

FRANCK LALOË

Do We Really Understand Quantum Mechanics?



CAMBRIDGE

CAMBRIDGE

more information - www.cambridge.org/9781107025011

DO WE REALLY UNDERSTAND QUANTUM MECHANICS?

Quantum mechanics is a very successful theory that has impacted on many areas of physics, from pure theory to applications. However, it is difficult to interpret, and philosophical contradictions and counter-intuitive results are apparent at a fundamental level. In this book, Laloë presents our current understanding of the theory.

The book explores the basic questions and difficulties that arise with the theory of quantum mechanics. It examines the various interpretations that have been proposed, describing and comparing them and discussing their successes and difficulties. The book is ideal for researchers in physics and mathematics who want to know more about the problems faced in quantum mechanics but who do not have specialist knowledge in the subject. It will also appeal to philosophers of science and scientists who are interested in quantum physics and its peculiarities.

FRANCK LALOË is a Researcher at the National Center for Scientific Research (CNRS) and belongs to the Laboratoire Kastler Brossel at the Ecole Normale Supérieure. He is co-author of *Quantum Mechanics*, with Claude Cohen-Tannoudji and Bernard Diu, one of the best-known textbooks on quantum mechanics.

DO WE REALLY UNDERSTAND QUANTUM MECHANICS?

FRANCK LALOË

Ecole Normale Supérieure

and

National Centre for Scientific Research (CNRS)



CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town,
Singapore, São Paulo, Delhi, Mexico City

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org

Information on this title: www.cambridge.org/9781107025011

© F. Laloë 2012

This publication is in copyright. Subject to statutory exception
and to the provisions of relevant collective licensing agreements,
no reproduction of any part may take place without the written
permission of Cambridge University Press.

First published 2012

Printed in the United Kingdom at the University Press, Cambridge

A catalogue record for this publication is available from the British Library

Library of Congress Cataloguing in Publication data

Laloë, Franck, 1940–

Do we really understand quantum mechanics? / Franck Laloë.
p. cm.

Includes bibliographical references and index.

ISBN 978-1-107-02501-1 (hardback)

1. Quantum theory. 2. Science–Philosophy. I. Title.

QC174.12.L335 2012

530.12–dc23 2012014478

ISBN 978-1-107-02501-1 Hardback

Cambridge University Press has no responsibility for the persistence or
accuracy of URLs for external or third-party internet websites referred to
in this publication, and does not guarantee that any content on such
websites is, or will remain, accurate or appropriate.

Contents

<i>Foreword</i>	<i>page</i> ix
<i>Preface</i>	xi
1 Historical perspective	1
1.1 Three periods	2
1.2 The state vector	7
2 Present situation, remaining conceptual difficulties	17
2.1 Von Neumann's infinite regress/chain	19
2.2 Schrödinger's cat	21
2.3 Wigner's friend	26
2.4 Negative and "interaction-free" measurements	27
2.5 A variety of points of view	31
2.6 Unconvincing arguments	37
3 The theorem of Einstein, Podolsky, and Rosen	38
3.1 A theorem	39
3.2 Of peas, pods, and genes	40
3.3 Transposition to physics	45
4 Bell theorem	56
4.1 Bell inequalities	57
4.2 Various forms of the theorem	66
4.3 Cirelson's theorem	77
4.4 No instantaneous signaling	80
4.5 Impact of the theorem: where do we stand now?	89
5 More theorems	100
5.1 GHZ contradiction	100
5.2 Generalizing GHZ (all or nothing states)	105
5.3 Cabello's inequality	108

5.4	Hardy's impossibilities	111
5.5	Bell–Kochen–Specker theorem: contextuality	114
6	Quantum entanglement	120
6.1	A purely quantum property	121
6.2	Characterizing entanglement	126
6.3	Creating and losing entanglement	133
6.4	Quantum dynamics of a sub-system	142
7	Applications of quantum entanglement	150
7.1	Two theorems	150
7.2	Quantum cryptography	154
7.3	Teleporting a quantum state	160
7.4	Quantum computation and information	163
8	Quantum measurement	168
8.1	Direct measurements	168
8.2	Indirect measurements	176
8.3	Weak and continuous measurements	181
9	Experiments: quantum reduction seen in real time	195
9.1	Single ion in a trap	196
9.2	Single electron in a trap	200
9.3	Measuring the number of photons in a cavity	201
9.4	Spontaneous phase of Bose–Einstein condensates	204
10	Various interpretations	211
10.1	Pragmatism in laboratories	212
10.2	Statistical interpretation	220
10.3	Relational interpretation, relative state vector	222
10.4	Logical, algebraic, and deductive approaches	225
10.5	Veiled reality	230
10.6	Additional (“hidden”) variables	231
10.7	Modal interpretation	261
10.8	Modified Schrödinger dynamics	264
10.9	Transactional interpretation	280
10.10	History interpretation	281
10.11	Everett interpretation	292
10.12	Conclusion	300
11	Annex: Basic mathematical tools of quantum mechanics	304
11.1	General physical system	304
11.2	Grouping several physical systems	316
11.3	Particles in a potential	320

<i>Appendix A</i>	<i>Mental content of the state vector</i>	328
<i>Appendix B</i>	<i>Bell inequalities in non-deterministic local theories</i>	330
<i>Appendix C</i>	<i>An attempt for constructing a “separable” quantum theory (non-deterministic but local)</i>	332
<i>Appendix D</i>	<i>Maximal probability for a state</i>	335
<i>Appendix E</i>	<i>The influence of pair selection</i>	336
<i>Appendix F</i>	<i>Impossibility of superluminal communication</i>	341
<i>Appendix G</i>	<i>Quantum measurements at different times</i>	345
<i>Appendix H</i>	<i>Manipulating and preparing additional variables</i>	350
<i>Appendix I</i>	<i>Correlations in Bohmian theory</i>	353
<i>Appendix J</i>	<i>Models for spontaneous reduction of the state vector</i>	357
<i>Appendix K</i>	<i>Consistent families of histories</i>	362
<i>References</i>		364
<i>Index</i>		390

Foreword¹

Quantum Mechanics is an essential topic in today's physics curriculum at both the undergraduate and graduate levels. Quantum mechanics can explain the microscopic world with fantastic accuracy; the fruits from its insights have created technologies that have revolutionized the world. Computers, lasers, mobile telephones, optical communications are but a few examples. The language of quantum mechanics is now an accepted part of the language of physics and day-to-day usage of this language provides physicists with the intuition that is essential for achieving meaningful results. Nevertheless, most physicists acknowledge that, at least once in their scientific career, they have had difficulties understanding the foundations of quantum theory, perhaps even the impression that a really satisfactory and convincing formulation of the theory is still lacking.

Numerous quantum mechanics textbooks are available for explaining quantum formalism and applying it to understand problems such as the properties of atoms, molecules, liquids, and solids; the interactions between matter and radiation; and more generally to understand the physical world that surrounds us. Other texts are available for elucidating the historical development of this discipline and describing the steps through which it went before quantum mechanics reached its modern formulation. In contrast, books are rare that review the conceptual difficulties of the theory and then provide a comprehensive overview of the various attempts to reformulate quantum mechanics in order to solve these difficulties. The present text by Franck Laloë does precisely this. It introduces and discusses in detail results and concepts such as the Einstein–Podolsky–Rosen theorem, Bell's theorem, and quantum entanglement that clearly illustrate the strange character of quantum behavior. Within the last few decades, impressive experimental progress has made it possible to carry out experiments that the founding fathers of quantum mechanics considered only as “thought experiments”. For instance, it is now possible to follow the

¹ Translated by D. Kleppner.

evolution of a single atom in real time. These experiments are briefly reviewed, providing an updated view of earlier results such as convincing violations of the Bell inequalities.

This book provides a clear and objective presentation of the alternative formulations that have been proposed to replace the traditional “orthodox” theory. The internal logics and consistency of these interpretations is carefully explained so as to provide the reader with a clear view of the formulations and a broad view of the state of the discipline. At a time when research is becoming more and more specialized, I think that it is crucial to keep some time for personal thought, to step back and ask oneself questions about the deep significance of the concepts that we employ routinely. In this text, I see the qualities of clarity, intellectual rigor, and deep analysis that I have always noticed and appreciated in the work of the author during many years of friendly collaboration. I wish the book a well-deserved great success!

Claude Cohen-Tannoudji

Preface

In many ways, quantum mechanics is a surprising theory. It is known to be non-intuitive, and leads to representations of physical phenomena that are very different from what our daily experience could suggest. But it is also very surprising because it creates a big contrast between its triumphs and difficulties.

On the one hand, among all theories, quantum mechanics is probably one of the most successful achievements of science. It was initially invented in the context of atomic physics, but it has now expanded into many domains of physics, giving access to an enormous number of results in optics, solid-state physics, astrophysics, etc. It has actually now become a general method, a frame in which many theories can be developed, for instance to understand the properties of fluids and solids, fields, elementary particles, and leading to a unification of interactions in physics. Its range extends much further than the initial objectives of its inventors and, what is remarkable, this turned out to be possible without changing the general principles of the theory. The applications of quantum mechanics are everywhere in our twenty-first century environment, with all sorts of devices that would have been unthinkable 50 years ago.

On the other hand, conceptually this theory remains relatively fragile because of its delicate interpretation – fortunately, this fragility has little consequence for its efficiency. The reason why difficulties persist is certainly not that physicists have tried to ignore them or put them under the rug! Actually, a large number of interpretations have been proposed over the decades, involving various methods and mathematical techniques. We have a rare situation in the history of sciences: consensus exists concerning a systematic approach to physical phenomena, involving calculation methods having an extraordinary predictive power; nevertheless, almost a century after the introduction of these methods, the same consensus is far from being reached concerning the interpretation of the theory and its foundations. This is reminiscent of the colossus with feet of clay.

The difficulties of quantum mechanics originate from the object it uses to describe physical systems, the state vector $|\Psi\rangle$. While classical mechanics describes a system by directly specifying the positions and velocities of its components, quantum mechanics replaces them by a complex mathematical object $|\Psi\rangle$, providing a relatively indirect description. This is an enormous change, not only mathematically, but also conceptually. The relations between $|\Psi\rangle$ and physical properties leave much more room for discussions about the interpretation of the theory than in classical physics. Actually, many difficulties encountered by those who tried (or are still trying) to “really understand” quantum mechanics are related to questions pertaining to the exact status of $|\Psi\rangle$. For instance, does it describe the physical reality itself, or only some (partial) knowledge that we might have of this reality? Does it describe ensembles of systems only (statistical description), or one single system as well (single events)? Assume that, indeed, $|\Psi\rangle$ is affected by an imperfect knowledge of the system; is it then not natural to expect that a better description should exist, at least in principle? If so, what would be this deeper and more precise description of the reality?

Another confusing feature of $|\Psi\rangle$ is that, for systems extended in space (for instance, a system made of two particles at very different locations), it gives an overall description of all its physical properties in a single block, from which the notion of space seems to have disappeared; in some cases, the properties of the two remote particles are completely “entangled” in a way where the usual notions of space-time and of events taking place in it seem to become diluted. It then becomes difficult, or even impossible, to find a spatio-temporal description of their correlations that remains compatible with relativity. All this is of course very different from the usual concepts of classical physics, where one attributes local properties to physical systems by specifying the density, the value of fields, etc. at each point of space. In quantum mechanics, this separability between the physical contents of different points of space is no longer possible in general. Of course, one could think that this loss of a local description is just an innocent feature of the formalism with no special consequence. For instance, in classical electromagnetism, it is often convenient to introduce a choice of gauge for describing the fields in an intermediate step; in the Coulomb gauge, the potential propagates instantaneously, while Einstein relativity forbids any communication that is faster than light. But this instantaneous propagation is just a mathematical artefact: when a complete calculation is made, proper cancellations of the instantaneous propagation take place so that, at the end, the relativistic limitation is perfectly obeyed. But, and as we will see below, it turns out that the situation is much less simple in quantum mechanics: in fact, a mathematical entanglement in $|\Psi\rangle$ can indeed have important physical consequences on the result of experiments, and even lead to predictions that are, in a sense, contradictory

with locality. Without any doubt, the state vector is a curious object to describe reality!

It is therefore not surprising that quantum mechanics should have given rise to so many interpretations. Their very diversity makes them interesting. Each of them introduces its own conceptual frame and view of physics, sometimes attributing to it a special status among the other natural sciences. Moreover, these interpretations may provide complementary views on the theory, shedding light onto some interesting features that, otherwise, would have gone unnoticed. The best-known example is Bohm's theory, from which Bell started to obtain a theorem illustrating general properties of quantum mechanics and entanglement, with applications ranging outside the Bohmian theory. Other examples exist, such as the use of stochastic Schrödinger dynamics to better understand the evolution of a quantum sub-system, the history interpretation and its view of complementarity, etc.

This book is intended for the curious reader who wishes to get a broad view on the general situation of quantum physics, including the various interpretations that have been elaborated, and without putting aside the difficulties when they occur. It is not a textbook designed for a first contact with quantum mechanics; there already exist many excellent reference books for students. In fact, from Chapter 1, the text assumes some familiarity with quantum mechanics and its formalism (Dirac notation, the notion of wave function, etc.). Any student who has already studied quantum mechanics for a year should have no difficulty in following the equations. Actually, there are relatively few in this book, which focuses, not on technical, but on logical and conceptual difficulties. Moreover, a chapter is inserted as an annex at the end of the book in order to help those who are not used to the quantum formalism. It offers a first contact with the notation; the reader may, while he/she progresses in the other chapters, choose a section of this chapter to clarify his/her ideas on such or such technical point.

Chapters 1 and 2 recall the historical context, from the origin of quantum mechanics to the present situation, including the successive steps from which the present status of $|\Psi\rangle$ emerged. Paying attention to history is not inappropriate in a field where the same recurrent ideas are so often rediscovered; they appear again and again, sometimes almost identical over the years, sometimes remodelled or rephrased with new words, but in fact more or less unchanged. Therefore, a look at the past is not necessarily a waste of time! Chapters 3, 4, and 5 discuss two important theorems, which form a logical chain, the EPR (Einstein, Podolsky, and Rosen) theorem and the Bell theorem; both give rise to various forms, several of which will be described. Chapter 6 gives a more general view on quantum entanglement, and Chapter 7 illustrates the notion with various processes that make use of it, such as quantum cryptography and teleportation. Chapter 8 discusses quantum measurement, in particular weak and continuous measurements. A few experiments

are described in Chapter 9; among the huge crowd of those illustrating quantum mechanics in various circumstances, we have chosen a small fraction of them – those where state vector reduction is “seen in real time”. Finally, Chapter 10, the longest of all chapters, gives an introduction and some discussion of the various interpretations of quantum mechanics. The chapters are relatively independent and the reader may probably use them in almost any order. Needless to say, no attempt was made to cover all subjects related to the foundations of quantum mechanics. A selection was unavoidable; it resulted in a list of subjects that the author considers as particularly relevant, but of course this personal choice remains somewhat arbitrary.

The motivation of this book is not to express preference for any given interpretations, as has already been done in many reference articles or monographs (we will quote several of them). It is even less to propose a new interpretation elaborated by the author. The objective is, rather, to review the various interpretations and to obtain a general perspective on the way they are related, their differences and common features, their individual consistency. Indeed, each of these interpretations has its own logic, and it is important to remember it; a classical mistake is to mix various interpretations together. For instance, the Bohmian interpretation has sometimes been criticized by elaborating constructions that retain some elements of this interpretation, but not all, or by inserting elements that do not belong to the interpretation; one then obtains contradictions. The necessity for logical consistency is general in the context of the foundations of quantum mechanics. Sometimes, the EPR argument or the Bell theorem have been misunderstood because of a confusion between their assumptions and conclusions. We will note in passing a few occasions where such mistakes are possible in order to help avoiding them. We should also mention that it is out of the question to give an exhaustive description of all interpretations of quantum mechanics here! They may be associated in many different ways, so that it is impossible to account for all possible combinations or nuances. A relatively abundant bibliography is proposed to the reader, but, in this case also, reaching any exhaustiveness is impossible; some choices have been made, sometimes arbitrary, in order to keep the total volume within reasonable limits.

To summarize, the main purpose of this book is an attempt to provide a balanced view on the conceptual situation of a theory that is undoubtedly one of the most remarkable achievements of the human mind, quantum mechanics, without hiding either difficulties or successes. As we already mentioned, its predictive power constantly obtains marvelous results in new domains, sometimes in a totally unpredictable way; nevertheless this intellectual edifice remains the object of discussions or even controversy concerning its foundations. No one would think of discussing classical mechanics or the Maxwell equations in the same way. Maybe

this signals that the final and optimum version of the theory has not yet been obtained?

Acknowledgments

Many colleagues played an important role in the elaboration of this book. The first is certainly Claude Cohen-Tannoudji, to whom I owe a lot. Over the years I benefited, as many others did, from his unique and deep way to use (and even to think) quantum mechanics; more than 40 years of friendship (and of common writing) and uncountable stimulating discussions were priceless for me. Alain Aspect is another friend with whom, from the beginning of his thesis in the seventies, a constant exchange of ideas took place (it still does!). At the time, the foundations of quantum theory were not very well considered among mainstream physicists, even sometimes perceived as passé or mediocre physics; Alain and I could comfort each other and make progress together in a domain that we both found fascinating, with the encouragement of Bernard d’Espagnat. Jean Dalibard and Philippe Grangier have been other wonderful discussion partners, always open minded with extreme intellectual clarity; I wish to thank them warmly. The title “Do we really understand quantum mechanics?” was suggested to me long ago by Pierre Fayet, on the occasion of two seminars on this subject he was asking me to give. This book arose from a first version of a text published in 2001 as an article in the *American Journal of Physics*, initiated during a visit at the Theoretical Physics Institute at the California University of Santa Barbara. During a session on Bose–Einstein condensation, I was lucky enough to discuss several aspects of quantum mechanics with its organizer, Antony Leggett; another lucky event favoring exchanges was to share Wojciech Zurek’s office! A little later, a visit at the Lorentz Institute of Leiden was also very stimulating, in particular with the help of Stig Stenholm. As for Abner Shimony, he guided me with much useful advice and encouraged the writing of the first version of this text.

Among those who helped much with the present version of the text, Michel Le Bellac played an important role by reading the whole text and giving useful advice, which helped to improve the text. He and Michèle Leduc also chose a wonderful (anonymous) reviewer who also made many very relevant remarks; I am grateful to all three of them. Among the other friends who also helped efficiently on various aspects are Roger Balian, Serge Reynaud, William Mullin, Olivier Darrigol, Bernard d’Espagnat and Catherine Chevalley; I am very grateful for many comments, advice, questions, etc. Markus Holzmann kindly read the whole manuscript when it was completed and made many interesting suggestions. The careful editing work of Anne Rix has been invaluable for improving the homogeneity and the quality of the text. I am grateful to Elisabeth Blind who kindly agreed that one of her wonderful paintings could be reproduced on the cover of this book.

Last but not least, concerning the last chapter describing the various interpretations of quantum mechanics, I asked specialists of each of these interpretations to be kind enough to check what I had written. I thank Sheldon Goldstein for reading and commenting the part concerning the Bohmian theory, Philip Pearle and Giancarlo Ghirardi for their advice concerning modified Schrödinger dynamics, Robert Griffiths and Roland Omnès for their comments on the history interpretation, Bernard d’Espagnat for clarifying remarks on veiled reality, Richard Healey for his help on the modal interpretation, Carlo Rovelli for his comments and suggestions on the relational interpretation, Alexei Grinbaum for illuminating comments concerning quantum logic and formal theories, and Thibault Damour for his helpful reading of my presentation of the Everett interpretation. According to the tradition it should be clear that, if nevertheless errors still subsist, the responsibility is completely the author’s. Finally, without the exceptional atmosphere of my laboratory, LKB, without the constant interaction with his members, and without the intellectual environment of ENS, nothing would have been possible.

1

Historical perspective

The founding fathers of quantum mechanics had already perceived the essence of the difficulties of quantum mechanics; today, after almost a century, the discussions are still lively and, if some very interesting new aspects have emerged, at a deeper level the questions have not changed so much. What is more recent, nevertheless, is a general change of attitude among physicists: until about 1970 or 1980, most physicists thought that the essential questions had been settled, and that “Bohr was right and proved his opponents to be wrong”. This was probably a consequence of the famous discussions between Bohr, Einstein, Schrödinger, Heisenberg, Pauli, de Broglie, and others (in particular at the Solvay meetings [1–3], where Bohr’s point of view had successfully resisted Einstein’s extremely clever attacks). The majority of physicists did not know the details of the arguments. They nevertheless thought that the standard “Copenhagen interpretation” had clearly emerged from the infancy of quantum mechanics as the only sensible attitude for good scientists. This interpretation includes the idea that modern physics must contain indeterminacy as an essential ingredient: it is fundamentally impossible to predict the outcome of single microscopical events; it is impossible to go beyond the formalism of the wave function (or its generalization, the state vector $|\Psi\rangle$) and complete it. For some physicists, the Copenhagen interpretation also includes the difficult notion of “complementarity” – even if it is true that, depending on the context, complementarity comes in many varieties and has been interpreted in many different ways! By and large, the impression of the vast majority was that Bohr had eventually won the debate against Einstein, so that discussing again the foundations of quantum mechanics after these giants was pretentious, passé, and maybe even bad taste.

