 374, 380
PBR basis, 427
PBR theorem, 426, 427
Periodic table of elements, 688, 689
Phase, 3, 11
— shift gate, 372
— shifter, 142
Geometric —, 290
Phase space, 85, 87, 138
Phase transition, 699
Phenomenism
Epistemic —, 260, 279, 411, 426, 489, 490
Quantum —, 221, 350, 491
Photoelectric effect, 3–5
Photon, 4, 5, 7, 610
Idler —, 287, 307
Signal —, 287, 307
Physical theory, 130
Planck area, 667
Planck equation, 2, 107
Planck length, 249
Planck mass, 249
Planck's constant, 2, 7, 556, 557
— as quantum of action, 2
Reduced —, 15
Planck scale, 441
Planck time, 249
Poincaré group, 184, 612, 626
Poincaré sphere, 263, 373, 380, 465, 470
Poincaré's recurrence theorem, 552
Poincaré transformation, 182
Pointer, 152
Poisson's brackets, 65
Polar decomposition, 463

Polarisation vector, 608
Population inversion, 470
POSet, 728, 758
Position, 44–46
— eigenfunction, 46
— eigenstate, 44
— eigenvalue equation, 45
— representation, 46
1D — observable, 44
3D — observable, 51
Positive operator valued measure, 462, 463, 474, 477–479, 481, 527, 528
Positron, 103
Possibility, 340, 347
— as pre–determined, 487
Potential
Centrifugal —, 677
— barrier, 681
Coulomb —, 676, 677, 681
Potentiality, 340, 346–348, 350, 351
— as post–determined, 487
— as real resource, 347
PR box, 325, 326, 329
Prediction, 130
Conditional —, 226
Pre–established harmony, 188
Premeasurement, 216, 219, 228, 230, 236, 265, 317, 430, 447, 450, 451, 461
— as information availability, 281
— as interrogation, 450
— as observable selection, 231, 241, 486
— as physical context, 266, 268, 270, 550
— as reversible, 230, 452–458
— without interaction, 459
Preorder, 728, 729
Preparation, 132, 152, 264, 446–450, 558, 583
— as coarse-grained, 449
— as determinative measurement, 446
— as not unitary, 447, 448
— as state selection, 446, 486
Independence of —, 447, 583
Probabilistic theory, 27
Probability, 27
— amplitude, 26
— as real, 149
— as irreducible, 226
— as objective, 149, 261, 412, 413
— density, 45, 60
— in the continuous case, 45
Classical —, 140, 141
Conditional —, 320
Joint —, 140
Kolmogorov's axioms, 140
Quantum —, 141, 142, 147, 149
Sample space, 140
Probability amplitude
— damping, 473, 474
Amplitude operator, 462, 467, 473, 474, 485, 571, 572
Product, 723
Program
— state, 504
No–programming theorem, 504–507, 512
von Neumann's program paradigm, 504
Projection, 35, 158
— postulate, 158, 163, 479, 481
Projector, 34–36
—'s properties, 34, 35
Propagation vector, 53, 68, 607
Propagator, 537, 634, 654
Property, 131, 150, 461, 819
— ascription, 246, 479
— as equivalence class of detections, 479, 575
Proton, 627, 630, 675, 676
p-state, 80, 677, 678
Psycho–physical parallelism, 160
Pullback, 724
Purification, 123, 590–592
Pushout, 725
Pythagorean trigonometric identity, 144

Q

Quadrature, 528, 531
Quantisation principle, 31, 61, 84, 110
Quantum cryptography, 317–319
BB84 protocol, 317–319
Quantum Darwinism, 434, 779
Quantum discord, 429, 431–435, 515
Quantum dynamics
– as evolution of probability amplitudes, 139
Quantum eraser, 454, 456–458, 502, 503
Quantum features, 414–418, 423, 426, 428, 431, 436, 439, 515, 590
— as intrinsic correlations, 416
— as self–correlations, 416, 547
Quantum Field Theory (QFT), 612–617, 619–626
Locality of —, 612
Quantum jump, 543, 544, 546
— superoperator, 544
Quantum-mechanical laws, 130, 247

Quantumness, 431, 432
Quantum number
Azimuthal —, 72
Magnetic —, 72
Principal —, 678
Spin —, 96
Quantum of cost, 557
Quantum postulate, 5, 9, 31
Quantum potential, 287–289, 417
Quantum tomography, 532, 536, 588
Quantum trajectory, 543, 544, 821
Quantum volume, 372
Quark, 103, 627–630
Qubit, 264, 265, 317
Control —, 376
Target —, 376