Nowadays, the attitude of physicists is more open concerning these matters. One first reason is probably that the non-relevance of the “impossibility theorems” put forward by the defenders of the standard interpretation, in particular by Von Neumann [4], has now been better realized by the scientific community – see [5–7] and [8], as well as the discussion given in [9]). Another reason is, of course, the

great impact of the discoveries and ideas of J.S. Bell [6] in 1964. At the beginning of a new century, it is probably fair to say that we are no longer sure that the Copenhagen interpretation is the only possible consistent attitude for physicists – see for instance the doubts expressed by Shimony in [10]. Alternative points of view are considered with interest: theories including additional variables (or “hidden variables”¹) [11, 12]; modified dynamics of the state vector [7, 13–15] (non-linear and/or stochastic evolution); at the other extreme, we have points of view such as the so-called “many worlds interpretation” (or “many minds interpretation”, or “multibranched universe”) [16]; more recently, other interpretations such as that of “decoherent histories” [17] have been put forward (this list is non-exhaustive). These interpretations and several others will be discussed in Chapter 10. For a recent review containing many references, see [18], which emphasizes additional variables, but which is also characteristic of the variety of positions among contemporary scientists². See also an older but very interesting debate published in *Physics Today* [19]; another very useful source of older references is the 1971 *American Journal of Physics* “Resource Letter” [20]. But this variety of possible alternative interpretations should not be the source of misunderstandings! It should also be emphasized very clearly that, until now, no new fact whatsoever (or new reasoning) has appeared that has made the Copenhagen interpretation obsolete in any sense.

1.1 Three periods

Three successive periods may be distinguished in the history of the elaboration of the fundamental quantum concepts; they have resulted in the point of view that is called “the Copenhagen interpretation”, or “orthodox”, or “standard” interpretation. Actually, these terms may group different variants of the general interpretation, as we see in more detail below (in particular in Chapter 10). Here we give only a brief historical summary; we refer the reader who would like to know more about the history of the conceptual development of quantum mechanics to the book of Jammer [21] – see also [22] and [23]. For detailed discussions of fundamental problems in quantum mechanics, one could also read [10, 24, 25] as well as the references contained, or those given in [20].

1.1.1 Prehistory

Planck’s name is obviously the first that comes to mind when one thinks about the birth of quantum mechanics: in 1900, he was the one who introduced the famous

¹ As we discuss in more detail in §10.6, we prefer to use the words “additional variables” since they are not hidden, but actually appear directly in the results of measurements.

² For instance, the contrast between the titles of [10] and [18] is interesting.

constant h , which now bears his name. His method was phenomenological, and his motivation was actually to explain the properties of the radiation in thermal equilibrium (blackbody radiation) by introducing the notion of finite grains of energy in the calculation of the entropy [26]. Later he interpreted them as resulting from discontinuous exchange between radiation and matter. It is Einstein who, still later (in 1905), took the idea more seriously and really introduced the notion of quantum of light (which would be named “photon” only much later, in 1926 [27]) in order to explain the wavelength dependence of the photoelectric effect – for a general discussion of the many contributions of Einstein to quantum theory, see [28].

One should nevertheless realize that the most important and urgent question at the time was not so much to explain the fine details of the properties of interactions between radiation and matter, or the peculiarities of the blackbody radiation. It was more general: to understand the origin of the stability of atoms, that is of all matter which surrounds us and of which we are made! According to the laws of classical electromagnetism, negatively charged electrons orbiting around a positively charged nucleus should constantly radiate energy, and therefore rapidly fall onto the nucleus. Despite several attempts, explaining why atoms do not collapse but keep fixed sizes was still a complete challenge for physics³. One had to wait a little bit longer, until Bohr introduced his celebrated atomic model (1913), to see the appearance of the first ideas allowing the question to be tackled. He proposed the notion of “quantized permitted orbits” for electrons, as well as of “quantum jumps” to describe how they would go from one orbit to another, for instance during radiation emission processes. To be fair, we must concede that these notions have now almost disappeared from modern physics, at least in their initial forms; quantum jumps are replaced by a much more precise and powerful theory of spontaneous emission in quantum electrodynamics. But, on the other hand, one may also see a resurgence of the old quantum jumps in the modern use of the postulate of the wave packet (or state vector) reduction (§1.2.2.a). After Bohr, came Heisenberg, who, in 1925, introduced the theory that is now known as “matrix mechanics”⁴, an abstract intellectual construction with a strong philosophical component, sometimes close to positivism; the classical physical quantities are replaced by “observables”, corresponding mathematically to matrices, defined by suitable postulates without much help of intuition. Nevertheless, matrix mechanics contained many elements which turned out to be essential building blocks of modern quantum mechanics!

In retrospect, one can be struck by the very abstract and somewhat mysterious character of atomic theory at this period of history; why should electrons obey

³ For a review of the problem in the context of contemporary quantum mechanics, see [29].

⁴ The names of Born and Jordan are also associated with the introduction of this theory, since they immediately made the connexion between Heisenberg’s rules of calculation and those of matrices in mathematics.

such rules, which forbid them to leave a restricted class of orbits, as if they were miraculously guided on simple trajectories? What was the origin of these quantum jumps, which were supposed to have no duration at all, so that it would make no sense to ask what were the intermediate states of the electrons during such a jump? Why should matrices appear in physics in such an abstract way, with no apparent relation with the classical description of the motion of a particle? One can guess how relieved physicists probably felt when another point of view emerged, a point of view which looked at the same time much simpler and more in the tradition of the physics of the nineteenth century: the undulatory (or wave) theory.

1.1.2 The undulatory period

The idea of associating a wave with every material particle was first introduced by de Broglie in his thesis (1924) [30]. A few years later (1927), the idea was confirmed experimentally by Davisson and Germer in their famous electron diffraction experiment [31]. For some reason, at that time de Broglie did not proceed much further in the mathematical study of this wave, so that only part of the veil of mystery was raised by him (see for instance the discussion in [32]). It is sometimes said that Debye was the first, after hearing about de Broglie's ideas, to remark that in physics a wave generally has a wave equation: the next step would then be to try and propose an equation for this new wave. The story adds that the remark was made in the presence of Schrödinger, who soon started to work on this program; he successfully and rapidly completed it by proposing the equation which now bears his name, one of the most basic equations of all physics. Amusingly, Debye himself does not seem to have remembered the event. The anecdote may be inaccurate; in fact, different reports about the discovery of this equation have been given and we will probably never know exactly what happened. What remains clear is that the introduction in 1926 of the Schrödinger equation for the wave function⁵ [33] is one of the essential milestones in the history of physics. Initially, it allowed one to understand the energy spectrum of the hydrogen atom, but it was soon extended and gave successful predictions for other atoms, then molecules and ions, solids (the theory of bands for instance), etc. It is presently the major basic tool of many branches of modern physics and chemistry.

Conceptually, at the time of its introduction, the undulatory theory was welcomed as an enormous simplification of the new mechanics. This is particularly true because Schrödinger and others (Dirac, Heisenberg) promptly showed how it could be used to recover the predictions of matrix mechanics from more intuitive considerations, using the properties of the newly introduced “wave function” – the solution of the Schrödinger equation. The natural hope was then to extend this

⁵ See footnote 11 for the relation between the state vector and the wave function.

success, and to simplify all problems raised by the mechanics of atomic particles: one would replace it by a mechanics of waves, which would be analogous to electromagnetic or sound waves. For instance, Schrödinger initially thought that all particles in the universe looked to us like point particles just because we observe them at a scale which is too large; in fact, they are tiny “wave packets” which remain localized in small regions of space. He had even shown that these wave packets remain small (they do not spread in space) when the system under study is a harmonic oscillator – alas, we now know that this is a very special case; in general, the wave packets constantly spread in space!

1.1.3 Emergence of the Copenhagen interpretation

It did not take long before it became clear that the completely undulatory theory of matter also suffered from very serious difficulties, actually so serious that physicists were soon led to abandon it. A first example of difficulty is provided by a collision between particles, where the Schrödinger wave spreads in all directions, like a circular wave in water stirred by a stone thrown into it; but, in all collision experiments, particles are observed to follow well-defined trajectories and remain localized, going in some precise direction. For instance, every photograph taken in the collision chamber of a particle accelerator shows very clearly that particles never get “diluted” in all space! This stimulated the introduction, by Born in 1926, of the probabilistic interpretation of the wave function [34]: quantum processes are fundamentally non-deterministic; the only thing that can be calculated is probabilities, given by the square of the modulus of the wave function.

Another difficulty arises as soon as one considers systems made of more than one single particle: then, the Schrödinger wave is no longer an ordinary wave since, instead of propagating in normal space, it propagates in the so-called “configuration space” of the system, a space which has $3N$ dimensions for a system made of N particles! For instance, already for the simplest of all atoms, the hydrogen atom, the wave propagates in six dimensions⁶. For a collection of atoms, the dimension grows rapidly, and becomes an astronomical number for the ensemble of atoms contained in a macroscopic sample. Clearly, the new wave was not at all similar to classical waves, which propagate in ordinary space; this deep difference will be a sort of Leitmotiv in this text⁷, reappearing under various aspects here and there⁸.

⁶ This is true if spins are ignored; if they are taken into account, four such waves propagate in six dimensions.

⁷ For instance, the non-locality effects occurring with two correlated particles can be seen as a consequence of the fact that the wave function propagates locally, but in a six-dimensional space, while the usual definition of locality refers to ordinary space which has three dimensions.

⁸ Quantum mechanics can also be formulated in a way that does not involve the configuration space, but just the ordinary space: the formalism of field operators (sometimes called second quantization, for historical reasons). One can write these operators in a form that is similar to a wave function. Nevertheless, since they are quantum operators, their analogy with a classical field is even less valid.

In passing, it is interesting to notice that the recent observation of the phenomenon of Bose–Einstein condensation in dilute gases [35] can be seen, in a sense, as a sort of realization of the initial hope of Schrödinger: this condensation provides a case where a many-particle matter wave does propagate in ordinary space. Before condensation takes place, we have the usual situation: the gas has to be described by wave functions defined in a huge configuration space. But, when the atoms are completely condensed into a single-particle wave function, they are restricted to a much simpler many-particle state built with the same ordinary wave function, as for a single particle. The matter wave then becomes similar to a classical field with two components (the real part and the imaginary part of the wave function), resembling an ordinary sound wave for instance. This illustrates why, somewhat paradoxically, the “exciting new states of matter” provided by Bose–Einstein condensates are not an example of an extreme quantum situation; in a sense, they are actually more classical than the gases from which they originate (in terms of quantum description, interparticle correlations, etc.). Conceptually, of course, this remains a very special case and does not solve the general problem associated with a naive view of the Schrödinger waves as real waves.

The purely undulatory description of particles has now disappeared from modern quantum mechanics. In addition to Born and Bohr, Heisenberg [36], Jordan [37, 38], Dirac [39] and others played an essential role in the appearance of a new formulation of quantum mechanics [23], where probabilistic and undulatory notions are incorporated in a single complex logical edifice. The probabilistic component is that, when a system undergoes a measurement, the result is fundamentally random; the theory provides only the probabilities of the different possible outcomes. The wave component is that, when no measurements are performed, the Schrödinger equation is valid. The wave function is no longer considered as a direct physical description of the system itself; it is only a mathematical object that provides the probabilities of the different results⁹ – we come back to this point in more detail in §1.2.3.

The first version of the Copenhagen interpretation was completed around 1927, the year of the fifth Solvay conference [3]. Almost immediately, theorists started to extend the range of quantum mechanics from particle to fields. At that time, the interest was focussed only on the electromagnetic field, associated with the photon, but the ideas were later generalized to fields associated with a wide range of particles (electrons, muons, quarks, etc.). Quantum field theory has now enormously expanded and become a fundamental tool in particle physics, within a relativistic formalism (the Schrödinger equation itself does not satisfy Lorentz invariance).

⁹ In the literature, one often finds the word “ontological” to describe Schrödinger’s initial point of view on the wave function, as opposed to “epistemological” to describe the probabilistic interpretation.

A generalization of the ideas of gauge invariance of electromagnetism has led to various forms of gauge theories; some are at the root of our present understanding of the role in physics of the fundamental interactions (electromagnetic, weak, strong¹⁰) and led to the successful prediction of new particles. Nevertheless, despite all these remarkable successes, field theory remains, conceptually, on the same fundamental level as the theory of a single non-relativistic particle treated with the Schrödinger equation. Since this text is concerned mostly with conceptual issues, we will therefore not discuss field theory further.

1.2 The state vector

Many discussions concerning the foundations of quantum mechanics are related to the status and physical meaning of the state vector. In §§1.2.1 and 1.2.2, we begin by first recalling its definition and use in quantum mechanics (the reader familiar with the quantum formalism might wish to skip these two sections); then, in §1.2.3, we discuss the status of the state vector in standard quantum mechanics.

1.2.1 Definition, Schrödinger evolution, Born rule

We briefly summarize how the state vector is used in quantum mechanics and its equations; more details are given in §11.1.1 and following.

1.2.1.a Definition

Consider a physical system made of N particles with mass, each propagating in ordinary space with three dimensions; the state vector $|\Psi\rangle$ (or the associated wave function¹¹) replaces in quantum mechanics the N positions and N velocities which, in classical mechanics, would be used to describe the state of the system. It is often convenient to group all these positions and velocities within the $6N$ components of a single vector $\mathbf{V}$ belonging to a real vector space with $6N$ dimensions, called “phase space”¹²; formally, one can merely consider that the state vector $|\Psi\rangle$ is the quantum equivalent of this classical vector $\mathbf{V}$. It nevertheless belongs to a space that is completely different from the phase space, a complex vector space called “space of states” (or, sometimes, the “Hilbert space” for historical reasons) with infinite dimension. The calculations in this space are often made with the help of

¹⁰ There is a fourth fundamental interaction in physics, gravitation. The “standard model” of field theory unifies the first three interactions, but leaves gravitation aside. Other theories unify the four fundamental interactions, but for the moment they are not considered standard.

¹¹ For a system of spinless particles with masses, the state vector $|\Psi\rangle$ is equivalent to a wave function, but for more complicated systems this is not the case. Nevertheless, conceptually they play the same role and are used in the same way in the theory, so that we do not need to make a distinction here.

¹² The phase space therefore has twice as many dimensions as the configuration space mentioned above.

the Dirac notation [39], which actually we will use here, and where the vectors belonging to the space of states are often called “kets”.

Because the state vector belongs to a linear space, any combination of two arbitrary state vectors $|\Psi_1\rangle$ and $|\Psi_2\rangle$ belonging to the space of states:

$$|\Psi\rangle = \alpha |\Psi_1\rangle + \beta |\Psi_2\rangle \quad (1.1)$$

(where α and β are arbitrary complex numbers) is also a possible state for the system. This is called the “superposition principle” of quantum mechanics, and has many consequences.

Moreover, to each physical observable of the system, position(s), momentum(ta), energy, angular momentum, etc., the formalism of quantum mechanics associates a linear operator acting in the space of states, and provides rules for constructing these operators. For historical reasons (§1.1.1), each of these operators is often called “observable”; they belong to the category of mathematical operators called “linear Hermitian operators”.

1.2.1.b Schrödinger evolution

The evolution of the state vector $|\Psi(t)\rangle$ between time t_0 and t_1 is given by the Schrödinger equation:

$$i\hbar \frac{d}{dt} |\Psi(t)\rangle = H(t) |\Psi(t)\rangle \quad (1.2)$$

where $H(t)$ is the Hamiltonian evolution of the system (including the internal interactions of this system as well as the effects of classical external fields applied to it, for instance static or time-dependent magnetic fields). The Schrödinger equation is a linear differential equation, similar to many other such equations in physics. It leads to a progressive evolution of the state vector, without any quantum jump or discontinuity. It is as general as the Newton or Lagrange equations in classical mechanics, and can be applied to all physical situations, provided of course the system is well defined with a known Hamiltonian.

In particular, the Schrödinger equation can also be applied to a situation where the physical system interacts with a measurement apparatus (a spin 1/2 particle entering the magnetic field gradient created by a Stern–Gerlach apparatus for instance); it then does not select precise experimental results, but keeps all of them as potentialities (within a so-called “coherent superposition”). One more ingredient is then introduced into the theory, the Born probability rule.

1.2.1.c Born probability rule

We assume that, at time t_1 , when the solution $|\Psi(t)\rangle$ of equation (1.2) takes the value $|\Psi(t_1)\rangle$, the system undergoes a measurement, associated with an operator

M (observable) acting in the space of states. We note $|m_i\rangle$ the eigenvectors of M associated with eigenvalues m_i ($i = 1, 2, \dots$); if some eigenvalues are degenerate, several consecutive values in the series of m_i are equal, but associated with different vectors $|m_i\rangle$. Since M is an Hermitian operator, the $|m_i\rangle$ can be chosen as an orthonormal basis of the space of states.

The Born probability rule then states that, in an ideal measurement:

- (i) the result of a measurement associated with M can only be one of the m_i ; other results are never obtained.
- (ii) if a particular eigenvalue m_i is non-degenerate, the probability $\mathcal{P}_i$ of obtaining result m_i is given by the square modulus of the scalar product of $|\Psi(t_1)\rangle$ by $|m_i\rangle$:

$$\mathcal{P}_i = |\langle m_i | \Psi(t_1) \rangle|^2 \quad (1.3)$$

- (iii) the probability of measuring a degenerate eigenvalue is the sum of the probabilities (1.3) corresponding to all the orthonormal eigenvectors associated with this eigenvalue¹³.

Rules (ii) and (iii) may be grouped in a simple form, which will be useful in what follows. If the result corresponds to an eigenvalue m that is p times degenerate, the series of p numbers $m_i, m_{i+1}, \dots, m_{i+p}$ have the same value m . We can then introduce the sum of the projectors (§11.1.3) over the corresponding eigenvectors:

$$P_M(m) = |m_i\rangle\langle m_i| + |m_{i+1}\rangle\langle m_{i+1}| + \dots + |m_{i+p}\rangle\langle m_{i+p}| \quad (1.4)$$

This operator is also a projector (it is equal to its square), which can be applied to the state vector $|\Psi(t_1)\rangle$ before the measurement:

$$P_M(m) |\Psi(t_1)\rangle = |\Psi'_m\rangle \quad (1.5)$$

The probability of obtaining result m in the measurement is then nothing but the square of the norm of $|\Psi'\rangle$:

$$\mathcal{P}_m = \langle \Psi'_m | \Psi'_m \rangle = \langle \Psi(t_1) | P_M(m) | \Psi(t_1) \rangle \quad (1.6)$$

1.2.2 Measurement processes

The standard interpretation of quantum mechanics contains the progressive, deterministic, evolution of the wave function/state due to the Schrödinger equation. Usually, one also includes in this interpretation a second postulate of evolution;

¹³ Similarly, in the classical theory of probabilities, if an event E can be obtained either as event e_1 , or e_2 , ..., or e_i , ..., and if all events e_i are exclusive, the probability of E is the sum of the probabilities of the e_i .

this postulate is associated with the process of measurement, and completely different from the Schrödinger evolution since it is discontinuous. It is often called the “wave packet reduction”, or “wave function collapse”, or again “state vector reduction”, and was introduced by Von Neumann in his famous treatise (Chapter VI of [4]). This is the version found in most textbooks. But Bohr himself preferred another point view where state vector reduction is not used¹⁴; we discuss this point of view afterwards (there exist also other interpretations of quantum mechanics that do not make use of state vector reduction, as discussed in Chapter 10; see for instance §§10.1.2, 10.6, or 10.11).

1.2.2.a Von Neumann, reduction (collapse)

Suppose now that the system we study is prepared at time t_0 , evolves freely (without being measured) until time t_1 where it undergoes a first measurement, and then evolves freely again until time t_2 where a second measurement is performed. Just after the first measurement at time t_1 , when the corresponding result of measurement is known, it is very natural to consider that both the initial preparation and the first measurement are part of a single preparation process of the system. One then associates to this preparation a state vector that includes the information of the first result; this is precisely what the state vector reduction (or state collapse) postulate does. The new “reduced” state vector can then be used as an initial state to calculate the probabilities of the different results corresponding to the second measurement, at time t_2 .

Dirac also takes this point of view when he writes (page 9 of “Quantum mechanics” [39]): “There are, however, two cases when we are in general obliged to consider the disturbance as causing a change in state of the system, namely, when the disturbance is an observation and when it consists in preparing the system so as to be in a given state”.

We assume that the measurement is ideal¹⁵ – it preserves the integrity of the system, as opposed to destructive measurements such as the absorption of a photon in a detector. Then, after the measurement associated to M has provided result m_i corresponding to a non-degenerate eigenvalue (and therefore to a single $|m_i\rangle$), the reduced state vector is:

$$|\Psi'_{m_i}\rangle = |m_i\rangle \quad (1.7)$$

¹⁴ As stated in [40]: “Most importantly, Bohr’s complementarity interpretation makes no mention of wave packet collapse ... or a privileged role for the subjective consciousness of the observer. Bohr was also in no way a positivist. Much of what passes for the Copenhagen interpretation is found in the writings of Werner Heisenberg, but not in Bohr”.

¹⁵ We come back in more detail on the Von Neumann model of measurement in §8.1.1 and on the notion of QND (quantum non-demolition) measurement.

In other words, at time t_1 when the first measurement is performed and provides result m_i , the state vector jumps discontinuously from $|\Psi(t_1)\rangle$ to $|m_i\rangle$. If the same measurement is repeated very shortly after, by applying the Born rule (1.3) to state $|\Psi'\rangle$, one finds that all probabilities are zero but one. The result is then certain: one obtains m_i again.

The generalization of (1.7) to a degenerate eigenvalue is given by (1.5):

$$|\Psi'_m\rangle = c P_M(m) |\Psi(t_1)\rangle \quad (1.8)$$

where a normalization coefficient c has been added¹⁶.

The rule can easily be generalized to more than two successive measurements. Each time a measurement is performed and a result obtained, the state vector jumps to a new value that includes this new information (but may also erase some previous information). One exception occurs if the same measurement is performed repeatedly, at times that are sufficiently close to avoid any Schrödinger evolution of the system between the measurements. Then all results are necessarily the same and, after the first measurement, the state vector reduction has no effect (but it becomes effective again as soon as a different observable is measured).

Clearly, state vector reduction is closely related to the Born probability rule. Actually, if one generalizes this rule to multi-time measurements (§10.1.2.a and Appendix G), one can derive the reduction of the state vector as a convenient rule to calculate probabilities. From this point of view, the state reduction no longer appears as a postulate, but just as a convenient way to make calculations that can be derived from the generalized Born rule.

To summarize, the general scheme then includes different stages in the evolution of the state vector. Between preparation and measurements, it evolves continuously according to the Schrödinger equation. When the system undergoes a measurement, it interacts with a measurement apparatus. The probabilities of the various outcomes can be calculated from the state vector in a perfectly well defined way from this equation and from the Born rule, but only the probabilities: the Schrödinger equation itself does not select a single result. The uniqueness of the outcome in the quantum description of the system is associated with state vector reduction, a process that makes the state vector jump discontinuously (and randomly) to a new value. The emergence of a single result is then obtained (one could say “forced”) explicitly in the state vector by retaining only the appropriate component corresponding to the observed outcome – while all components of the state vector associated with the other results are put to zero (hence the name “reduction”) – we will come back to this question in §2.1. The reduction process is discontinuous and irreversible. In this general scheme, separate rules and equations are therefore introduced, one for

¹⁶ This coefficient is the inverse of the square root of $\langle \Psi(t_1) | P_M(m) | \Psi(t_1) \rangle$.

the “natural” continuous evolution of the system between measurements, another for the measurements performed on it. The difficulty is then to understand precisely how to avoid possible conflicts between these two different postulates.

1.2.2.b Bohr

Bohr does not make explicit use of the notion of state vector reduction. He prefers to see the ensemble of all measurements performed at different times as being part of one single big experiment, which includes all experimental devices necessary to perform the series of measurements (and presumably also those involved in the preparation stage of the studied system as well). The rules of quantum mechanics then provide the probabilities corresponding to all possible series of results. In this view, one should not ask (as we just did in §1.2.2.a) what is the quantum state of the measured system between the first measurement and the second: separating the system from the whole experimental apparatus has no meaning in Bohr’s interpretation of quantum mechanics (non-separability, §3.3.3.b).

Any possible conflict between two different evolution postulates then disappears. It is nevertheless replaced by another difficulty since, in each experiment, one has to distinguish between two different parts: observed system(s) and measurement apparatus(es). Only the latter are directly accessible to human experience and can be described with ordinary language, as in classical physics. The measurement apparatuses and observers then have a very specific role. They provide the results of measurements to observers, so that they are at the origin of our perception of the physical world; but at the same time they also introduce an irreducible non-deterministic component into the theory and into the evolution of physical systems. The difficulty is then to decide where exactly to put the frontier between the two different parts¹⁷. For instance, if the distinction is to be made in terms of size of the systems, one could ask from what size a physical system is sufficiently macroscopic to be considered as directly accessible to human experience, and will behave as a measurement apparatus. If the distinction is made in another way than size, then more elaborate rules should be specified to clarify this vague concept. In [41], Bohr writes (see also Bell’s quotation in §2.5): “This necessity of discriminating in each experimental arrangement between those parts of the physical system considered which are to be treated as measuring instruments and those which constitute the objects under investigation may indeed be said to form a *principal distinction between classical and quantum-mechanical description of physical phenomena*”.

¹⁷ This does not necessarily mean that one has to choose a frontier between two parts of the physical world that is fixed forever. In Bohr’s view, it may depend on the experiment considered (one could imagine that an element that is part of the measured system in one experiment becomes part of the measurement apparatus in another).

1.2.3 Status

Under these conditions, what is the status of the state vector (or wave function) in standard quantum mechanics?

1.2.3.a Two extremes

When discussing the status of the state vector, two opposite mistakes should be avoided, since both “miss the target” on different sides. The first is to endorse the initial hopes of Schrödinger and to decide that the wave function, propagating in a space with many dimensions, directly describes the physical properties of the system. In such a purely undulatory view, the positions and velocities of particles are replaced by the amplitude of a complex wave, and the very notion of point particle becomes diluted. Nevertheless, the difficulties introduced by this view are now so well known – see discussion in the preceding section – that nowadays few physicists seem to be tempted to support it, at least within standard theory¹⁸. At the opposite extreme, one considers that the wave function does not attempt to describe the physical properties of the system itself, but just the information that some observer has on it. It then becomes analogous to a classical probability distribution in usual statistical theory; the wave function should then get a relative (or contextual) status depending on the observer. Of course, at first sight, this seems to bring an elementary solution to the difficulties introduced into quantum mechanics by the state vector reduction and its discontinuities (§1.2.2): we all know that classical probabilities may undergo sudden jumps, and nobody considers this as a special problem. For instance, as soon as new information on an (possibly remote) event becomes available to us, the probability distribution that we associate with it undergoes a sudden and complete change; by analogy, is this not the obvious way to explain the sudden state vector reduction?

One first difficulty of this point of view is that it would naturally lead to a relative character of the wave function: if two observers had different information on the same system, should they use different wave functions to describe the same system¹⁹? In classical probability theory, distributions of probabilities that

¹⁸ Within theories introducing modified Schrödinger dynamics (non-standard versions requiring a modification of the Schrödinger equation, see §10.8), a purely undulatory point of view can indeed be reintroduced.

¹⁹ We assume that the two observers use the same reference frame for space-time; they differ only by the amount of information they have. If they used different Galilean reference frames, one should apply simple mathematical transformations to go from one state vector to the other. But this has no more conceptual impact than the usual transformations which allow us, in classical mechanics, to transform positions and conjugate momenta when changing the frame of reference.

For completeness, we should add that there is also room in quantum mechanics for classical uncertainties arising from an imperfect knowledge of the system, in particular with the formalism of the density operator (§11.1.5). Here, we limit ourselves to the discussion of state vectors and wave functions (pure states).

Finally, let us mention that, in a more elaborate discussion, one would introduce other elements, for instance the notion of intersubjectivity, etc. [10, 24].

undergo sudden jumps are inherently “observer-dependent” and thus “subjective”: the observer who has more information describes the phenomenon with a distribution that is narrower than that of another, less informed, observer; in principle, one can always imagine an observer with perfect knowledge of the phenomenon for whom the probability is perfectly peaked at one value (the process is then deterministic for this observer). But standard quantum mechanics rejects the possibility of such a perfect description of all properties of a system, even for an observer who has all the possible information. Standard quantum mechanics does not attribute such a relative character to the wave function or state vector.

Moreover, when in ordinary probability theory a distribution undergoes a sudden “jump” to a more precise distribution, the reason is simply that more precise values of the variables already exist – they actually existed before the jump. In other words, the very fact that the probability distribution reflected our imperfect knowledge implies the possibility of a jump to a more precise description, closer to the reality of the system itself²⁰. But this is in complete opposition with orthodox quantum mechanics, which rejects the notion of any better description of the reality (§3.3.2). In this theory, the random character of the measurement process is absolute; it is vain to try and explain a particular result of measurement by causes, since they do not exist. Ignorance cannot be invoked to explain the jumps of the state vector in standard quantum mechanics.

We also note that classical probabilities can also be defined in the absence of observers and then called “objective”²¹: one considers a physical event taken among a whole ensemble of possible events, which all have well defined but different dynamics because some of their initial physical properties differ, or because they undergo uncontrolled perturbations. To introduce probabilities, one then postulates that all these distinct events have the same “weight”. By contrast, standard quantum mechanics introduces probabilities only if observers make measurements; it rejects the idea of different initial properties, or perturbations, explaining the occurrence of these probabilities.

Introducing the notion of pre-existing values is precisely the basis of “unorthodox” theories with additional variables (often called hidden variables – see §10.6 and footnote 1); uncontrolled perturbations are invoked in the theories with modified Schrödinger dynamics (§10.8). So, advocating this “information interpretation”

²⁰ Normally, information is about something! An information, or a probability, refers to an object or an event that is supposed to have its own reality, independently of the information, and of the fact that someone is acquiring it or not – see for instance §VII of [42]. It would be very unusual to define the object by the information itself, raising questions about logical circularity.

²¹ In classical theory, objective probabilities occur in classical statistical physics and the study of chaotic systems made of many particles. Nevertheless, these probabilities are then continuous functions of time; sudden jumps occur only for subjective classical probabilities.

for a straightforward explanation of state vector reduction amounts to advocating non-orthodox approaches to quantum mechanics. It is therefore important to keep in mind that, in the standard interpretation of quantum mechanics, the wave function (or state vector) gives THE ultimate physical description of the system, with all its physical properties; it is neither contextual, nor observer dependent (subjective), but absolute; if it gives probabilistic predictions on the result of future measurements, it nevertheless remains inherently completely different from an ordinary classical distribution of probabilities.

1.2.3.b The Copenhagen (orthodox) point of view

If none of these extremes is correct, how should we combine them to obtain the status of the state vector in orthodox quantum mechanics? To what extent should we consider that the wave function describes the physical system itself (realist interpretation), or rather that it contains only the information that we may have on it (positivistic interpretation), presumably in some sense that is more subtle than a classical distribution function? In orthodox quantum theory, the state vector has a really non-trivial status – actually it has no equivalent in all the rest of physics.

The Copenhagen/orthodox interpretation is not defined in exactly the same way by all authors²²; nuances exist between the different definitions of what is a quantum state. It may then be more appropriate to speak of the “standard” definition, namely that found in most textbooks. The standard view is that the state vector (or wave function) is associated with a preparation procedure of the physical system under consideration. Dirac, in Chapter I of [39], writes: “We must first generalize the meaning of a ‘state’ so that it can apply to an atomic system ... The method of preparation may then be taken as the specification of the state”. Or Stapp, when introducing the Copenhagen interpretation [43]: “The specifications A on the manner of preparation of the physical system are first transcribed into a wave function $\Psi_A(x)$ ”. Similarly, Peres writes [44]: “a state vector is not a property of a physical system, but rather represents an experimental procedure for preparing or testing one or more physical systems”. He also makes the general comment: “quantum theory is incompatible with the proposition that measurements are processes by which we discover some unknown and pre-existing property”. In this view, a wave function is an objective representation (independent of the observer) of a preparation procedure, rather than of the isolated physical system itself. The preparation procedure

²² For instance, Howard writes [40]: “Much of what passes for the Copenhagen interpretation is found in the writings of Werner Heisenberg, but not in Bohr. Indeed, Bohr and Heisenberg disagreed for decades in deep and important ways. The idea that there was a unitary Copenhagen point of view on interpretation was, it shall be argued, a post-war invention, for which Heisenberg was chiefly responsible. Many other physicists and philosophers, each with his own agenda, contributed to the promotion of this invention for polemical or rhetorical purposes”.

may imply the measurement of a physical quantity by some observer; for instance, one can prepare the state of a spin by sending the particle through a Stern–Gerlach magnet, and measuring the position of the particle at the output. But one can also prepare a system by letting it reach thermal equilibrium in interaction with a thermostat, as for instance is the case in most magnetic resonance experiments; then, the description of the system requires, not a single state vector, but many, which can all be summarized in a “density operator” (§11.1.5). In any case, it seems safe to associate the standard interpretation of the state vector (or density operator) with a physical preparation procedure of the system under study.

This does not exclude that the state vector may contain some information about the properties of the system itself. Indeed, when a given physical quantity has been measured, and when a given result has been obtained, quantum mechanics predicts that the same result is always obtained if the same measurement is repeated just after. It is then natural to think that, between the measurements, the system had a physical property related to the certainty of the result of the second measurement – see Chapter 3. The result of measurement has then become a property of the system. This idea can be expressed mathematically: all operators acting in the space of states that take $|\Psi\rangle$ as one of their eigenvectors correspond to physical quantities that have precise values in this state, and can be associated with properties of the system. But, for a quantum system in a given state $|\Psi\rangle$, most operators are not in this case; only a very small proportion of the properties that could be defined in classical mechanics can still be defined quantum mechanically; moreover, the list of those that are defined is not fixed, but depends on $|\Psi\rangle$. Hartle proposes the following definition [45]: “The state of an individual system in quantum mechanics is, therefore, defined as the list of all propositions (concerning the individual system) together with their truth values – true, false, or indefinite”.

At the end, we reach a sort of intermediate situation where neither the pure continuous undulatory interpretation nor the pure probabilistic interpretation is correct, while elements of both points of view are retained. A preparation procedure for a physical system is indeed an objective fact, and if the state vector describes such a procedure, it also necessarily has some objective component – it cannot be entirely mental and observer dependent. Sometimes some properties of the system exist, sometimes others, but most do not exist, so that their measurement will provide random results. Both interpretations are combined in a way that emphasizes the role of the whole experimental setup. Bohr described the situation by using the general concept “complementarity”; for instance, in [46], he writes “the viewpoint of complementarity presents itself as a rational generalization of the very idea of causality” (how to relate in a causal way a preparation procedure to an observation procedure).

2

Present situation, remaining conceptual difficulties

Conceptual difficulties still remain in quantum mechanics, even if they had already been identified by its inventors. This does not mean that the theory is not successful! The reality is quite the opposite: in fact, independently of these difficulties, quantum mechanics is certainly one of the most successful theories of all science. One can even consider that its ability to adapt to new situations is one of its most remarkable features. It continues to give efficient and accurate predictions while new experiments are performed, even in situations that the founding fathers had no way to imagine. Actually, there are very few theories that have been verified with the same accuracy in so many situations. Nevertheless, it remains true that conceptual difficulties subsist, and their discussion is the object of this chapter. As we will see, most of them relate to the process of quantum measurement, in particular to the very nature of the random process that takes place on this occasion.

We have seen that, in most cases, the wave function evolves gently, in a perfectly predictable and continuous way, according to the Schrödinger equation; in some cases only (when a measurement is performed), unpredictable changes take place, according to the Von Neumann postulate of state vector reduction. Obviously, having two different postulates for the evolution of the same mathematical object is very unusual in physics. The notion was a complete novelty when it was introduced, and still remains unique, but also the source of difficulties – in particular logical difficulties related to the compatibility between the two different postulates. Normally, one would tend to see a measurement process, not as very special, but just as an interaction process between a (possibly microscopic) measured system and a macroscopic measurement apparatus. The latter should be treated within the ordinary laws of physics: indeed, a theory where no distinction is necessary between normal evolution and measurements would seem much more general, and therefore preferable. In other words, why give such a special character to measurements so that two separate postulates become necessary? Where exactly does the range of application of the first postulate stop in favor of the second? More precisely, among

all the interactions – or perturbations – that a physical system can undergo, which ones should be considered as normal (Schrödinger evolution), which ones are a measurement (state vector reduction)? We need to understand better the reason why physics needs to introduce a split between two different worlds, and the nature of this split: Schrödinger’s world, which is perfectly continuous and deterministic, and Born’s measurement world, which is very different since the continuous evolution stops to be replaced by non-deterministic and discontinuous processes.

From Bohr’s point of view also, measurement processes and apparatuses are clearly “discriminated” from the system under study. In Bohr’s universe, a general evolution takes place in a continuous and deterministic way according to the Schrödinger equation. But, in the particular case of events involving a transfer of knowledge to an observer, randomness appears in the evolution of physical systems. Moreover, these measurement processes are so to say considered as “closed bubbles” inserted within this general evolution, closed events from their beginning to their end that extend over a whole region of space-time. They cannot be split into more detailed relativistic events, and are fundamentally characterized by the fact that an intelligent human being is asking a question to Nature; the outcome is a unique answer, but non-deterministic.

Whether we prefer Von Neumann’s or Bohr’s point of view, we are faced with a logical problem that did not exist before in classical mechanics, when nobody thought that measurements should be treated as special processes in physics. We learn from Bohr that we should not try to transpose our experience of the everyday world to microscopic systems; then, for each experiment, where exactly is the limit between these two worlds? Is it sufficient to reply that there is so much room between macroscopic and microscopic sizes that the exact position of the border does not matter?

It may also look very surprising that, in modern physics, the “observer” should play such a central role, giving to the theory an unexpected anthropocentric foundation, as in astronomy in the middle ages. Should we really reject the description of isolated systems as unscientific, just because we are not observing them? If observers are so important, how precisely is an observer defined? For instance, can an animal perform a measurement and reduce the state vector, or is a human being required? Bell once asked with humour [47]: “Was the world wave function waiting to jump for thousand of millions of years until a single-celled living creature appeared? Or did it have to wait a little longer for some highly qualified measurer, – with a Ph.D.?” – see also [48], in particular its title. These general questions are difficult, somewhat philosophical, and we will come back to them in Chapter 10 where, for instance, we mention that London and Bauer [49] have suggested that it is the faculty of introspection that qualifies a living creature to be an observer in quantum theory.

Another difficulty is related to the random character of the predictions of quantum mechanics, which provides only probabilities for the results of measurements. In itself, this situation is relatively usual in physics, where for instance statistical classical mechanics makes constant use of probabilities. But the probabilities are then associated with the notion of statistical ensemble: one assumes that the system under study is chosen randomly among many similar systems belonging to the same ensemble. Indeed, they share common properties contained in the statistical description, but a more accurate description can reveal that, at some finer level, they also differ by other individual properties. In other words, every single system has more physical properties than those specified for the whole ensemble. The same question can immediately be transposed to quantum mechanics: if its predictions are only probabilities, does it mean that it describes ensembles of systems only? Or should we consider that the state vector provides the most accurate possible description of a single system? We will often come back to this question in what follows, in particular in Chapters 3 and 10.

It is probably impossible to summarize all that has been written on the exact behavior and role of the state vector in a reasonable number of pages. As illustrations, we will discuss a few examples in this chapter: Von Neumann's infinite regress (§2.1), Schrödinger's cat (§2.2), Wigner's friend (§2.3), and negative measurements (§2.4). We will then propose a number of quotations (§2.5), which give an idea of the variety of possible positions on quantum mechanics, and finally mention a few points where errors have been made in the past (§2.6).

2.1 Von Neumann's infinite regress/chain

Von Neumann, in a treatise published in 1932, introduced an explicit theory of quantum measurement (Chapters. 4–6 of [4]). In contrast with Bohr's approach, Von Neumann considers the measuring apparatus as a quantum system. He assumes that the measured system S is put into contact with a measurement apparatus M and interacts with it for some time. M contains a macroscopic “pointer” P which, after the interaction has finished, reaches a position that depends on the initial state of S . In this chapter, we introduce the general ideas related to the Von Neumann infinite chain, without writing equations; we come back in §8.1 to Von Neumann's theory of measurement with more detail – see also, for instance, §9.2 of [21] and §11.2 of [50].

Assume that, initially, this system was in an eigenstate of the measured observable A characterized by one of its eigenvalues¹ a . After the end of the interaction, the

¹ Here, for the sake of simplicity, we assume that the eigenvalue is not degenerate, but the generalization to degenerate eigenvalues is possible.

pointer reaches a specific position that depends on the eigenstate; one can consider that observing the macroscopic position of P amounts to measuring the observable and obtaining a as a result.

Now, in general S is not initially in an eigenstate of the measured observable A , but in a superposition of such states. The linear Schrödinger equation then predicts that the whole physical system $S+M$ reaches a linear superposition² of states after the interaction. In the different components of this superposition, the pointer has several different positions, each corresponding to a different result of measurement; this situation is described as quantum entanglement (Chapter 6). Indeed, because the Schrödinger equation is linear, it cannot make a selection between these results: it can only lead to a superposition of all possible outcomes, as if no macroscopic result at all had emerged from the measurement!

To solve this problem, one can then try to add a second stage, and use another measurement apparatus M' in order to determine the position of the pointer of M . But the process repeats itself and the linearity of the Schrödinger equation leads to an even stranger superposition, this time containing $S+M+M'$ and different positions of the pointers of M and M' . By recurrence, if one adds more and more measurement apparatuses, M'' etc., one creates a longer and longer chain of correlated systems, without ever selecting one single outcome for the measurement. This recurrent process M, M', M'' is called the Von Neumann regress, or chain. One can summarize this discussion by saying that “Uniqueness of the results cannot emerge from the Schrödinger equation only; in fact the equation creates a chain of coherent superposition that propagates without any end”.