R
Raising operator, 469, 521
Random Access Code, 329
Randomness, 367, 548
Algorithmic —, 367
Random variable
Classical —, 531, 532
Rayleigh–Jeans formula, 107
Realism, 281, 284
Classical —, 171, 226, 227, 299, 489
Dynamical —, 299, 491
Realist
Classical —, 286
Reality, 422
— as a construction of mind, 223, 410
— as correction, 224
— as interaction, 223
— criterion, 165, 166, 171, 194, 275, 277, 278
— of observables, 278, 280
— of properties, 278
Element of –, 165
Fabric of —, 234, 497, 779
No independent —, 222
Real part, 14
Recovery operation, 519, 561
Redshift factor, 648, 669
Reductionism, 186, 187, 417
Local —, 587
Relational–interactional interpretation, 487, 655, 661, 723, 820
Relativistic constant, 173, 556
Relativity
General —, 190, 634–646, 648
Special —, 173–185
— as a theory of signal exchanges, 654
Renormalisation, 626, 633
Representation, 826
Reservoir, 243
Resolution of the identity, *see* Completeness relation
Resonance integral, 691
Resonator, 3
Retro–causation, 205, 550
Reversal operation, 763
Ricci tensor, 646
Riemann tensor, 644, 645
Right-hand rule, 66, 67
Rotation
— operator, 68
—s generator, 68, 91, 183, 184, 620

S
Scalar potential, 604
Scalar product, 27, 28
— of Euclidean vectors, *see* Dot product
Generalised —, 483
Properties of —, 28
Relativistic –, 182
Schmidt decomposition, 120–123, 509
Theorem, 122
Schmidt number, 121
Schrödinger cat, 162–164, 216, 227, 244–246
Schrödinger equation, 15, 16, 135, 136
— as deterministic, 27
— as first–order, 538
— for a free particle, 49
— in spherical coordinates, 287
1D —, 15, 16
3D —, 52
Linearity of —, 16
Reversibility of —, 16
Stationary —, 59, 60
Time–dependent —, 59, 60
Schrödinger–Pauli equation, 620, 625, 626
Schrödinger picture, 133–137
Generalised —, 486
Schwarzschild radius, 250
Scientific theory
— as complete, 164
— as correct, 164
Second law, 552
Second quantisation, 606–610
Selection
— as reduction of possibilties, 552

Selenocysteine, 790
Selenoprotein, 790
Self–dual variable, 764
Semiotic mechanism, 795, 796, 813
Separability, 117, 163, 227, 328
— condition, 165, 166, 170, 171, 186, 275, 276, 283, 300, 301, 336
Set, 718–720, 728
— theory, 718, 719
Empty —, 761
Universal —, 761
S gate, 379, 380
Shor's theorem, 387–389
Signal, 172, 517
—'s irreversibility, 550
—'s pattern, 517
—'s quantum–classical dynamics, 546
—'s reception, 548–550
—'s source as singularity, 555, 657
—'s spreading, 519, 550
— generate noise, 519
—'s structure, 517
Singular limit, 698, 699, 711
Smooth function, 605
Solipsism, 225, 228
Source determinism, 548
No —, 549, 573
Space
— homogeneity, 176
— isotropy, 176
— translation, 46
— generator, 48
— operator, *see* Unitary operator
Absolute —, 189
Affine —, 178
Euclidean —, 178
Grain of —, 659
Space-like interval, 180
Space-time, 178
— as emergent, 441
— distance, 178
— isotropy, 611
— quadrivector, 180
Span, 723
Spectral decomposition, 37, 44
Spectral theorem, 37
Spectrum, 31
Continuous —, 37, 57, 61
Discrete —, 37, 61
Spherical coordinates, *see* Coordinates
Spherical harmonics, 78–80
Spin, 95, 96, 678, 688
— down, 97
— foam, 661
— network, 659, 660
— observable, 96, 100
— up, 97
Spinor, 98, 104, 623
Left —, 621
Right —, 621
Weyl —, 622
Spontaneous emission, 468, 471, 474
Square–Not gate, 381, 382
s-state, 80, 677, 678
Standard model, 612, 627, 631
State, 132, 150
— as equivalence class of preparations, 450, 575
— change, 132
— mapping local environment, 651
— of space, 659, 660
— reconstruction, 532, 535, 536
— vector, 13, 25
— as mathematical tool, 221
Classical —, 151
Coherent —, 245, 528–531
Complete mixed —, 115
Dispersion–free —, 291–294
Distance between —s, 245
Eigen—, 43
Entangled —, 119, 219
Maximally —, 119, 254
Excited —, 468
Global —, 147
Ground —, 468
Input —, 465
Marginal —, 120
Mixed
Classical —, 158
Lüders —, 219, 236
Mixed —, 112, 220
Maximally —, 113, 252
Non–commuting —, 501
Non–orthogonal —s, 132
Output —, 465
Product —, 118
Pure —, 111, 112
— has zero entropy, 251
Purity, 115
Relative —, 226, 229, 231
Singlet —, 119, 168
Stationary —, 59, 198
Statement (logic)
Atomic —, 295
Statistical interpretation, 86, 226
Statistical mechanics