A simple example may be useful to illustrate the process in a more concrete way. Assume for instance that we consider a spin $1/2$ atom, which enters into a Stern–Gerlach spin analyzer. If the initial state of the spin is parallel (or anti-parallel) to the direction of the analyzer (the direction of the magnetic field which defines the eigenstates associated with the apparatus), the wave function of the atom is deflected upwards or downwards, depending on the initial eigenvalue of the longitudinal spin. But if the initial direction of the spin is transverse, the wave function splits into two different wave packets, one pulled upwards, the other pushed downwards; again, this is an elementary consequence of the linearity of the Schrödinger equation. Propagating further, each of the two wave packets may strike a detector, with which they interact by modifying its state (as well as theirs). For instance, the incoming spin $1/2$ atoms are ionized and produce electrons; as a consequence, the initial coherent superposition now encompasses new particles. Moreover, when a whole cascade of electrons is produced as in a photomultiplier, all these additional electrons also become part of the superposition. In fact, there is no intrinsic limit in

² This linear superposition is written explicitly in equation (8.7).

what soon becomes an almost infinite chain: rapidly, the linearity of the Schrödinger equation leads to a state vector which is the coherent superposition of states including a macroscopic number of particles, macroscopic currents, and, maybe pointers or recorders printing macroscopic zeros or ones on a piece of paper! If we stick to the Schrödinger equation, there is nothing to stop this infinite regress, which has its seed in the microscopic world but rapidly develops into a big macroscopic phenomenon. Can we for instance accept the idea that, at the end, it is the brain of the experimenter (who becomes aware of the results) and therefore a human being with consciousness who enters into such a superposition?

The very notion of a brain, or consciousness, in a macroscopic superposition is neither very intuitive nor clear; no-one has ever observed two contradictory results at the same time. Would this strange situation correspond to an experimental result printed on paper and looking more or less like two superimposed slides, or a double exposure of a photograph? In practice, we know that we always observe only one single result in a single experiment; linear superpositions somehow resolve themselves before they reach us, and presumably before they become sufficiently macroscopic to involve measurement apparatuses. It therefore seems obvious³ that a proper theory should break the Von Neumann chain at some point, and stop the regress when (or maybe before) it reaches the macroscopic world. But when exactly and how precisely?

Von Neumann concludes from his analysis that, indeed, it is not possible to formulate the laws of quantum mechanics in a complete and consistent way without reference to human consciousness. He considers that the emergence of a unique result of measurement is an irreducible element of the theory. The solution to the problem requires the introduction of a special postulate: the “projection postulate” of the state vector (§1.2.2), which forces the emergence of a single result as soon as an experimenter becomes aware of the result.

The notion of the Von Neumann regress is also the source of the phenomenon of decoherence (§6.3.3). The word decoherence is usually used to refer to the initial stage of the chain, when the number of degrees of freedom involved in the process is still relatively limited. The Von Neumann chain is more general and includes this initial stage as well as its continuation, which goes on until it reaches the other extreme where it really becomes paradoxical: the Schrödinger cat.

2.2 Schrödinger's cat

The famous story of the Schrödinger cat (1935) illustrates the same problem in a different way.

³ Maybe not so obvious after all? There is an interpretation of quantum mechanics that precisely rests on the idea of never breaking this chain: the Everett interpretation, which will be discussed in §10.11.

2.2.1 The argument

The cat appears in a few lines only in the context of a more general discussion written by Schrödinger [51] and entitled “The present situation of quantum mechanics”. His words (after translation into English [52]) are: “One can even set up quite ridiculous cases. A cat is penned up in a steel chamber, along with the following device (which must be secured against direct interference by the cat): in a Geiger counter, there is a tiny bit of radioactive substance, so small that perhaps in the course of the hour, one of the atoms decays, but also, with equal probability, perhaps none; if it happens, the counter tube discharges, and through a relay releases a hammer that shatters a small flask of hydrocyanic acid. If one has left this entire system to itself for an hour, one would say that the cat still lives if meanwhile no atom has decayed. The psi-function of the entire system would express this by having in it the living and dead cat (pardon the expression) mixed or smeared out in equal parts.

It is typical of these cases that an indeterminacy originally restricted to the atomic domain becomes transformed into macroscopic indeterminacy, which can then be resolved by direct observation. That prevents us from so naively accepting as valid a ‘blurred model’ for representing reality. In itself, it would not embody anything unclear or contradictory. There is a difference between a shaky or out-of-focus photograph and a snapshot of clouds and fog banks.”

In other words, Schrödinger considers a Von Neumann chain starting from one (or a few) radioactive atomic nucleus, which can emit a photon, which is detected by a gamma ray detector, which in turn is followed by electronic amplification, which then triggers a macroscopic mechanical system that automatically opens a bottle of poison, which finally kills the cat when the decay takes place (Figure 2.1). Quantum entanglement therefore occurs at many stages. When the probability that a photon has been emitted is about $1/2$, the system reaches a state with two components of equal weight, one where the cat is alive and one where it is dead. Schrödinger considers this as an obvious impossibility (a quite ridiculous case), and concludes, therefore, that something must have happened to stop the Von Neumann chain before it went too far. Again, the challenge is to explain macroscopic uniqueness: why, at a macroscopic level, a unique result (alive, or dead, cat) emerges, while this does not happen within the linear Schrödinger equation.

The cat is of course a symbol of any macroscopic object – Einstein, in a letter to Schrödinger the same year [53], used the image of the macroscopic explosion of a barrel of powder⁴. Such an object can obviously never be in a “blurred” state containing possibilities that are contradictory (open and closed bottle, dead and alive cat, etc.). Therefore, as Schrödinger points out, his equation should not be pushed

⁴ Einstein writes “No clever interpretation will be able to transform this function Ψ into an adequate description of real things; in reality, there is nothing between exploded and non-exploded”.

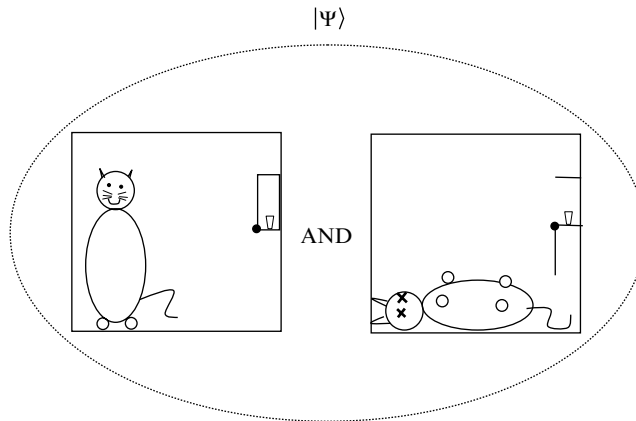


Figure 2.1. The story of the Schrödinger cat illustrates how the rules of standard quantum mechanics may lead to very strange situations, where macroscopic objects are at the same time in completely different physical states. Schrödinger imagines a cat in a box where a device is triggered by the emission of a particle by a radioactive nucleus. If the particle is emitted, the device starts and opens (or breaks) a bottle of poison, which kills the cat. If the particle is not emitted, the cat remains alive. After some time, when the probability of decay of the nucleus is $1/2$, the state $|\Psi\rangle$ of the whole system contains components associated with completely different macroscopic situations, both realized at the same time. Uniqueness of macroscopic reality would require that the “AND” in the middle of the figure should be replaced by “OR”, which cannot be explained within the Schrödinger equation. In the words of Schrödinger, this is a “ridiculous” situation.

too far and include macroscopic objects. Standard quantum mechanics is not only incapable of avoiding these paradoxical cases, it actually provides a general recipe for creating them. The logical conclusion is that some additional ingredients are needed in the theory in order to select one of the branches of the superposition, and avoid such stupid macroscopic superpositions⁵. Needless to say, the limit of validity of the linear equation does not have to be related to the macroscopic object itself: the branch selection process may perfectly take place before the linear superposition reaches the cat. But the difficult question is exactly where and when the process takes place.

The usual interpretation of quantum mechanics postulates that the only way to stop the linear propagation of the Schrödinger equation and break the chain is to perform an act of measurement. The question then becomes: does an elaborate animal such as a cat, or a primitive living creature such as a bacteria, have the

⁵ It is amusing to note that, historically, Schrödinger's name is associated to two somewhat opposite concepts. One is contained in a very powerful continuous equation of evolution that applies to all systems; the other is the cat, which is the symbol of a limit that the same equation should never reach.

intellectual abilities that are necessary to perform a measurement and resolve several Von Neumann branches into one? At what stage of evolution can a living creature perceive its own state, projecting itself onto one of the alive or dead states? Or do humans only have access to a sufficient level of introspection to become conscious of their own observations, and to reduce the state vector? Some theories take this point of view, a case in which, when the wave function includes a cat component, the animal could remain simultaneously dead and alive for an arbitrarily long period of time, a paradoxical situation indeed.

The last sentence of Schrödinger's quotation concerning photographs has often been considered as obscure. Schrödinger probably wishes to emphasize the difference between an incomplete knowledge (out-of-focus photograph) of a well defined object, and an object that is inherently not sharply defined in space (a cloud) – between an indeterminacy that is related to lack of information or an inherent indeterminacy. In other words, he is already questioning the complete character of quantum mechanics (§3).

2.2.2 Misconceptions

A common misconception is that the paradox is easily solved by just invoking decoherence (§6.3.3), which explains why it is impossible in practice to observe quantum interferences between states where a cat is alive or dead. We come back to this point in §6.3.3.b in more detail, and discuss it only briefly here. Actually, (de)coherence is irrelevant in Schrödinger's argument: the cat is actually a symbol of the absurdity of a quantum state that encompasses two incompatible possibilities in ordinary life, coherent or not. It does not change the absurdity of the final situation whether the state in question is a pure state (sensitive to decoherence) or a statistical mixture (insensitive to decoherence). Moreover, the cat itself is already part of the environment of the radioactive atom (the detector and the bottle of poison are also in this case)⁶. The chain (the tree of possibilities) starts to propagate at a microscopic level (from the radioactive atom) and continues further and further without apparent limits; the real difficulty is to stop it from reaching the macroscopic world. It certainly does not help to remark that the chain propagates even further than the cat; invoking decoherence is not answering the question, it is repeating it.

⁶ The cat itself is never in a coherent superposition of alive and dead states. When the curious state is created, the cat is already correlated with the radioactive source, the mechanical system, the bottle of poison (open or closed), the gas in the box, etc. All these components already act as an environment and produce complete decoherence. Restoring this coherence would imply an operation of putting back all of them into the same quantum state, a clearly impossible task. The propagation of decoherence further into the environment does not add anything new to the argument.

The real problem addressed by Schrödinger is how to explain the uniqueness of the macroscopic world. How can something that was indeterminate become determinate, and what is the process responsible for this? Until what point exactly does the theory remain deterministic, to become stochastic beyond this point? This question is not trivial and much more difficult to answer than just explaining the absence of macroscopic coherence⁷. Schrödinger was well aware of the properties of entanglement in quantum mechanics, a word that he introduced (and uses explicitly in the article on the cat; see §3.3.3.b); he was certainly not sufficiently naive to think that dead and alive cats can give rise to interferences. In [51], he actually never mentions the coherent or incoherent character of the superposition⁸: for him, the question is not to avoid a coherent superposition of macroscopically different states, it is to have no superposition at all⁹!

2.2.3 Modern cats

The meaning of words sometimes changes in physics. In the recent literature in quantum electronics and optics, it has become more and more frequent to use the words “Schrödinger cat (SC)” with a different meaning. Initially, the cat was the symbol of an impossibility, an animal that can never exist (a Schrödinger gargoyle?); it was the final step of a “reductio ad absurdum” reasoning, in short something that obviously has (and will) never been observed. But, nowadays, the same words are often used to describe states that are actually perfectly accessible physically, namely any coherent superposition of states that are more or less macroscopically distinguishable (the coherent character is then essential). In this new sense, Schrödinger cats have been predicted and observed in a variety of systems, for instance an ion located in two different places in a trap; they of course undergo rapid decoherence (§6.3.3) through correlation to the environment. Theoretical calculations of this decoherence are possible within the Schrödinger equation, which is used to calculate how the initial stages of the Von Neumann chain take place, and how rapidly the state vector tends to ramify into branches containing the environment.

To summarize §2.2, the paradox addresses the core of most of our difficulties with quantum mechanics: as Wigner writes [54] “measurements which leave the system object-plus-apparatus in one of the states with a definite position of the pointer cannot be described by the linear laws of quantum mechanics”. The question is

⁷ The impossibility of observing interferences between dead and alive cat states is, of course, a necessary condition for macroscopic uniqueness. But it is also far from sufficient.

⁸ The words do not appear even once in his article.

⁹ This is for instance the purpose of theories with a modified non-linear Schrödinger dynamics (§10.8): they provide equations of motion where, at a macroscopic level, all probabilities dynamically go to zero, except one that goes to 1.

then: what is exactly the process that forces Nature to break the linearity and to make its choice among the various possibilities for the results of experiments? Indeed, the emergence of a single result in a single experiment is a major issue. As Pearle expresses it concisely [14], the problem is to “explain why events occur”!

2.3 Wigner’s friend

In a theory such as standard quantum mechanics, where the observer plays such an essential role, who is entitled to play it? Wigner discusses the role of a friend, who has been asked to perform an experiment, a Stern–Gerlach measurement for instance [55]. The friend is working in a closed laboratory, so that an outside observer cannot be aware of the result before he/she opens the door (Figure 2.2). What happens just after the particle has emerged from the analyzer and when its position has been observed inside the laboratory, but is not yet known outside? For the outside observer, it is natural to consider the whole ensemble of the closed laboratory, containing the experiment as well as his friend, as the “system” to be described by a big wave function. As long as the door of the laboratory remains closed and the result of the measurement unknown, this state vector will continue to contain a superposition of the two possible results; it is only later, when the

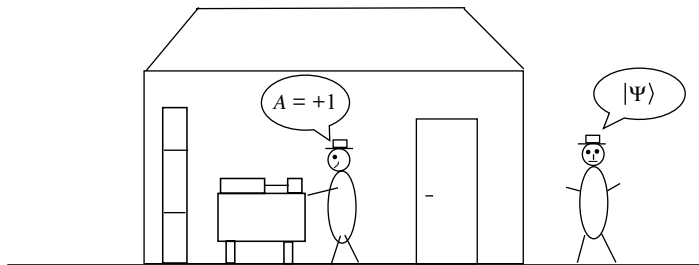


Figure 2.2. Wigner assumes that a physicist outside of a closed laboratory has a friend inside who performs a quantum experiment providing results $A = \pm 1$. When the friend observes a result, for instance $A = +1$, he describes the physical system with a state vector that includes this information and therefore has undergone state vector reduction. Nevertheless, as long as the door of the laboratory remains closed, the physicist outside should describe the ensemble of the experiment and of his friend with a state vector still containing all different possible results; for him, state vector reduction occurs only when he opens the door and becomes aware of the result. One then reaches a situation where the same physical reality is described at the same time by two different state vectors. To avoid this, one may assume that the first perception of the result by the mind of any observer (the friend in this case) reduces the state vector; the system then has a single state vector even if for some time it is not known by the experimenter outside.

result becomes known outside, that the state vector reduction should be applied. But, clearly, for Wigner’s friend who is inside the laboratory, this reasoning is just absurd! He/she will much prefer to consider that the state vector is reduced as soon as the result of the experiment is observed inside the laboratory. We are then back to a point that we have already discussed (§1.2.3), the absolute/relative character of the wave function: does this contradiction mean that we should consider two state vectors, one reduced, one not reduced, during the intermediate period of the experiment¹⁰? For a discussion by Wigner of the problem of the measurement, see [54].

An interpretation, sometimes associated with Wigner’s name¹¹, assumes that the reduction of the state is a real effect taking place when a human mind interacts with the surrounding physical world and acquires some consciousness of its state; we will come back to it in §10.1.1.b. In other words, the electrical currents in the human brain may be associated with a reduction of the state vector of measured objects by some as yet unknown physical process. Of course, in this view, the reduction takes place under the influence of the experimentalist inside the laboratory, and the question of the preceding paragraph is settled. But, even if one accepts the somewhat provocative idea of possible action of the mind (or consciousness) on the environment, this point of view does not suppress all logical difficulties: what is a human mind; what level of consciousness¹² is necessary to reduce the state, etc.?

2.4 Negative and “interaction-free” measurements

Other paradoxical situations occur with “negative measurements” or “interaction-free measurements” in quantum mechanics. These concepts may be illustrated with the following examples:

- (i) Consider a source that emits one particle with a spherical wave function (its values are independent of the direction in space). The source is surrounded by two detectors, as shown in Figure 2.3; one of the detectors, D_1 , captures the particle emitted in almost all directions, except a small solid angle Ω ; the second detector D_2 captures the particle inside this solid angle, but at a larger distance. We assume a perfect situation where both detectors have 100% efficiency.

¹⁰ Hartle considers that the answer to this question is yes [45]; see also the relational interpretation of quantum mechanics (§10.3.1).

¹¹ The title of [55] is indeed suggestive of this sort of interpretation; moreover, Wigner writes in this reference that “it follows (from the Wigner friend argument) that the quantum description of objects is influenced by impressions entering my consciousness”. At the end of the article, he also discusses the influence of nonlinearities which would put a limit on the validity of the Schrödinger equation, and be indications of life.

¹² See in §10.1.1 for a relation between consciousness and introspection (London and Bauer).

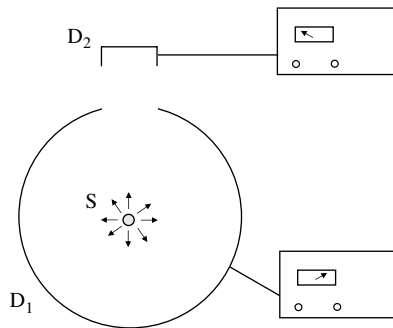


Figure 2.3. A source S surrounded by a large detector D_1 emits a particle. This detector captures the particle and provides a signal to a first recording device, except if the particle escapes through a hole corresponding to a small solid angle (in the upper direction in the figure). In this case, the particle is necessarily detected by the second detector D_2 sitting at a larger distance and providing a signal to a second recording device. In the discussion, it is assumed that the detectors are ideal and have 100% efficiency.

Suppose that the wave packet describing the wave function of the particle reaches the first detector. Two possibilities may occur:

- Either the particle is detected by D_1 . It then disappears, and the state vector is projected onto a state containing no particle and an excited detector (the first); later, the second detector D_2 will never record a particle. This first possibility occurs in most realizations of the experiment.
- Or the particle is not detected. This event occurs with a small probability, given by the solid angle subtended by the hole in the first detector seen from S , divided by 4π (we assume that the source is isotropic), but it sometimes does occur. Then the only fact that the first detector has *not* recorded the particle implies a reduction of the wave function to its component contained within Ω , implying that the second detector will *always* detect the particle later. We then have a curious situation where the probability of detection at this second detector has been greatly enhanced by a sort of “non-event” at the first detector, that is without any interaction between the particle and the first measurement apparatus. This illustrates that the essence of quantum measurement is something much more subtle than the often invoked “unavoidable perturbations of the measurement apparatus” (Heisenberg microscope, etc.); state vector reduction may take place even when the interactions play no role in the process.

Of course, if one assumes that, initially, the emitted particle already had a well-defined direction of emission, then the interpretation of this thought experiment becomes straightforward: detector D_2 records only the fraction of the particles

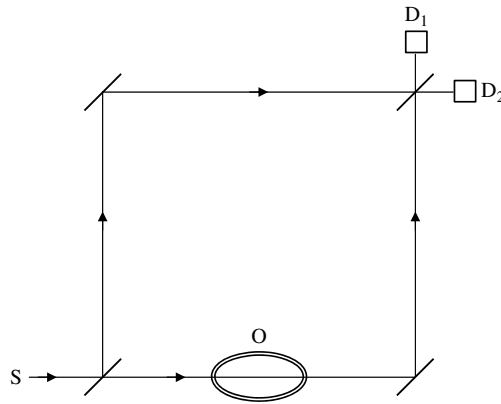


Figure 2.4. A source S emits a series of particles, one by one, into an interferometer. The path differences are adjusted in such a way that all particles reach detector D_1 and no particle reaches D_2 . When an opaque object O is inserted to block one path for the particle, the destructive interference does not occur anymore and particles are sometimes recorded at D_2 . In such an event, it seems that the presence of the object has been measured with certainty, while no interaction whatsoever with the object is involved since the detected particle necessarily went through the upper path of the interferometer (otherwise it would have been absorbed).

that were emitted from the beginning in its direction. It is then not surprising that no interaction at all occurred with detector D_1 . But standard quantum mechanics postulates that this well-defined direction of emission does not exist before any measurement. Assuming that it does would actually amount to introducing additional variables to quantum mechanics, a subject that we will discuss in much more detail below (Chapter 3 and §10.6), but which does not belong to the standard interpretation of quantum mechanics.

- (ii) Consider now Figure 2.4, with a Mach–Zehnder interferometer into which a source emits a series of particles, one by one, which are registered at detectors D_1 and D_2 . Following [56], we assume that the path differences are adjusted to create a destructive interference effect at detector D_2 ; no particle can then reach this detector and all particles are necessarily detected in D_1 (again we assume ideal detectors with 100% efficiency). Now, an opaque object O is inserted in the lower path of the interferometer, which cancels the destructive interference effect and allows some particles to reach D_2 . If we assume that the two input and output beam splitters have a 50% reflectivity, this event happens one time out of four. What happened in such an event? Since the particle was not blocked by the object, it seems that it necessarily went through the upper path of the interferometer, meaning that it never interacted with the object. But, if the object had not been inserted, it could never have reached D_2 ! The net

result is that the sole observation of a detection at D_2 reveals the presence of an object, during an event that apparently excludes any interaction with it. This phenomenon is called “interaction-free measurement” process¹³. It illustrates that, as long as the localization of the particle has not been measured, one should not attribute any position or single trajectory to it in standard quantum mechanics. The positive measurement in D_2 “cancels” at a distance, so to say, the effects of the interaction between the particle and the object, which ends up totally unaffected in this particular experiment¹⁴.

For refinements of the ideas and experiments related to interaction-free measurements, see [57, 58]. Hardy has proposed a version involving a double interferometer, one for an electron and one for a positron, and the mutual annihilation of these particles [59].

It has also been pointed out that negative measurements can be useful in the context of quantum cryptography (§7.2). The idea proposed in [60] is that two remote partners, Alice and Bob, make random choices between two orthogonal photon polarizations; Alice sends a photon with the polarization chosen locally, which Bob will send back with a mirror to Alice only if its polarization is different from his own local random choice. The whole setup contains an interferometer that is designed so that, if the photon comes back from Bob to Alice, the probability that Alice will observe a photon at detector D_1 vanishes due to a destructive interference effect similar to that used in example (ii) above. Now, by selecting only the events where Alice observes the photon at D_1 (then necessarily Bob does not detect the particle), one automatically selects events where the two remote random choices made by Alice and Bob turn out to be the same. If Alice and Bob communicate to each other the result of each run of the experiment (which detector has clicked or not), but keep secret their choice of polarization, by writing down their binary local choices for the selected events they progressively build a shared secret key. The remarkable feature of this scheme is that the events used for secret transmission are actually those where the photon has remained confined within Alice’s apparatus¹⁵: it has not been transmitted from Alice to Bob, in perfect analogy with the particle

¹³ Even if the effect requires that the test particles *may* interact with the object, since clearly the effect would not take place if the object was completely transparent to the particle.

¹⁴ In other words, the measurement process selects a component of the state vector that has propagated in the upper path of the interferometer only. But this event is made possible by the fact that, while propagating in the lower path, the Schrödinger wave has been absorbed. The curious thing is that the particle and this component of the wave seem to be dissociated in the process, since they propagate in two different arms at an arbitrary distance from each other.

¹⁵ The useful events for building the key are those where Bob did not return the Schrödinger wave to Alice’s location, but without absorbing the particle. After Alice and Bob choose identical polarizations, the state vector develops a component where the particle propagates to Bob’s site, so that it can be measured along the transmission line or at Bob’s site, but this component vanishes when Alice observes the particle at D_1 .

in example (i) above (since the detection of the particle meant that the particle had not taken the path containing the object). Because the particle does not follow the transmission line between the two sites in the events retained for building the secret key, the flux of particles along this line contains no information whatsoever on the random choices made by Alice and Bob; hence an enhanced security.

2.5 A variety of points of view

The following quotations may be useful to get an idea¹⁶ of the variety of interesting points of view that have been expressed since the appearance of quantum mechanics.

Copenhagen interpretation:

(i) Bohr ([21], 2nd edition, page 204 and [61]): “There is no quantum world. There is only an abstract physical description. It is wrong to think that the task of physics is to find out how Nature is. Physics concerns what we can say about Nature”. Or, similarly: “there is no quantum concept” [62].

Concerning physical phenomena: “one may strongly advocate limitation of the use of the word *phenomenon* to refer exclusively to observations obtained under specified circumstances, including an account of the whole experiment” [46].

He also defines the purpose of physics in the following way [63, 64]: “Physics is to be regarded not so much as the study of something a priori given, but rather as the development of methods of ordering and surveying human experience. In this respect our task must be to account for such experience in a manner independent of individual subjective judgement and therefore objective in the sense that it can be unambiguously communicated in ordinary human language”.

While quantum mechanics is often (including nowadays) considered as a theory that is completely new with respect to classical physics (a revolution), Bohr prefers to see it as its logical continuation. He considers [65, 66] that, when introducing quantum mechanics, “the problem with which the physicists were confronted was to develop a rational generalization of classical physics¹⁷, which would permit the harmonious incorporation of the quantum of action”. For this purpose, the “principle of correspondence” should be used [67]: “The correspondence principle expresses the tendency to utilize in the systematic development of the quantum theory every feature of the classical theories in rational transcription appropriate to the fundamental contrast between the postulates and the classical theories”. It is therefore natural that, on many occasions, he emphasized the importance of classical concepts to give

¹⁶ With, of course, the usual proviso: short quotations taken out of their context may, sometimes, give a superficial view on the position of their authors.

¹⁷ Similarly, Bohr would probably have seen Einsteinian relativity as a rational generalization of classical electromagnetism (Maxwell equations).

a proper meaning to the formalism of quantum mechanics; in particular, classical concepts are indispensable to describe any process of measurement.

Bohr actually wrote many texts on quantum mechanics, sometimes rather philosophically oriented [1, 68]; we give more quotations in §3.3.2. In his famous Como lecture (1927), he introduced the notion “complementarity”, the relation between contradictory attributes of the same object in quantum mechanics, and then extended it far beyond, including other sciences than physics (see [68] or for instance [46] where he mentions biology¹⁸, sociology, and psychology; see also §7.2 of [21] or Chapter 4 of [50]).

(ii) Born in 1926 (page 804 of [34]): “The motion of particles conforms to the laws of probability, but the probability itself is propagated in accordance with the laws of causality”.

Or, shortly later in [69]: “The quantum theoretical description ... does not answer ... the question of where a certain particle is at a given time ... This suggests that quantum mechanics only answers properly put statistical questions, and says nothing about the course of individual phenomena. It would then be a singular fusion of mechanics and statistics”.

(iii) Heisenberg [21, 70]: “But the atoms or the elementary particles are not real; they form a world of potentialities or possibilities rather than one of things and facts”.

In *Physics and Philosophy* [70] (Chapter V): “Natural science does not simply describe and explain nature; it is a part of the interplay between nature and ourselves; it describes nature as exposed to our method of questioning”.

Chapter III of [70] is entitled “The Copenhagen interpretation of quantum theory”, and he writes in this chapter: “We cannot completely objectify the result of an observation, we cannot describe what ‘happens’ between this observation and the next”. Later, he adds: “Therefore, the transition from the ‘possible’ to the ‘actual’ takes place during the act of observation. If we want to describe what happens during an atomic event, we have to realize that the word ‘happens’ can apply only to the observation, not to the state of affairs between two observations”.

He concludes this chapter with: “... the measuring device has been constructed by the observer, and we have to remember that what we observe is not nature in itself but nature exposed to our method of questioning. Our scientific work in physics consists in asking questions about nature in the language that we possess and trying to get an answer from experiment by the means that are at our disposal.

¹⁸ At the end of this article, he discusses the “complementary mode of description” and illustrates its generality by writing: “An example is offered by biology where mechanistic and vitalistic arguments are used in a typically complementary manner. In sociology too such dialectics may be often useful, particularly in problems confronting us with the study and comparison of human cultures ...”.

In this way, quantum theory reminds us, as Bohr has put it, of the old wisdom that when searching for harmony in life one must never forget that in the drama of existence we are ourselves both players and spectators. It is understandable that in our scientific relation to nature our own activity becomes very important when we have to deal with parts of nature into which we can penetrate only by using the most elaborate tools”.

(iv) Jordan (as quoted by Bell in [71]): “observations not only disturb what has to be measured, they *produce* it. In a measurement of position, the electron is forced to a decision. We compel it to assume a definite position; previously it was neither here nor there, it had not yet made its decision for a definite position ...”.

(v) Landau and Lifshitz, in the first section of their book on quantum mechanics [72]: “The possibility of a quantitative description of the motion of an electron requires the presence also of physical objects which obey classical mechanics to a sufficient degree of accuracy. If an electron interacts with such a ‘classical object’ the state of the latter is, generally speaking, altered ... In this connection the ‘classical object’ is usually called *apparatus*, and its interaction with the electron is spoken of as *measurement*. However, it must be most decidedly emphasized that we are here not discussing a process of measurement in which the physicist-observer takes part. By *measurement* in quantum mechanics, we understand any process of interaction between classical and quantum objects, occurring apart from and independently of any observer. The importance of the concept of measurement in quantum mechanics was elucidated by N. Bohr”.

(vi) Dirac, page 7 of [39]: “*The only object of theoretical physics is to calculate results that can be compared with experiment*, and it is quite unnecessary that any satisfying description of the whole course of the phenomena should be given”.

Critics of the Copenhagen interpretation:

(vii) Schrödinger [73]: “the world is given to me only once, not one existing and one perceived. Subject and object are only one. The barrier between them cannot be said to have broken down as a result of recent experience in the physical sciences, for this barrier does not exist”. In §2.2, we give more quotations by Schrödinger.

(viii) Einstein, in a letter to Schrödinger in 1928 [74, 75]: “The Heisenberg–Bohr tranquilizing philosophy – or religion? – is so delicately contrived that, for the time being, it provides a gentle pillow for the true believer from which he cannot very easily be aroused.”

In 1936 [76]: “The Ψ function does not in any way describe a condition which could be that of a single system; it relates rather to many systems, to an ‘ensemble of systems’ in the sense of statistical mechanics ... If the function Ψ furnishes only

statistical data concerning measurable magnitudes ... the reason lies ... in the fact that the function Ψ does not, in any sense, describe the condition of *one* single system”.

(ix) de Broglie [77]: “The interpretation of wave mechanics by Bohr and Heisenberg has many consequences opening new philosophical perspectives. The corpuscle is no longer a well-defined object within the frame of space and time; it is only an ensemble of potentialities to which probabilities are assigned, it is only an entity manifesting itself to us in a fugitive way, sometimes taking an aspect, sometimes another. Professor Bohr, who is in a way the Rembrandt of contemporary physics, since he sometimes shows a clear taste for ‘chiaroscuro’, said that the corpuscles are ‘unsharply defined individuals within finite space-time limits’.”

More recent comments:

(x) Bell [47], describing “modern” quantum theory (Copenhagen interpretation) and its relations with cosmology: “it never speaks of events in the system, but only of outcomes of observations upon the system, implying the existence of external equipment¹⁹” (how, then, do we describe the whole universe, since there can be no external equipment in this case?).

“The problem is this: quantum mechanics is fundamentally about observations. It necessarily divides the world into two parts, a part which is observed and a part which does the observing. The results depend on how this division is made, but no definite prescription for it is given. All we have is a recipe which, because of practical human limitations, is sufficiently unambiguous for practical purposes”.

See also his text “Against measurement” where he discusses and criticizes various presentations of the orthodox interpretation [78].

(xi) Mermin [9], summarizing the “fundamental quantum doctrine” (orthodox interpretation): “the outcome of a measurement is brought into being by the act of measurement itself, a joint manifestation of the state of the probed system and the probing apparatus. Precisely how the particular result of an individual measurement is obtained – Heisenberg’s transition from the possible to the actual – is inherently unknowable”.

(xii) Shimony [10]: “According to the interpretation proposed by Bohr, the change of state is a consequence of the fundamental assumption that the description of any physical phenomenon requires reference to the experimental arrangement”.

(xiii) Rosenfeld [79], speaking of the orthodox interpretation: “the human observer, whom we have been at pains to keep out of the picture, seems irresistibly to intrude into it ...”.

¹⁹ One could add “and of external observers”.

(xiv) Gottfried [80] (page 188 of first edition): “The reduction postulate is an independent axiom ... The outcome of these considerations is that quantum mechanics cannot give a complete description of the physical world because there must exist systems (called ‘conscious’ by Wigner) that are beyond the theory’s power of description, i.e. that cannot be incorporated into the part of the world that we treat with the Schrödinger equation”.

(xv) Stapp [42] “The interpretation of quantum theory is clouded by the following points: (1) Invalid classical concepts are ascribed fundamental status; (2) The process of measurement is not describable within the framework of the theory; (3) The subject–object distinction is blurred; (4) The observed system is required to be isolated in order to be defined, yet interacting to be observed”.

Or, as cited by Bell in [47]: “How can a theory which is *fundamentally* a procedure by which gross macroscopic creatures, such as human beings, calculate predicted probabilities of what they will observe under macroscopically specified circumstances ever be claimed to be a complete description of physical reality?” (the completeness of quantum mechanics will be discussed in Chapter 3).

(xvi) Leggett discusses the Copenhagen interpretation [81] in these terms: “... the formalism of QM which goes under the name of the Copenhagen interpretation (though it should probably more correctly be called the Copenhagen non-interpretation, since its whole point is that any attempt to interpret the formalism in intuitive terms is doomed to failure) ... while denying that microscopic objects (electrons, photons, atoms, ...) necessarily have definite properties in the absence of observation, emphatically asserts (or at least implies) that macroscopic objects (counters, cats, etc.) do have such properties at all times, whether they are observed or not. This insistence on the necessity of drawing a sharp line ... between the microscopic world and the macroscopic one of everyday life (including apparatus) is a pervasive theme in the writings of Niels Bohr ...”.

In [82], he comments “The real trouble only starts when we take seriously the fact that the measuring apparatus ... is itself a physical system made up of atoms and electrons, and therefore should in principle be describable in quantum-mechanical terms. It should therefore be legitimate to ask what happens if, instead of treating measurement as something quite extraneous to the ordinary behavior of physical systems, we treat it as merely a particular type of physical process and describe it by the linear time-dependent Schrödinger equation”. And then, in [83], he comments: “Within the conventional interpretation of quantum mechanics, a system does not possess definite properties until we, as it were, force it to declare them by carrying an appropriate measurement. But is this the only possible interpretation? ... it is perfectly possible that at a deeper level systems do, in fact, have objective objective properties, whether or not anyone is measuring them ... The apparently random

outcomes predicted by the quantum formalism would then simply be due to our ignorance of the details of a deeper level description²⁰”.

(xvii) Van Kampen [84] has proposed the following interesting caveat: “Whoever endows the state vector with more meaning than needed for computing observable phenomena is responsible for the consequences!” In Appendix A, we discuss how two other references [85, 86] make good use of this warning.

Present situation:

As one can guess from the above quotations, even nowadays no general consensus has been reached among physicists concerning the precise meaning of the state vector or wave function. They agree on the formalism and how to use the state vector in practice. For all experiments realized so far, a pragmatic choice between the two postulates can be left to the judgment of the physicist; to make practical predictions concerning an experiment, common sense is sufficient (cf. §10.1.1), so that problems related to the foundations of the theory may be put aside. Nevertheless, it would be preferable to have well-defined mathematical laws than just physically reasonable recipes! It is therefore not surprising that, when the question comes to discussing the physical meaning of the mathematical objects and the interpretation of the theory, the debate soon re-appears, and may even sometimes become passionate. Moreover, even those who claim to be in perfect agreement with orthodox quantum theory will use, in fact, a variety of nuances (and even sometimes contradictory points of view and inconsistencies) when they are asked to explain in detail their point of view.

To summarize, the orthodox status of the wave function is indeed a subtle mixture between different concepts concerning reality and the knowledge that we have of this reality. Do the fantastic achievements of standard quantum theory mean that it is the ultimate and most precise description of a quantum system that physics will ever provide? Is quantum theory compatible with realism, or does it require positivism? The questions are not settled. Interestingly Bohr himself is generally considered more as a realist²¹ than a positivist or an operationalist [21]. In the words of Jammer ([50], p. 157): “Bohr, as Von Weizsäcker [87] emphasized, never rejected the notion of reality, he only modified it”. If asked about the relations between the wave function and reality, Bohr would probably have said that the wave function is indeed a useful tool, but that the concept of reality cannot properly be defined at its level only; it has to include all macroscopic measurement apparatuses that

²⁰ Section 3 gives a much more detailed account of the discussions on the completeness (or incompleteness) of quantum mechanics.

²¹ See for instance in §3.3.2 how he considers that physical reality can be defined without ambiguity. Bohr accepts the notion of reality provided it is properly defined (with a full description of the experiment).

are used to have access to microscopic information (we come back to this point in more detail in §3.3).

2.6 Unconvincing arguments

We have already emphasized that the invention of the Copenhagen/standard interpretation of quantum mechanics has been, and remains, one of the biggest achievements of all science. One can admire even more, in retrospect, how early its founders conceived it, at a time when experimental data were relatively scarce. Since that time, numerous ingenious experiments have been performed, precisely with the hope of seeing the limits of this interpretation; nevertheless, until now, not a single fact has disproved the theory. It is really a wonder of pure logic that has allowed the early emergence of such an intellectual construction.

This being said, one has to admit that, in some cases, the brilliant authors of this construction may sometimes have gone too far, pushed by their great desire to convince. For instance, authoritative statements have been made concerning the absolute necessity of the orthodox interpretation which now, in retrospect, seem exaggerated. According to these statements, the orthodox interpretation would give the only possible ultimate description of physical reality; no finer description would ever become possible. From this perspective, the fundamental probabilistic character of microscopic phenomena should be considered as a proven fact, a rule that should be carved into marble and accepted forever by scientists. But, now, we know that this is not necessarily true: yes, one may prefer the orthodox interpretation for various reasons, but this is not the only choice allowed by pure logic. Other interpretations are still perfectly possible; determinism in itself is not disproved at all²². As discussed for instance in [9], and initially clarified by Bell [5, 6] and Bohm [7, 8], the “impossibility proofs” put forward by the proponents of the Copenhagen interpretation are logically unsatisfactory for a simple reason: they arbitrarily impose conditions that may be relevant to quantum mechanics (linearity), but not to the theories they aim to dismiss – any theory with additional variables such as the Bohm theory, for instance. Because of the exceptional scientific stature of the authors of the impossibility theorems, it took a long time for the physics community to realize that the theorems were irrelevant; now that this is more widely recognized, the plurality of interpretations is more easily accepted.

²² Provided one accepts non-locality, see Chapter 4.

3

The theorem of Einstein, Podolsky, and Rosen

More than 70 years after its publication, the article by Einstein, Podolsky, and Rosen (EPR) [88] is still cited hundreds of times every year in the literature; this is a very exceptional case of longevity for a scientific article! There is some irony in this situation since, for a long time, the majority of physicists did not pay much attention to the EPR reasoning. They considered it as historically interesting, but with no precise relevance to modern quantum mechanics; the argument was even sometimes completely misinterpreted. A striking example is given in the Einstein–Born correspondence [89] where Born, even in comments that he wrote after Einstein’s death, clearly shows that he never completely understood the nature of the objections raised by EPR. Born went on thinking that the point of Einstein was a stubborn rejection of indeterminism (“look, Albert, indeterminism is not so bad!”), while actually the major concern of EPR was locality and/or separability (we come back later to these terms, which are related to the notion of space-time). If giants like Born could be misled in this way, no surprise that, later on, many others made similar mistakes!

This is why, in what follows, we will take an approach that may look elementary, but at least has the advantage of putting the emphasis on the logical structure of the arguments and their generality. Doing so, we will closely follow neither the historical development of the ideas nor the formulation of the original article¹, but rather will emphasize the generality of the EPR reasoning. For a more precise historical report, see Chapter 6 of [50], or [93], as well as references contained therein; [94] relates the circumstances under which the EPR article was written, and was sent not only to *Physical Review* but also to a

¹ The published version of the EPR article was written by Podolsky. In later comments on the subject, Einstein gave the impression that he thought that the essence of the argument could have been expressed in a simpler way; see for instance a letter from Einstein to Schrödinger [90] (“the writing is not really what I was expecting: what is essential is sort of buried under erudition”); see also [76], [91], and [92], where Einstein explains the argument in his own terms.

newspaper (*New York Times*), which published a report² – an initiative that Einstein disapproved.

3.1 A theorem

One often speaks of the “EPR paradox”, although the word “paradox” is not completely appropriate in this case. For Einstein, the basic motivation was not to invent paradoxes or to entertain colleagues inclined to philosophy, it was to build a strong logical reasoning that, starting from well-defined assumptions (roughly speaking: locality and some form of realism), would lead ineluctably to a clear conclusion (quantum mechanics is incomplete, and even: physics is deterministic³). To emphasize this logical structure and generality, here we will speak of the “EPR theorem”⁴, which formally could be stated as follows:

Theorem: If the predictions of quantum mechanics are correct (even for systems made of remote correlated particles) and if physical reality can be described in a local (or separable) way, then quantum mechanics is necessarily incomplete: some “elements of reality”⁵ exist in Nature that are ignored by this theory.

The theorem is certainly valid; it has been scrutinized by many scientists who have found no flaw in its derivation. Indeed, the logic which leads from the assumptions to the conclusions is perfectly sound. It would therefore be an error to repeat (a classical mistake!) “the theorem was shown by Bohr to be incorrect” or, even worse, “the theorem is incorrect since experimental results are in contradiction with it”⁶. Bohr himself, of course, did not make the error: in his reply to EPR [41], he explained why he thought that the assumptions on which the theorem is based were not relevant to the quantum world, which made it inapplicable to a discussion on quantum mechanics. More precisely, he used the word “ambiguous” to characterize these assumptions, but never claimed that the reasoning is faulty (for more details, see §3.3.2). A theorem that is not applicable in a particular case is not necessarily incorrect: theorems of Euclidean geometry are not wrong, or even useless, because one can also build a consistent non-Euclidean geometry! Concerning possible contradictions with experimental results we will see that, in a sense, they make a theorem even more interesting, mostly because it can then be used within a “*reductio ad absurdum*” reasoning.

² This reports ends with comments by E. Condon (Princeton University) which show that he also did not realize the exact point of the EPR article; maybe he did not have the time to read the article in detail before the newspaper was published.

³ Born’s mistake, therefore, was to confuse assumptions and conclusions.

⁴ Einstein himself nevertheless sometimes used the word “paradox”, for instance in §5 of [76].

⁵ These words are carefully defined by the authors of the theorem; see the beginning of §3.3.