Classical —, 543, 704–706
Quantum —, 706–709
Steering, 561
Stern–Gerlach experiment, 95, 96
Stokes' law, 710
Stokes' theorem, 601
Stone's theorem, 90, 375
Stress–energy tensor, 616, 646
Strong force, 103, 192, 627, 628
Residual —, 676
Subjectivist interpretation, 159–161, 216, 219, 220, 222, 223, 225, 248, 249, 417, 422
Sufficient reason
Principle of —, 193, 423, 818
Superluminal communication, 321
Superposition, 24, 149
— as intrinsic, 415
— as relative, 39
— principle, 24, 88, 110
— state, 24, 39
Surprisal, 250, 797
Swap gate, 379
Symbolic system, 802, 803, 813
Symmetry, 90, 91
Continuous —, 89
Discrete —, 93
Synchronisation, 407, 408
System
— theory, 489
Chaotic —, 200
Classical —
—s cannot be purified, 552
Complex —, 200
Composite —, 116
Free —, 61, 262
Object — as information processor, 578, 579
Open —, 237, 411
Physical —, 187, 189, 489, 498
— as passible of the 3 information operations, 578
Quantum —s as information processors, 365, 449
Quantum —s as self-similar, 264, 270
Quantum —s as sensitive to the environment, 487, 503, 650, 825
Quantum —s mapping environmental information, 488

T

Tachyon, 179
Tau lepton, 103
Taylor series, 18
Teleportation, 311–317, 567–574, 576, 756
— as reversed measurement, 576
— of logic gates, 395, 397, 398, 400, 401
Tensor, 70, 71
Test, 131, 484
Observation —, 559
Completeness of —, 584
Preparation —, 559
Completeness of —, 584
T gate, 379, 380
Thermal conduction, 414
Tidal force, 637–639, 673
Time
— asymmetry, 205, 206, 443
— dilation, 176
Gravitational —, 639
— homogeneity, 176
— reversal, 183
Absolute —, 173
Proper —, 407, 441, 638
Relative —, 176
Time-like interval, 179
Toffoli gate, 370, 371
Quantum —, 382–384
Token, 262, 318, 339, 340, 344
Trace, 114
Cyclic property of —, 114
Partial —, 120, 235, 236
Transactional interpretation, 441, 487, 655
Transmissivity, 453
Triangle inequality, 82
Triorthogonal decomposition, 241, 242, 330
Tsirelson's bound, 327, 329
Tsirelson's theorem, 326, 327
Turing machine, 366, 367, 372, 731
Twins paradox, 184
2-form, 54, 55
Type, 262, 318, 339, 340, 344
Type–token, 339, 340, 344

U

Uncertainty, 81
— relation, 84, 87, 88, 110, 156
— between quadratures, 528
Energy–time —, 83, 84, 648, 649
Entropic —, 255, 256
Generalised —, 82
Position–momentum –, 154
Position–momentum —, 80, 83–85
Undulatory behaviour, 12, 109, 139

Unitary operator, 16
Momentum translation —, 89
Parity —, 93, 533
Displaced —, 534
Permutation —, 103, 104, 273
Rotation —, 67, 89, 373
Space translation —, 47, 88, 89
Spin Rotation —, 99
Time translation —, 16, 17, 91, 92
Unitary transformations
— as entropy preserving, 251
Universal Darwinism, 696
Universal entrenchment, 189, 247, 253
Universe
Partecipatory —, 443–445
Universe in a stationary state, 658
Universe's fragment, 432–434
Unruh effect, 633, 667

V
Vacuum, 611
— decay, 633
— state, 530, 531, 611
Valence Bond Method (VBM), 694, 695
Variance, 80
Variational method, 690
Variety, 270
Vector
— space, 13, 31
Logical —, 771
—'s magnitude, 28
—'s norm, 28, 30
Axial —, 66
Normalised —, 30
Polar —, 66
Vector potential, 602, 604
Vector product, *see* Cross product
Venn diagram, 430
Viscosity, 415
Volume contraction, 176
von Neumann equation, 137, 138
von Neumann formula, 220
von Neumann's cut, 160, 225, 246, 479
von Neumann's problem, 156, 219, 231

W
Water molecule, 700, 702
Wave
Advanced —, 655
Empty —, 286, 287
Pilot —, 286, 287, 289, 488
Plane —, 53
Spherical —, 53
Wave equation
Classical —, 606
Wave function
— in momentum space, 50
— in position space, 44
Covalent —, 695
Ionic —, 695
Multiversal —, 228
Wavelength, 3
Compton —, 11, 250
de Broglie thermal —, 243
Wave packet, 11, 26, 83
Reduction of the —, 161, 477
Wave–particle dualism, 12, 13, 109
Wavicle, 13, 439
Weak force, 103, 627, 631, 632
Wheeler–DeWitt equation, 663, 664
Which path
— detectors, 458
— observable, 477
Wien's displacement law, 556
Wigner function, 528, 532–536
Wigner's friend, 225
Wigner's theorem, 93
World-picture, 636

X
X gate, *see* Pauli-X gate
X rays, 7

Y
Y gate, *see* Pauli-Y gate
Young's experiment, 9, 10, 24, 201

Z
Z gate, *see* Pauli-Z gate