⁶ The contradiction in question occurs through the Bell theorem (which is therefore sometimes criticized for the same reason), which was introduced as a continuation of the EPR theorem.

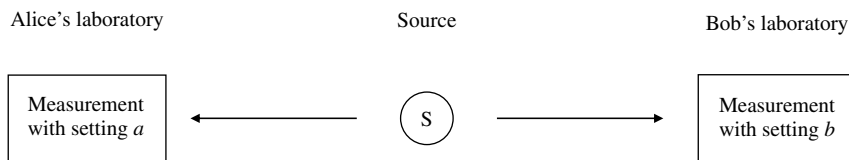


Figure 3.1. A source S emits two particles, which then propagate in space and reach different regions of space, where Alice and Bob perform measurements with them in their respective laboratories; a and b are the settings of the two corresponding measurement apparatuses (for instance, orientations of Stern–Gerlach magnets).

Good texts on the EPR argument are abundant; for instance, a classic is the little article by Bell [71]. Another excellent introductory text is, for instance, [48], which contains a complete description of the EPR thought experiment in a particular case (two settings only are used) and provides a general discussion of many aspects of the problem. For a detailed list of references, see for instance [95]. The basic scheme considered is summarized in Figure 3.1: a source S emits two correlated particles, which propagate towards two remote regions of space where they undergo measurements; the type of these measurements are defined by “settings”, or “parameters”⁷ a in the first region, b in the second (typically orientations of Stern–Gerlach analyzers), which are the free choice of the experimentalists; in each region, a result is obtained, which can take only two values symbolized by ± 1 in the usual notation. Finally, a crucial assumption is made: for the particular experimental arrangement considered, every time both settings are chosen to be the same value, the results of both measurements are always the same (perfect correlations).

Here, rather than trying to paraphrase the good texts on EPR, we will intentionally take a different presentation, based on an analogy, a sort of a parable. Our purpose is to emphasize an important feature of the argument: the essence of the EPR reasoning is actually nothing but what is usually called “the scientific method”, in the sense discussed by Francis Bacon and Claude Bernard. For this purpose, we will leave pure physics for botany! Indeed, in both disciplines, one needs rigorous scientific procedures in order to prove the existence of relations and causes, which is precisely what we want to do.

3.2 Of peas, pods, and genes

Inferring the properties of microscopic objects from macroscopic observations requires a combination of ingenuity, in order to design meaningful experiments, with a good deal of logic, in order to deduce these microscopic properties from the macroscopic results. Because it is impossible to observe an electron or even

⁷ Depending on the context, we will use the words “settings” and “parameters” indifferently.

a macromolecule with the naked eye, or to take it in one's hand, some abstract reasoning is necessary. The scientist of past centuries who, like Mendel, was trying to determine the genetic properties of plants, had exactly the same problem: he did not have access to any direct observation of the DNA molecules, so that he had to base his reasoning on adequate experiments and on the observation of their macroscopic outcome. In our parable, the scientist grows plants from peas, and later observes the color of flowers (the “result” of the measurement, +1 for red, -1 for blue) as a function of the conditions (temperature for instance) in which the peas are grown – these conditions correspond to the “settings” a or b , also called “experimental parameters” which determine the nature of the measurement. The basic purpose is to infer the intrinsic properties of the peas (the EPR “element of reality”) from these observations.

3.2.1 Simple experiments: no conclusion yet

It is clear that many external parameters such as temperature, humidity, amount of light, etc. may influence the growth of vegetables and the color of a flower; in a practical experiment, it seems therefore very difficult to identify and control all the relevant parameters with certainty. Suppose that one observes that the flowers growing in a series of experiments are sometimes blue, sometimes red; the reason behind these variations may therefore be some irreproducibility of the conditions of the experiment, but it can also be something more fundamental. In more abstract terms, a completely random character of the observed result of the experiments may originate either from the fluctuations of uncontrolled external perturbations, or from some intrinsic property that the measured system (the pea) initially possesses. It may even result from the fact that the growth of a flower (or, more generally, life?) is a fundamentally indeterministic process – needless to say, all three reasons can be combined in any complicated way. Transposing the issue to quantum physics leads to the following formulation of the question: are the results of the experiments random because of the influence of some uncontrolled fluctuating process acting somewhere (fluctuations in the macroscopic apparatus, fluctuations of some microscopic property of the measured particles, etc.)? Or are they fundamentally random because they are the result of an irreducible quantum process?

The scientist may repeat the “experiment” a thousand times and even more: as long as the results are always totally random, there is no way to decide which interpretation should be selected; it is just a matter of personal taste. Of course, philosophical arguments may be built to favor or reject one of them, but from a pure scientific point of view, at this stage, there is no compelling reason to make one choice or another. Such was the situation of quantum physics before the EPR argument.

3.2.2 Correlations: causes unveiled

The stroke of genius of EPR was to realize that studying quantum correlations could lead to a big step forward in the analysis. They exploit the crucial assumption mentioned above: when the choices of the settings are the same, the observed results are always identical.

3.2.2.a Same measurement parameters

In our botanical analogy, we assume that the botanist observes correlations between colors of flowers under appropriate conditions. Peas come together in pods, so that it is possible to grow peas taken from the same pod and observe their flowers in remote places. When no special care is taken to give equal values to the experimental parameters (temperature, etc.), or when peas are taken from different pods, nothing special is observed. But when the two peas come from the same pod and when the parameters for the growth of each of them are chosen to the same values, then the colors are systematically the same (both remain random from one pod to the next, but they always take the same value).

What can we then conclude? Since the peas grow in remote places, there is no way that they can be influenced by any single uncontrolled fluctuating phenomenon, or that they can somehow influence each other in the determination of the colors. If we believe that causes always act locally, and that perfect correlations cannot appear by pure chance (without any cause), we are led to the following conclusion: the only possible explanation of the common color is the existence of some common property of both peas, which determines the color⁸. It may be difficult to detect the property in question directly, since it is presumably encoded inside some tiny part of a biological molecule, but the property exists and is sufficient to determine the results of the experiments.

This is the essence of the argument, and we could stop at this point. Nevertheless, for completeness, let us make every step of the EPR reasoning even more explicit. The key idea is that the nature and the number of “elements of reality” associated with each pea cannot vary under the influence of some remote experiment, performed on the other pea. Let us first assume that the two experiments are performed at different times: one week, the experimenter grows a pea; the next week another pea from the same pod, in exactly the same conditions of temperature, humidity, etc. We then assume that perfect correlations of the colors are always observed, without any special influence of the delay between the experiments. Just after completion of the first experiment (observation of the first color), but still before the second

⁸ The fact that the correlations disappear if the parameters are no longer set to the same value shows that the color is a function of both this common property and of the local parameters of the experiment (settings).

experiment, the result of that future experiment has a perfectly determined value; therefore, there must already exist one element of reality attached to the second pea that corresponds to this certainty. Clearly, it can not be attached to any other object than the pea, for instance one of the measurement apparatuses; this is because the observation of perfect correlations only arises when making measurements with peas taken from the same pod. Symmetrically, the first pod also had an element of reality attached to it, which ensured that its measurement would always provide a result that coincides with that of the future measurement. We can assume that the elements of reality associated with both peas are coded in some genetic information, and that the values of the codes are the same for all peas coming from the same pod; but other possibilities exist and the precise nature and mechanism involved in the elements of reality do not really matter here. The important point is that, since these elements of reality cannot appear by any action at a distance, they necessarily also existed before any measurement was performed – presumably even before the two peas were separated.

It seems difficult not to agree that the method which led to these conclusions is indeed the scientific method; no tribunal would believe that, in any circumstance, perfect correlations could be observed in remote places by pure chance, without being the consequence of some common characteristics shared by both objects. Such perfect correlations can then only reveal the initial common value of some variable attached to them, which is in turn a consequence of some fluctuating common cause in the past (a random choice of pods in a bag for instance).

3.2.2.b *Different measurements parameters*

Now, let us consider any pair of peas, when they are already spatially separated, but before the experimentalist decides what type of measurements they will undergo (values of the parameters, delay or not, etc.). We know that, if the decision turns out to favor measurements with exactly the same parameter, perfect correlations will always be observed; this is true for any local choice of the parameters, provided they are the same at both places. Since elements of reality cannot appear or change their values depending of experiments that are performed in a remote place, the two peas necessarily carry some elements of reality that completely determine the color of the flowers for any experimental condition – not only for the conditions that turn out to be chosen for the experiment. Any theory ignoring these elements of reality is necessarily incomplete.

To take another analogy, let us for instance assume that the peas are replaced with elaborate automata, using the most elaborate technology available, able to measure the external parameters such as temperature, etc. and to process them

with powerful computers⁹ to calculate the results ± 1 from the parameters. All this computing power is actually useless to simulate the predictions of quantum mechanics for remote experiments, unless the memory of each computer contains initially a common random number. Because each automaton does not have access to the remote parameters, the only way to reproduce quantum mechanics is to feed the memory of both computers with the same random number, and to program both for calculating in the same way all the possible results as a function of the local parameters and the common random number.

3.2.2.c Summary

To summarize, the reasoning shows that each result of a measurement may be a function of two kinds of variables¹⁰:

- (i) intrinsic properties of the peas, which they carry along with them;
- (ii) the local parameters of the experiment (temperature, humidity, etc.); because correlations disappear when the parameters are different, a given pair that turned out to provide two blue flowers with one choice of common experimental conditions could have provided a pair of red flowers with another choice.

We may also add that:

- (iii) the results are well-defined functions of these variables, which means that no fundamentally indeterministic process takes place in the experiments. The only source of randomness is the choice of the initial pair.
- (iv) when taken from its pod, a pea cannot “know in advance” to which sort of experiment it will be submitted, since the decision may not yet have been made by the experimenters; when separated, the two peas therefore have to take with them all the information necessary to determine the color of flowers for any kind of experimental conditions.

In other words, each pea carries with it as many elements of reality as necessary to provide “the correct answer”¹¹ to all possible questions it might be asked in the future.

⁹ We are assuming here that the computers are not quantum computers (if quantum computers can ever be built, which is another question).

¹⁰ In Bell’s notation, the A functions depend on λ as well as on a .

¹¹ In §13 of [51], Schrödinger offers a comparison with a “schoolboy under examination” who may be asked two different questions by a teacher. During many tests, it appears that “the pupil will correctly answer the first question that is put to him. From that it follows that in every case he knows the answer to both questions”, even if afterwards he does not necessarily answers the second correctly. “No school principal will judge otherwise”.

3.3 Transposition to physics

We now come back to microscopic physics, as in the original EPR argument.

3.3.1 The EPR argument for two correlated microscopic particles

Historically, EPR introduced their arguments for two correlated particles and measurements of their position and momentum, with continuous results. It is nevertheless convenient to use a version of the argument, often called EPRB, that involves discrete results and spins, initially introduced by Bohm [96].

3.3.1.a Assumptions

We consider two spin $1/2$ particles, which are emitted by one source S in a singlet state¹² where their spins are correlated, and then propagate towards two different regions of space, where they undergo the measurement of their spin component along a direction defined by angle a on the left, b on the right (Figure 3.2). We call Alice and Bob the two operators who perform these experiments in two remote laboratories, separated by an arbitrary distance. Alice makes a free choice of the direction a , which defines her “type of measurement”, and can only obtain result $+1$ or -1 whatever type of measurement has been chosen; similarly, Bob chooses direction b arbitrarily and obtains one of the results $+1$ or -1 . In the EPR thought experiment, it is assumed that the two spins interact only with the measurement apparatuses, with no proper evolution of the singlet state. Standard quantum mechanics then predicts (§4.1.1) that the distance and times at which the spin measurements are performed is completely irrelevant for calculating the probabilities of the various combined results. If for instance a and b are chosen equal (parallel directions of measurements), the prediction is that the results are always opposite, even if the measurements take place in very remote places. This remains true whatever common choice of $a = b$ is made, even by two experimenters who operate independently in the two regions of space and make a random choice at the last moment (after the emission of the pair of particles).

The starting point of EPR is to assume that the predictions of quantum mechanics concerning the results of experiments are correct. More specifically, the reasoning assumes that the perfect correlations predicted by this theory are always observed, whatever the distance between the experimental apparatuses may be. In the parable of the peas, the red/blue colors are obviously the analogue of the two results ± 1 ,

¹² In this chapter, we do not need to know the precise definition of a singlet spin state. This definition is given in §4.1.1, as well as the quantum probabilities of the results associated with any values of a and b . For equal values of a and b , these probabilities give perfect correlations, which is the only result we need to know to understand the EPR reasoning. See also see the Annex, equation (11.69).

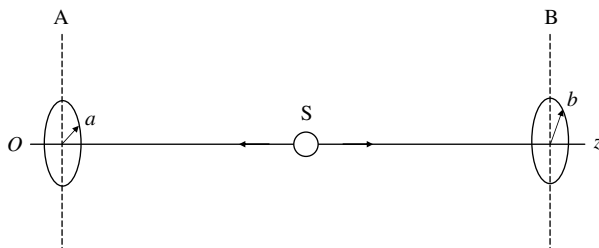


Figure 3.2. Scheme of an EPRB experiment. A source S emits pairs of particles in a singlet spin state. The particles propagate along direction Oz towards two remote regions of space A and B , where Stern–Gerlach magnets are used to measure the components of their spins along directions in the plane perpendicular to Oz . For the first particle, the direction is defined by angle a , for the second by angle b . Each measurement provides result $+1$ or -1 , and correlations between these results are measured when the experiment is repeated many times.

the experimental parameters (temperature, etc.) are the analogous to the orientation of these apparatuses. We have assumed that the same colors are always observed for the same pair of peas, provided the experimental conditions are the same for both peas, while for quantum mechanics we just saw that the results are always opposite for parallel settings; to obtain a perfect correspondence with the parable, we have to insert a sign change. For instance we may assume that, in the quantum experiment, Alice attributes a red color to result $+1$ and a blue color to result -1 , while Bob chooses the opposite convention¹³. In practice, Alice could complete her measurement apparatus with a device that makes a red lamp flash for result $+1$, or a blue lamp flash for result -1 ; Bob could use a similar device with different connections providing the opposite correspondence between results and colors. This question of sign does not introduce any difficulty; in the context of the EPR argument, there is no fundamental difference between a perfect correlation and a perfect anti-correlation.

Another essential ingredient of the EPR reasoning is the notion of “elements of reality”; EPR first remark that “these elements cannot be determined by a priori philosophical considerations, but must be found by an appeal to results of experiments and measurements”. They then propose the following criterion: “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”. In other words, certainty cannot emerge from nothing: an experimental result that is known in advance is necessarily the consequence of

¹³ Alternatively, one may make no change on the quantum mechanics side and attribute values $+1$ to red flowers and -1 to blue flowers in one place where peas are grown, and choose the opposite convention in the other place.

some pre-existing physical property. In our botanical analogy, we implicitly made use of this idea in the reasoning of §3.2.2.

A last, but essential, ingredient of the EPR reasoning is the notion of space-time and locality: the elements of reality in question are attached to the region of space where the experiment takes place¹⁴, and they cannot vary suddenly under the influence of events taking place in a very distant region of space. They can even less appear under these conditions¹⁵. The peas of the parable were in fact not so much the symbol of microscopic objects, electrons or spin 1/2 atoms for instance. Rather, they symbolize regions of space where we just know that “something is propagating inside”; it can be a particle, a field, a biological molecule, or anything else, with absolutely no assumption on its structure or physical description. Actually, in the EPR quotation of the preceding paragraph, one may replace the word “system” by “region of space”, without altering the rest of the reasoning. One may summarize the situation by saying that the basic belief of EPR is that regions of space can contain elements of reality attached to them (attaching distinct elements of reality to separate regions of space is sometimes called “separability” – see §3.3.3.b) and that they evolve locally – for brevity, this is often called “local realism” in the literature.

3.3.1.b Conclusions

From these assumptions, by the same reasoning as above, EPR obtain the equivalent of the conclusions of §3.2.2: whatever values are chosen for a , b , the results of the measurements are functions:

- (i) of intrinsic individual properties of the spins that they carry with them (the EPR elements of reality);
- (ii) of course, also of the orientations a , b of the Stern–Gerlach analyzers.
In addition, they show that:
- (iii) the results are given by well-defined functions of these variables, which implies that no indeterministic process is taking place; in other words, a particle with spin carries along with it all the information necessary to provide the result for the measurement, whatever choice is made for the orientation a (for the first particle) or b (for the second). All components of each spin have determined values simultaneously.

¹⁴ Einstein writes in [91]: “Things in physics are imbedded in a space-time continuum. These things require an autonomous existence inasmuch as they are in separate ‘parts’ of space”.

¹⁵ In standard quantum mechanics, if one applies the state vector reduction, the unmeasured spin suddenly “jumps” to a state with a perfectly well-defined spin component along the direction of measurement for the other spin. In this sense, quantum mechanics does attribute an element of reality to the unmeasured spin before it undergoes any measurement, in partial agreement with the EPR conclusion. But the disagreement occurs before the first measurement, since then standard quantum mechanics contains no such element of reality whatsoever.

- (iv) it is possible to envisage future measurements of observables that correspond to two different values b and b' for instance, that is to two different components of the spin that are called “incompatible” in quantum mechanics; the EPR reasoning shows that, in reality, incompatible observables can simultaneously have perfectly well-defined values.

Item (i) may be called the EPR-1 result; it implies that something is missing in quantum mechanics (the description of these intrinsic individual properties before the measurements), which is thus incomplete (EPR require from a complete theory that “every element of physical reality must have a counterpart in the physical theory”). The state vector may be a sufficient description for a statistical ensemble of pairs, but not for one single pair of spins; in this case, it should be completed by some additional information. In other words, inside the ensemble of all pairs, one can distinguish between sub-ensembles with different physical properties.

Item (iii) goes further and establishes the validity of determinism from a locality assumption, combined with correct predictions of quantum mechanics.

Item (iv) can be called the EPR-2 result; it shows that the notion of incompatible observables is not fundamental, but just a consequence of the incomplete character of the theory. It actually provides a reason to reject complementarity. Curiously, EPR-2 is often presented as the major EPR result, sometimes even with no mention of the others; actually, the rejection of complementarity is almost marginal or, at least, less important for EPR than the proof of incompleteness. In fact, in our reasonings, we will only need EPR-1. Einstein himself did not give much importance to the relation between the reasoning and non-commuting observables¹⁶, and it seems likely that this component was introduced into the article by Podolsky.

Finally, we should also mention that the EPR article contains another logical element, which is often overlooked (maybe because it is not mentioned in the abstract of the EPR article). Nevertheless, Einstein considered it as important – for instance it is the main aspect that he emphasizes in his 1936 article [76], one year after the publication of the EPR article¹⁷. The result may be called EPR-3: the description of physical reality given by quantum mechanics is “redundant”, since it describes the same physical reality for particle 2 with several different

¹⁶ At the end of the already mentioned letter to Schrödinger [90], he writes “as for the fact that the different states of system 2 may be considered as eigenvectors of different operators, I really do not care” (“Das ist mir Wurst” in German).

¹⁷ Einstein’s words in this article are (he calls systems A and B the systems we called 1 and 2): “Since there can be only *one* physical condition of B after the interaction, and which can reasonably not be considered as dependent on the particular measurement we perform on the system A separated from B, it may be concluded that the Ψ function is *not* unambiguously coordinated with the physical condition. The coordination of several Ψ functions with the same physical condition of system B shows again that the Ψ function cannot be interpreted as a (complete) description of a physical condition of a unit system”.

state vectors. Indeed, if all pairs of particles are identical as assumed in quantum mechanics, the physical reality attached to particle 2 cannot depend on the type of measurement performed on particle 1; it will necessarily be the same just after this measurement is performed, whatever physical quantity is measured on particle 1. But quantum mechanics implies that state vector reduction should be applied in a basis of the space of states that does depend on the measurement; therefore, particle 2 will reach several different states, depending on the measurement performed on particle 1. We then end up with too many different state vectors to describe the same physical reality for particle 2; hence a contradiction. It is somewhat paradoxical that the same theory should be both incomplete and redundant! By contrast, if we accept the idea of the EPR elements of reality, the measurement on particle 1 reveals pre-existing properties of this particle, and therefore of the emitted pair, and finally of particle 2 if the two particles are initially correlated; the situation can then be understood within the usual frame of classical correlations.

3.3.2 Bohr's reply

Bohr, in his reply [41, 97], does not criticize the EPR reasoning, but their assumptions, which he considers as inappropriate in quantum physics. For Bohr, the criterion of physical reality proposed by EPR “contains an essential ambiguity when applied to quantum phenomena”; he adds: “their argumentation does not seem to me to adequately meet the actual situation with which we are faced in atomic physics” (in this context, “atomic” is equivalent to “microscopic” or “quantum” in modern language). His text has been studied in detail and discussed by many authors (for an historical review, see for instance Chapter 6 of [50]), but remains delicate to understand in detail. One reason may be that, instead of concentrating his arguments on the precise situation considered by EPR, Bohr emphasizes in general the consistency of the mathematical formalism of quantum mechanics and the “impossibility of controlling the reaction of the object on the measuring instruments”. But, precisely, the main point of the EPR argument is to select a situation where these unavoidable perturbations do not exist! EPR locality implies that a measurement performed in region A can create no perturbation on the elements of reality in region B.

Only the second part of Bohr's article really deals with the EPR argument. Bohr then writes the following: “The wording of the above mentioned criterion (the EPR criterion for elements of reality) . . . contains an ambiguity as regards the expression ‘without in any way disturbing a system’. Of course there is, in a case like that considered (by EPR), no question of a mechanical disturbance of the system under investigation during the last critical stage of the measuring procedure. But even at this stage there is essentially the question of an influence on the very conditions

which define the possible types of predictions regarding the future behavior of the system ... the quantum description may be characterized as a rational utilization of all possibilities of unambiguous interpretation of measurements, compatible with the finite and uncontrollable interactions between the objects and the measuring instruments in the field of quantum theory”.

Several authors have wondered how exactly these words should be interpreted, what is meant by “mechanical disturbance” and, especially, how to understand the central sentence “there is essentially the question of an influence on the very conditions which define the precise types of predictions regarding the future behavior of the system”. What is probably meant is “an influence *of the first measurement performed in A* on the conditions which define the predictions on the future behavior of the system *during the second measurement performed in B*” – or maybe “the future behavior of *the whole system extending in both regions A and B*”. In Bohr’s view, physical reality cannot be properly defined without reference to a complete and well-defined experiment. This includes, not only the systems to be measured (the microscopic particles), but also all the measurement apparatuses: “these (experimental) conditions must be considered as an inherent element of any phenomenon to which the term physical reality can be unambiguously applied”. As a consequence, a spin measurement performed in A on one spin, as well as the direction of the measured component, change the physical reality of the whole physical system, which of course includes the second spin. This leads to a rejection of the EPR assumption, according to which the physical reality contained within the space region B of one spin is independent of measurements performed on the other in an arbitrarily remote region A, and the EPR conclusion is no longer valid.

In the same line, more than ten years later (in 1948), Bohr summarized his point of view by writing [46]: “Recapitulating, the impossibility of subdividing the individual quantum effects and of separating a behavior of the objects from their interaction with the measuring instrument serving to define the conditions under which the phenomena appear implies an ambiguity in assigning conventional attributes to atomic objects which calls for a reconsideration of our attitude towards the problem of physical explanation”. What is questioned by Bohr is therefore the very necessity of an explanation in the situation considered. The quotations given above in §2.5 also show that, for Bohr, a consistent interpretation of the quantum formalism cannot be obtained without the inclusion of the classical concepts associated with each experiment¹⁸.

¹⁸ Heisenberg did not agree with Einstein either, and characterized his position on quantum mechanics in the following terms (Chapter V of [70]): “When Einstein has criticized quantum theory he has done so from the basis of dogmatic realism. This is a very natural attitude. Every scientist who does research work feels that he is looking for something that is objectively true. His statements are not meant to depend upon the conditions under which they can be verified”.

J.S. Bell summarizes Bohr's reply concisely by writing [71] that, in Bohr's view, "there is no reality below some classical macroscopic level". Indeed, if one assumes that physical reality is necessarily macroscopic, then EPR's attempt to assign elements of reality to one of the spins only, or to a region of space containing it, is incompatible with quantum mechanics and therefore unphysical¹⁹ – even if the region in question is very large and isolated from the rest of the world. Expressed differently, a physical system that is extended over a large region of space is to be considered as a single entity, within which no attempt should be made to distinguish physical sub-systems or any sub-structure; trying to attach physical reality to regions of space is then automatically bound to failure. In terms of our Leitmotiv of §1.1.3, the difference between ordinary space and configuration space, we could say the following: the system has a single wave function for both particles that propagates in a configuration space with more than three dimensions, and this should be taken very seriously; no attempt should be made to come back to three dimensions and implement locality arguments in a smaller space. Bohr's point of view is, of course, not contradictory with relativity, but it certainly minimizes the impact of basic notions such as space-time, or events (a measurement process in quantum mechanics is not local; therefore it is not an event *stricto sensu*). His point of view does not fit well with a very strict interpretation of relativity.

Many physicists admit that a precise characterization of Bohr's attitude, in terms for instance of exactly what traditional principles of physics should be given up, is delicate (see for example the discussion of [10]). What is clear is that Bohr considers that it is vain to attempt to give any physical explanation that goes beyond orthodox quantum mechanics. In his reply to EPR [41] in *Physical Review*, there seems to be an influence of previous discussions that he had with Einstein at the Solvay conferences, explaining why he just repeats the orthodox point of view for a single particle submitted to incompatible measurements, rather irrelevant in the EPR context. Locality is not explicitly discussed. Did he fully appreciate how interesting the discussion becomes for two remote correlated particles, and the novelty of the EPR argument, which is the starting point of the Bell theorem for instance²⁰? In Pearle's words: "Bohr's rebuttal

In his Nobel lecture in 1933, he had already written: "The very fact that the formalism of quantum mechanics cannot be interpreted as a visual description of a phenomenon occurring in space and time shows that quantum mechanics is in no way concerned with the objective determination of space-time phenomena".

¹⁹ One could add that the EPR disproof of the notion of incompatible observables (EPR-2) implies that, at least, two different settings are considered for one of the measurement apparatuses; but this should correspond, in Bohr's view, to two different physical realities (corresponding to every different couple a, b), instead of a single one as assumed in the EPR reasoning.

²⁰ If Bohr had known the Bell theorem, he could merely have replied to EPR that their logical system was inconsistent (see §4.1.3)!

was essentially that Einstein's opinion disagreed with his own" [98]. Even Bell confessed that he had strong difficulties understanding Bohr ("I have very little idea what this means ..." – see the appendix of [71]). But, in any case, the Bohrian point of view remains presently as strong as ever, while we now know that the EPR assumptions contain self contradictions, as we will see in the next chapter.

3.3.3 *Locality and separability*

Locality and separability are two different notions, even if they are often (but not always) related.

3.3.3.a *Various aspects of locality*

The notion of locality itself contains various concepts. Most physicists, because of the strong impact of relativity, tend to immediately associate it with the notion of lightcone and the existence of a maximum possible velocity for the propagation of influences, interactions, or messages. It is indeed perfectly possible to interpret the EPR reasoning in these terms, assuming for instance that the two measurements are made simultaneously (in the reference frame of the source) at very large distances. But locality is also a much more general basic concept of physics, which predates any kind of relativity (Galileo or Einstein). It could be expressed simply as "the mutual influence of events decreases when their distance is increased"; or even "the influence of distant objects can be ignored if they are sufficiently remote". This purely spatial notion (no time is involved) is indeed one of the foundations of all experimental sciences: one assumes that the observations made in a laboratory depend on what is contained in the laboratory, but not on arbitrarily remote events (including human choices for experimental settings). In practice, if the observations made in each physics laboratory depended on the parameters chosen in all other laboratories in the world, it would probably become impossible to do any meaningful experiment! This simple notion of locality, which appears as a basic component of the scientific method, is sufficient for the EPR reasoning. See §3.3 of [99] for a generalization of the locality concept to systems with stochastic evolution; we come back to the importance of locality in the context of the Bell theorem in §4.2.3.

3.3.3.b *Quantum non-separability*

Instead of invoking the role of the measurement apparatuses to define physical reality, as Bohr does to reject the EPR reasoning, one can introduce the notion of "non-separability". The idea is that different quantum systems, when they have

interacted in the past, no longer have in general their own physical properties; they are both part of a larger system, which is the only one possessing physical properties. One should then not try to separate (conceptually) the whole system into two smaller physical systems and attribute them properties; the whole system is non-separable²¹.

In general, separability is a notion that is conceptually distinct from locality. It is not necessarily related to space: two physical systems could occupy the same region of space and remain distinct with their own physical properties (separable is not the same thing as separate). In the EPR reasoning, nevertheless, since the two particles are in remote regions of space, separation is assumed to entail separability. By contrast, quantum non-separability can be stated as “even when two regions of space are disconnected and very far apart, one cannot always attribute separate physical properties to what they contain”.

Quantum non-separability is rooted in the way the quantum formalism describes systems and sub-systems, and clearly related to the notion of entanglement (§6.1): a perfect description of the whole does not contain a perfect description of the parts. We mentioned earlier that Schrödinger considered entanglement as one of the most fundamental properties of quantum mechanics. Entanglement drastically restricts the number of physical properties that can be attributed to the sub-systems; this number may even vanish in some cases. In other words, the “best possible description” (with a state vector) is not available to the sub-systems; they have an additional level of indeterminacy, which never occurs in classical mechanics.

Invoking intrinsic quantum non-separability is appealing, since the difficulties associated with the precise definition of a measurement apparatus do not immediately appear. It nevertheless remains delicate, in particular because correlations and entanglement can, according to the Schrödinger equation, propagate towards the macroscopic world (Schrödinger cat paradox) so that this absence of physical properties can reach any scale; it can end up with a situation where only the whole universe has physical properties!

Generally speaking, if the world was completely non-local (or completely non-separable), all physical phenomena being totally interlaced, unravelling them would be beyond reach, and their scientific analysis by the experimental method would become impossible. The scientific progress of mankind have been made possible by the fact that, fortunately, non-locality and non-separability remain relatively rare phenomena, appearing mostly in physical situations designed by physicists to observe them.

²¹ Peres has proposed a mathematical criterion defining the separability of an ensemble of two physical systems described by a density operator (see §6.2.5).

3.3.4 *The EPR argument for macroscopic systems*

Interestingly, the EPR argument also applies in the context of another discovery made by Einstein, namely Bose–Einstein condensation. The original EPR argument involved two microscopical particles, atoms for instance. The essence of Bohr’s point of view then hinges on the idea that microscopic systems do not possess physical reality independently of the measurement apparatuses, since physical reality cannot be defined at this level. But it turns out that quantum mechanics also predicts that similar correlations should be observed with systems that are arbitrarily large, and can therefore be macroscopic [100]; this requires that they should be initially in a special quantum state, a “double Fock state”. Such a state is not very common, but could be reached for instance through the phenomenon of Bose–Einstein condensation.

Here we just summarize the general idea, since we come back to this subject in §9.4.3. We assume that two large condensates, associated with two Fock states with different spin states, overlap in two different regions of space. Initially, according to quantum mechanics, the relative phase of the two condensates is totally undetermined, so that none of the overlap region contains any transverse spin orientation. But, if spin measurements are performed along transverse directions in one of the regions, a transverse spin polarization spontaneously appears under the effect of quantum measurement in this region. The direction of this spontaneous polarization is completely random. Nevertheless, since this process fixes the relative phase of the two condensates, theory predicts that a spin polarization also appears in the other region of overlap, with a transverse direction that is parallel to that in the measurement region. One then has a remote consequence of the measurement, without any interaction between the two regions, but just under the effect of state vector reduction.

Clearly, the EPR argument then directly applies: How can a spin polarization spontaneously appear in an arbitrarily remote region of space, without any interaction to create it? How can the physical reality associated to a region of space be affected by an arbitrarily distant measurement? But in this case the big difference with the original EPR situation is that the number of particles involved is arbitrarily large, so that the transverse spin polarizations can be macroscopic. If the spins carry magnetic moments, they will create a macroscopic transverse magnetization, which can be detected directly with macroscopic devices such as ordinary compasses. For arbitrarily large macroscopic objects, it seems difficult to invoke the Bohrian argument, and to decide that they have access to physical reality only when they are associated with a well-defined measurement apparatus!

In this case, the EPR element of reality pre-existing the first measurement is the relative phase of the two condensates with spin. The notion is related to spontaneous

symmetry breaking and the spontaneous appearance of a phase in a system that undergoes a superfluid transition (Anderson's phase [101]). Of course, we do not know what Bohr would have said concerning this macroscopic version of the EPR thought experiment. The relative phase of Bose–Einstein condensates also contains some interesting effects leading to violations of quantum non-locality [102], but non-local effects will be discussed in the next chapter.

4

Bell theorem

The Bell theorem [103] can be seen in several different ways, as the EPR argument. Initially, Bell invented it as a direct logical continuation of the EPR theorem: the idea is to take the existence of the EPR elements of reality seriously, and to push it further by introducing them explicitly into the mathematics with the notation λ ; one then proceeds to study all possible kinds of correlations that can be obtained from the fluctuations of one or several variables λ , making the condition of locality explicit in the mathematics (locality was already useful in the EPR theorem, but not used in equations). The reasoning develops within determinism (considered as proved by the EPR reasoning) and classical probabilities; it studies in a completely general way all kinds of correlations that can be predicted from the fluctuations of some classical common cause in the past – if one prefers, from some random choice concerning the initial state of the system. This leads to the famous inequalities. But subsequent studies have shown that the scope of the Bell theorem is not limited to determinism; for instance, the λ variables may determine the probabilities of the results of future experiments, instead of the results themselves (see Appendix B), without canceling the theorem. We postpone the discussion of the various possible generalizations to §4.2.3. For the moment, we just emphasize that the essential condition for the validity of the Bell theorem is locality: all kinds of fluctuations can be assumed, but their effect must affect physics only locally. If we assume that throwing dice in Paris may influence physical events taking place in Tokyo, or even in other galaxies, the proof of the theorem is no longer possible. For non-specialized general discussions of the Bell theorem, see for instance [48, 71, 104, 105].

Another, less general, view of the Bell theorem is obtained by disconnecting it from the EPR reasoning. It can then be seen as an impossibility theorem for additional (or “hidden”) variables (see §10.6): if these variables are arbitrarily introduced, and if their evolution is local, the resulting theory cannot reproduce all predictions of standard quantum mechanics, even if a very complicated evolution is postulated.

4.1 Bell inequalities

The Bell inequalities are relations satisfied by the average values of products of random variables that are correlated classically – by this we mean that their correlations originate from some fluctuating past event that has influenced their values, as in the preceding chapter for the peas. The inequalities are especially interesting in cases where they are contradictory with quantum mechanics; one of these situations occurs in the spin version of the EPR argument [96], already introduced in §3.3.1, where two spin 1/2 particles undergo measurements in remote regions of space. This is why we begin this section by briefly recalling the predictions of quantum mechanics for such a physical system – the only ingredient we need from quantum mechanics at this stage is the predictions concerning the probabilities of results. Then we will leave again standard quantum mechanics and come back to the EPR–Bell argument, discuss its contradictions with quantum mechanics, and finally emphasize the generality of the theorem.

4.1.1 Quantum mechanics: two spins in a singlet state

We come back to the experiment shown in Figure 3.2, where two spins 1/2 are emitted by a source S in a singlet spin state and propagate in opposite directions. Their spin state is described by:

$$|\Psi\rangle = \frac{1}{\sqrt{2}} \left[|+, -\rangle - |-, +\rangle \right] \quad (4.1)$$

with the usual notation: the two-spin state $|\pm, \mp\rangle$ contains the first spin in an eigenstate with eigenvalue $\pm\hbar/2$ of its component along Oz ¹, and the second in an eigenstate with eigenvalue $\mp\hbar/2$. When they reach distant locations, the spin on the left is submitted to a spin measurements with a Stern–Gerlach apparatus oriented along the direction defined by angle a , while the other spin undergoes a similar measurement along the direction defined by b .

If:

$$\theta_{ab} = a - b \quad (4.2)$$

is the angle between the directions defined by a and b , quantum mechanics predicts that the probability for a double detection of results $+1, +1$ (or of $-1, -1$) is:

$$\mathcal{P}_{(+,+)} = \mathcal{P}_{(-,-)} = \frac{1}{2} \sin^2 \frac{\theta_{ab}}{2} \quad (4.3)$$

¹ One can show that the singlet state is invariant under any rotation, which implies that it has the same form (4.1) if any quantization axis Oz is used.

while the probability of two opposite results is:

$$\mathcal{P}_{(+,-)} = \mathcal{P}_{(-,+)} = \frac{1}{2} \cos^2 \frac{\theta_{ab}}{2} \quad (4.4)$$

This is all that we need to know, for the moment, of quantum mechanics: probability of the results of measurements. An important remark is that, if $\theta_{ab} = 0$ (when the orientations of the measurements apparatuses are parallel), the formulas predict that the two probabilities (4.3) vanish, while the others are equal to 1/2. This means that the condition of perfect correlations required by the EPR reasoning is fulfilled. More precisely, the results of the experiments are always opposed instead of equal, but it is sufficient to change the arbitrary direction of one of the axes to make them equal – see the discussion of §3.3.1.a.

A state such as (4.1), where the properties of the two sub-systems (two spins in this case) are correlated inside the state vector itself, is called “entangled state” in quantum mechanics. We will come back to the notion of quantum entanglement in Chapter 6 in more detail.

4.1.2 Local realism: proof of the BCHSH inequality

We start from the EPR theorem and, following Bell, we assume that λ represents the “elements of reality” associated with the spins; it should be understood that λ is only a concise notation which may summarize a vector with many components. The number of elements of reality is arbitrary – no limitation is introduced by this notation – and, in fact, one can even include in λ components that play no special role in the problem. The only thing that really matters is that λ should contain all the information concerning the results of possible measurements performed on the spins. Another usual notation is to write these results A and B , and to use small letters a and b for the settings (parameters) of the corresponding apparatuses. Clearly, A and B may depend, not only on λ , but also on the settings a and b ; nevertheless, locality requests that b has no influence on the result A (since the distance between the locations of the measurements can be arbitrarily large); conversely, a has no influence on result B . We therefore call $A(a, \lambda)$ and $B(b, \lambda)$ the corresponding functions; their values are either +1 or –1.

In what follows, it is sufficient to consider two directions only for each separate measurement; we then use the simpler notation:

$$A(a, \lambda) \equiv A \quad A(a', \lambda) \equiv A' \quad (4.5)$$

and:

$$B(b, \lambda) \equiv B \quad B(b', \lambda) \equiv B' \quad (4.6)$$

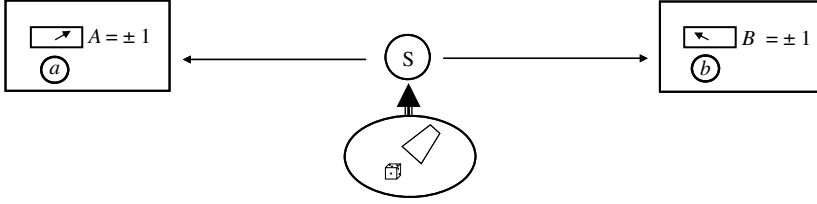


Figure 4.1. A source S emits two particles, which propagate to two remote measurement apparatuses having respective settings a and b ; each apparatus provides a result ± 1 . The oval containing a dice below the source symbolizes a random fluctuating process controlling the conditions under which the two particles are emitted, which in turn determine their properties. Correlations are observed between the results obtained with the two remote measurement apparatuses; they are consequences of the common random properties shared by the particles of each pair, initially created by the random process at the source.

For each pair of particles, λ is fixed, and the four numbers have well-defined values (which can only be ± 1). With Eberhard [106], we notice that the sum of products:

$$M = AB - AB' + A'B + A'B' = A(B - B') + A'(B + B') \quad (4.7)$$

is always equal to either $+2$ or -2 ; this is because one of the brackets in the right-hand side of this equation always vanishes, while the other is ± 2 . Now, if we take the average value $\langle M \rangle$ of M over a large number of emitted pairs (average over λ), since each instance of M is limited to these two values, we necessarily have:

$$-2 \leq \langle M \rangle \leq +2 \quad (4.8)$$

This is the so-called BCHSH (Bell, Clauser, Horne, Shimony, and Holt) form [107] of the Bell theorem. The inequality is necessarily satisfied² by the average values of all possible kinds of measurements providing random results ± 1 , whatever the mechanism creating the correlations may be, as long as locality is obeyed: A does not depend on setting b , and B does not depend on setting a .

A natural assumption is to consider that randomness arises from the fluctuations of some common cause in the past. Figure 4.1 illustrates this situation, and Figure 4.2 shows the corresponding space-time representation of the events; the lines connecting the cause and the effects must remain within the light cone $x = \pm ct$ if relativity is obeyed. But the inequality remains valid if, for instance, other fluctuating processes perturb the particles while they are on their way to the measurement apparatuses, or while they interact with the measurement apparatuses themselves (Figure 4.3).

² In our definition (4.7) of M , the term AB' has a minus sign and the three others a plus sign, but this is arbitrary. By writing $AB + AB' \pm A'B \mp A'B' = A(B + B') \pm A'(B - B')$ and $\pm AB \mp AB' + A'B + A'B' = \pm A(B - B') + A'(B + B')$, we can obtain four inequalities where the minus sign goes to any of the four terms.

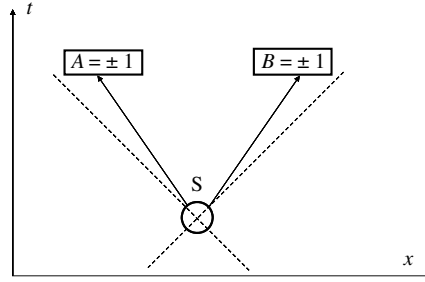


Figure 4.2. Space-time diagram associated with the events shown in Figure 4.1. According to relativity, the two arrows joining the emission of the particles to the measurement events should remain inside the light cone $x = \pm ct$ (dotted lines).

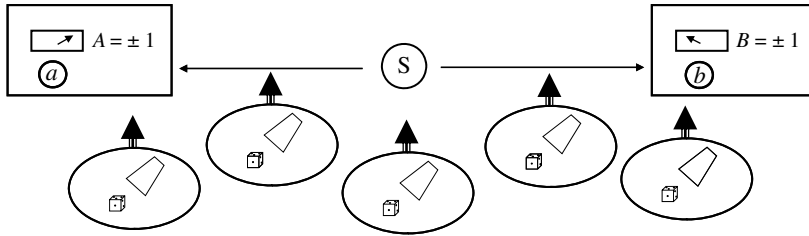


Figure 4.3. Fluctuating uncontrolled causes may influence, not only the emission of particles, but also their propagation and the measurement apparatuses, without changing the validity of the BCHSH inequality.

It then becomes necessary to include into λ new components associated with the corresponding random processes, which may change the distribution of this multi-dimensional variable, but not the fact that the average of a number M that is always equal to ± 2 necessarily obeys (4.8).

The simplicity of the proof is such that one can expect the validity of the inequality to apply in many situations; this is indeed true, as we will discuss in more detail in §§4.2.2 and 4.2.3. Here we just remark that the result is independent of the interpretation of the variable λ , which does not have to be defined as an additional variable or an element of reality. For instance, we can assume that λ is simply used to label the realization of the experiment: $\lambda = 1$ corresponds to the first experiment, $\lambda = 2$ to the second, $\dots$, $\lambda = N$ to the last of a series of experiments. If, for each realization, the four numbers A , B , A' , and B' have well-defined values, all equal to ± 1 , the number M is also defined for each realization and is necessarily equal to -2 or $+2$. Whatever arbitrary values of the four numbers may occur during any series of experiments, mathematically, there is no way for the sum of the M to exceed $2N$ or to be less than $-2N$. The average obtained by dividing by N then necessarily obeys (4.8); the mere existence of the four numbers is sufficient to obtain this result.

4.1.3 Contradiction with quantum mechanics

The simplicity and the generality of the proof are such that one could reasonably expect that any sensible physical theory should automatically give predictions that obey this inequality. The surprise is to realize that quantum mechanics does not. From (4.3) and (4.4), we can calculate the average value of the product of the two results obtained by measuring the components of the two spins along directions a and b with relative angle θ_{ab} ; the result is:

$$\mathcal{P}_{(+,+)} + \mathcal{P}_{(-,-)} - \mathcal{P}_{(+,-)} - \mathcal{P}_{(-,+)} = -\cos\theta_{ab} \quad (4.9)$$

This gives the quantum equivalent of the average over λ of the product $A(a, \lambda)B(b, \lambda)$ in a local realist theory. To obtain the quantum equivalent $\langle Q \rangle$ of the combination of four products in expression (4.7), we calculate the same combination of averages of products of results:

$$\langle Q \rangle = -\cos\theta_{ab} + \cos\theta_{ab'} - \cos\theta_{a'b} - \cos\theta_{a'b'} \quad (4.10)$$

For some appropriate choices of the four angles a, a', b, b' , $\langle Q \rangle$ reaches values $\pm 2\sqrt{2}$. Figure 4.4 shows the relative positions of the four vectors $\mathbf{a}, \mathbf{a'}, \mathbf{b}, \mathbf{b'}$ defining the directions of measurements for which these extrema are obtained³. But we have seen that all local realist theories obey the BCHSH inequality (4.8) with bounds ± 2 . Quantum mechanics therefore predicts a violation by a factor $\sqrt{2}$, more than 40% (it turns out that $\sqrt{2}$ is the maximal possible violation predicted by quantum mechanics – cf. §4.3). As simple as the cosine variation of (4.9) may look, no local realist theory can reproduce it. The EPR–Bell reasoning therefore leads to a quantitative and significative contradiction with quantum mechanics, which is not a local realist theory in the EPR sense.

How is this contradiction possible, and how can a reasoning that is so simple be incorrect within quantum mechanics? Different answers are possible:

- (i) In Bohr's view, as we already discussed in §3.3.2, only the whole experiment should be considered, without distinguishing in it two separate measurements that would be performed on each particle: a single indivisible two particle measurement takes place. A fundamentally random process then takes place simultaneously over the whole region of space occupied by the experimental setup; it is delocalized, as schematized in Figure 4.5, to be contrasted with Figures 4.1 and 4.3.

³ The four vectors $\mathbf{a}, \mathbf{b}, \mathbf{a'}$ and $\mathbf{b'}$ are in the same plane and each of them makes an angle of 45° with the preceding vector; one then immediately sees that $\langle Q \rangle = -2\sqrt{2}$. If both directions of $\mathbf{b}$ and $\mathbf{b'}$ are reversed, one obtains $\langle Q \rangle = 2\sqrt{2}$.

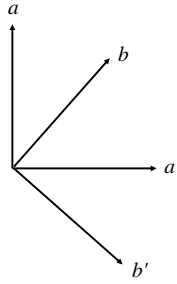


Figure 4.4. The position of the four vectors a , b , a' , and b' corresponding to a maximal violation of the BCHSH inequality for two spins $1/2$ in a singlet state. The only pair of vectors giving rise to a negative correlation coefficient is (a, b') , since the angle between the vectors exceeds 90° .

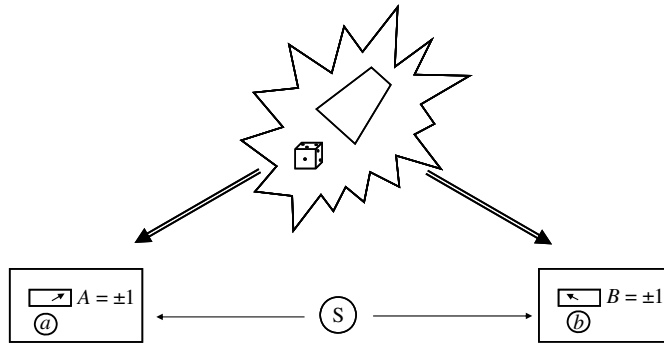


Figure 4.5. Scheme associated with the quantum description of a measurement involving two remote apparatuses. This scheme may be compared to those of Figures 4.1 and 4.3. The big flash at the top of the figure symbolizes a fundamentally random and non-local process, by contrast with ovals below the events used in previous figures to represent fluctuating processes resulting from uncontrolled local perturbations. The quantum process is inherently extended in space, so that no space-time description similar to that of Figure 4.2 is possible.

Functions A and B then depend on the two settings, so that they should be written $A(a, b)$ and $B(a, b)$. Instead of two numbers A and A' , we now have four, which are $A = A(a, b)$, $A' = A(a', b)$, as well as $A'' = A(a, b')$ and $A''' = A(a', b')$; similarly, B and B' are replaced by four numbers. Their products then have 16 possible values instead of 4; clearly, the proof of the BCHSH inequality is then no longer possible so that the contradiction disappears.

- (ii) One may prefer a more local view of the process of measurement and retain the concept of single-particle measurement as meaningful in this context. To avoid the contradiction with the predictions of quantum mechanics, one then considers that it is wrong to attribute well-defined values A , A' , B , B' to each emitted pair, since only two of them can be measured in any realization of the experiment. Therefore, it is not legitimate to speak of these four quantities,

or reason on them, even as unknown quantities (we come back to this point with the discussion of counterfactuality in §4.5.2). As nicely summarized by Peres [108], “unperformed experiments have no result⁴”! Wheeler expresses a similar idea when he writes: “No elementary phenomenon is a phenomenon until it is a registered (observed) phenomenon” (p. 184 of [109]).

- (iii) As for Wigner, he emphasizes in [110] that the proof of the Bell inequalities relies on a very simple notion within realism (local or non-local): the number of categories into which one can classify all pairs of particles⁵ or, equivalently, all realizations of the experiment. Each category is associated with well-defined results of measurements for the various choices of the settings a and b (configurations of the experiment) that are considered; in any long sequence of repeated experiments, each category contributes with some given weight equal to its probability of occurrence (a positive or zero number). For one single type of experiment that can give R different results, the number of categories is R . When one considers P different possible configurations, each with R different possible results, the number of categories becomes R^P .

For the sake of simplicity, here we assume that each setting can take two different values (but, in §4.2.1, we study Wigner’s original case with three different values). If, following Bohr, we consider that the experiment as a whole must be considered, then each configuration is given by the choice of a pair a, b among $P = 4$ possibilities; since each of them can give 4 different pairs of results, the number of categories is 4^4 . One can then attribute to each of them appropriate weights and reproduce the predictions of quantum mechanics, and no special contradiction is obtained.

But then Wigner notes that, if one adds the notion of locality, for each pair the result on one side becomes independent of the setting on the other side, keeping only a dependence on the local setting. Each category then becomes the intersection of two sub-ensembles: one associated with the first side of the experiment depending on a only, which is chosen among $2 \times 2 = 4$ possibilities (2 for the possible choices of a , and 2 for the possible results); another sub-ensemble associated with the second side of the experiment depending on b only, which is also chosen among 4 possibilities. Altogether the total number of categories is now only 2^4 . To each of these categories one can ascribe four numbers A, A', B and B' , which are well defined and all equal to ± 1 , so

⁴ In Bohr’s view, it is not forbidden to consider that “unperformed experiments have results”, but experiments and results can only be expressed in terms of the whole experimental setup, and therefore of both parameters a and b .

⁵ In this reference, Wigner actually reasons explicitly in terms of hidden variables; he considers domains for these variables, which correspond to given results for several possible choices of the settings. But, from the EPR point of view, these domains also correspond to categories of pairs of particles, which is why we use this notion of categories.

that their combination in (4.7) has values $M = \pm 2$; the BCHSH inequalities are then immediately obtained. In other words, Wigner points out that the mathematical origin of the Bell inequalities lies precisely in the possibility of distributing all pairs with positive probabilities into a smaller number of categories than from a non-local point of view; the difference between these numbers is the origin of the contradiction⁶.

4.1.4 Logical content

A general way to express the Bell theorem in logical terms is to state that the following system of three assumptions⁷ (which could be called the EPR assumptions) is self-contradictory:

- (1) validity of the EPR notion of “elements of reality”,
- (2) locality,
- (3) the predictions of quantum mechanics are always correct.

Indeed, if one starts from the three assumptions, one can then discover situations where the third assumption is not true, hence a self-contradiction. The Bell theorem then becomes a useful tool to build a “reductio ad absurdum” reasoning: it shows that, among all three assumptions, one (at least) has to be given up. If Einstein’s program was to assume local realism and, at the same time, that all predictions of quantum mechanics are correct, this program is impossible. This conclusion is purely logical, independent of any experimental result.

One may notice that the reasoning actually contains a fourth assumption:

- (4) the measurement settings a and b are freely chosen by the experimenters⁸, and not the physical consequence of some past event.

This is the “free will” hypothesis. It is so general in all experimental scientific disciplines that, often, it is not even mentioned: one naturally assumes that experimenters can freely decide what kind of experiments they will perform, and what parameters they will choose for them. The opposite attitude would be to consider that these decisions are in fact predetermined by some past event and by the propagation of its influences to the distant experimenters (see in §4.5.1.c

⁶ It has also been noticed [111] that the Bell inequalities can be seen as applications of theorems on marginal distributions in the theory of probabilities [112]; the mathematical proof of the inequalities relies on the existence of one common probability space for the relevant random variables (this is also related to the notion of counterfactuality).

⁷ Here we focus ourselves only on one set of assumptions leading to the theorem, the set we have used to derive it. Other possible sets of assumptions to derive the theorem will be discussed in §4.2.2.

⁸ One can even assume that they choose their value after the emission of the two particles by the source while they are still flying towards them.

the discussion of fatalism, or superdeterminism). What is then at stake is the notion of a free parameter in a physical theory: a and b are considered as free parameters that are external to the theory, and not as solutions of some dynamical equations starting from a common initial condition in the past.

When stated in this way, the Bell theorem appears as general conceptually, but experimentally inaccessible since assumption 3 is obviously too broad to be tested. One may then prefer another form of the logical self-contradiction, where this assumption is replaced by two more specific statements concerning one particular experiment only (for instance: two $1/2$ spins in a singlet state, or two photons emitted in an atomic 0-1-0 cascade):

- (3') in this experiment, the prediction of quantum mechanics concerning the total correlations observed when the settings are the same ($a = b$) are correct (this assumption leads to the existence of the EPR elements of reality).
- (3'') the predictions concerning the correlations for different settings are also correct.

Removing either (3') or (3'') from the set of assumptions is also sufficient to remove the self-contradiction. The motivation of the experimental tests of the Bell inequalities was precisely to check if it was not assumption (3') or (3'') that should be abandoned. Maybe, after all, the Bell theorem is nothing but an efficient pointer towards unexpected situations where the predictions of quantum mechanics are so paradoxical that they are actually wrong? Such was the hope of some theorists, as well as an exciting challenge to experimentalists.

4.1.5 Contradiction with experiments

In 1967, an experiment was undertaken by Kocher and Commins [113] with photons emitted during an atomic cascade between three levels of atomic Mercury having angular momenta $J = 0 \rightarrow 1 \rightarrow 0$. In this case, quantum mechanics predicts correlations that are similar to those with two spins $1/2$ in a singlet state; the directions of photon polarization analyzers play the role of the orientations of the Stern–Gerlach analyzers for spins $1/2$ (it is sufficient to divide all angles by 2 to transpose the results from spins to photons). This experiment tested that the predictions of quantum mechanics are correct when the two analyzers are parallel (or perpendicular), corresponding to perfect correlations – in other words, assumption (3') was tested successfully in this case.

Two years later Clauser, Horne, Shimony, and Holt [107] realized that this kind of experiment could be extended to provide an experimental test of assumption (3''), with measurements of correlations rates at various angles of the polarizers. They proposed a generalization of the Bell theorem to a new form (the BCHSH

inequality) that was more suitable to experimental tests than the original Bell inequality. Following this line, in 1972 correlations for different settings (and therefore assumption 3'') were tested by Freedman and Clauser [114] in an atomic cascade $J = 0 \rightarrow 1 \rightarrow 0$ of the Calcium atom; this provided a confirmation of the quantum predictions and a lead to violation of local-realist inequalities by 6 standard deviations. In 1976, three other experiments were performed, two with photons emitted during $J = 1 \rightarrow 1 \rightarrow 0$ cascades in Mercury [115, 116], and one with protons [117]; again the predictions of quantum mechanics were fully confirmed.

In the eighties, the results were made more and more precise and convincing in a series of experiments performed in Aspect's group, using the same Calcium transition as [114]. One of these new experiments [118] included a study of the effect of the distance between the detectors and the source on the quantum correlations for various polarizations, in order to check that this distance plays no role as predicted by quantum mechanics. Another used two-channel detections [119] providing actual ± 1 signals, while in previous experiments one polarization was just absorbed and gave no signal at all (this experiment led to violations by 15 standard deviations!). The third included a time-component in a random choice of the polarizations of the detections [120] – see also [121] for experiments using a two-photon transition between two levels of zero angular momentum in Deuterium.

Ever since, the experiments have been constantly improved, including the use of parametric down conversion of light to obtain efficient two-photon sources, leading to violations by 22 standard deviations [122], and even measurements of correlations over more than 10 km. [123], and now even more than 100 km. [124]. Violations with systems of four photons, equivalent to two particles of spin 1 correlated in a singlet state, have been reported [125]. The list of all references is too long to be given here, but we give a few more in §4.5.1 during the discussion of “loopholes”. A summary of the present situation is that, even in these most intricate situations invented and tested by the experimentalists, no one has been able to disprove quantum mechanics. In this sense, we can say that Nature obeys laws which are non-local, or non-realist, or both. It goes without saying that no experiment in physics is perfect, and it is always possible to invent *ad hoc* scenarios where some physical processes, for the moment totally unknown, “conspire” in order to give us the illusion of correct predictions of quantum mechanics – we come back to this point in §4.5.1 – but the quality and the number of the experimental results does not make this attitude very attractive intellectually.

4.2 Various forms of the theorem

What is generally meant by the “Bell theorem” is actually not limited to a single inequality applying in one set of specific circumstances, but encompasses a whole

set of inequalities that are valid in various cases. We will give a few examples of such derivations, inequalities, and possible generalizations.

4.2.1 Other inequalities

Many forms of inequalities⁹ have been proposed as consequences of local realism (or of the other logical sets discussed above), while in contradiction with the predictions of quantum mechanics; here we will give three examples as illustrations, starting with the initial 1964 Bell inequality. Following a well-established tradition, we will call Alice and Bob the partners who, each in one of two remote laboratories, perform experiments on the particle they receive for each emitted pair by the source.

4.2.1.a Bell 1964

In his initial 1964 article [103], Bell did not introduce the BCHSH inequality (4.7)–(4.8), but another similar but mathematically distinct inequality. His derivation was also different from that given above, but we now show that this inequality can be obtained as a special case of the BCHSH inequality. Assume we replace (4.5) and (4.6) by:

$$A = A(a, \lambda) \quad ; \quad B = B(b, \lambda) \quad ; \quad A' = A(-b, \lambda) \quad ; \quad B' = B(c, \lambda) \quad (4.11)$$

where a , b , and c are three angles defining three different settings (orientations of the Stern–Gerlach analyzers), instead of four as above; notation $-u$ is used for the direction opposite to u , obtained by turning the analyzer by half a turn. Since this operation interchanges results $+1$ and -1 , we have, for any u :

$$A(-u, \lambda) = -A(u, \lambda) \quad (4.12)$$

Now, we assume that, as quantum mechanics predicts, the results obtained by Alice and Bob are always perfectly correlated when they use opposite directions for their measurements. We therefore have:

$$A'B = A(-b, \lambda)B(b, \lambda) = 1 \quad (4.13)$$

so that (4.8) becomes:

$$-2 \leq \langle A(a, \lambda)B(b, \lambda) - A(a, \lambda)B(c, \lambda) - A(b, \lambda)B(c, \lambda) \rangle + 1 \leq +2 \quad (4.14)$$

⁹ It has been noticed [111] that the Bell inequalities can be seen as applications of theorems on marginal distributions in the theory of probabilities [112]; the mathematical proof of the inequalities relies on the existence of one common probability space for the relevant random variables (this is related to the notion of counterfactuality).

If we note $E(a, b)$ the average of the product of results obtained by Alice and Bob:

$$E(a, b) = \langle A(a, \lambda) B(b, \lambda) \rangle \quad (4.15)$$

the upper bound condition in (4.14) then becomes:

$$E(a, b) - E(a, c) \leq 1 + E(b, c) \quad (4.16)$$

Now, if we replace a by $-a$ and perform the same calculation, we obtain:

$$-E(a, b) + E(a, c) \leq 1 + E(b, c) \quad (4.17)$$

or, grouping (4.16) and (4.17):

$$|E(a, b) - E(a, c)| \leq 1 + E(b, c) \quad (4.18)$$

which is the 1964 Bell inequality. We remark that its proof requires the perfect correlations contained in (4.12), which is not the case in the BCHSH inequalities.

Quantum mechanics predicts violations of this inequality: if we choose $\theta_{ab} = 45^\circ$ and $\theta_{ac} = 135^\circ$, using (4.9) we obtain $\sqrt{2}$ for the left-hand side of (4.18) while the right-hand side is equal to 1, corresponding to a violation by a factor $\sqrt{2}$.

4.2.1.b Wigner inequalities

At the end of §4.1.3, we discussed Wigner's point of view, who ascribes the successive realizations of an experiment to different categories; this classification stems from a realist point of view that can be either local or non-local, with a smaller number of categories in the former case. The choice of the experimental setting by each partner was limited to two different values; here, as in the original article [110], we assume that Alice may choose among three different values a , a' , and a'' , and, similarly, that Bob may choose among three values, b , b' , and b'' . The number of pairs of possible results $(\pm 1, \pm 1)$ for each realization remains the same as above, $R = 4$, but the number of experimental configurations now increases to $P = 3 \times 3 = 9$. From a non-local realist point of view, the number of categories is then 4^9 ; if locality is added, one has to consider the intersection of 2^3 classes by 2^3 others, which eventually leads to $2^6 = 4^3$ categories, a much smaller number indeed.

Within local realism, how to simply define these categories? It is actually sufficient to specify the results $A = \pm 1$, $A' = \pm 1$, and $A'' = \pm 1$ observed by Alice if she chooses respectively the settings a , a' , or a'' (locality implies that the value of the setting chosen by Bob does not affect her own result) as well as the three values B , B' , and B'' observed by Bob if he chooses respectively settings b , b' , or b'' . We now attempt to reproduce, within this scheme, the predictions of quantum mechanics

for two spins $1/2$ in a singlet state, when $b = a$, $b' = a'$, and $b'' = a''$. For these configurations, we have seen in §4.1.1 that the results obtained by Alice and Bob are always opposed. This leads us to write $B = -A$, $B' = -A'$, and $B'' = -A''$; it then becomes useless to specify the values of Bob's results, and the number of categories of realizations is now reduced to $2^3 = 8$. The predictions of quantum mechanics are reproduced each time Alice and Bob choose parallel directions; but we will see that, when they choose different directions, it cannot always be the case.

We consider a large number N of realizations of the experiment, each associated with results (A, A', A'') – it is not necessary to specify the values of B, B' , and B'' since they are opposite. We denote $p_3(A, A', A'')$ the proportion of these realizations falling into the category defined by three given results, among the eight possible triplets. Similarly, we note $p_2(A, A')$ the proportion of realizations for which only the two variables A and A' are specified, but not A'' ; in the same way, we introduce the proportions $p_2(A, A'')$ and $p_2(A', A'')$. We then have:

$$p_2(A = +1, A'' = -1) = p_3(+1, +1, -1) + p_3(+1, -1, -1) \quad (4.19)$$

while the proportion $p_2(A = +1, A' = -1)$ is given by:

$$p_2(A = +1, A' = -1) = p_3(+1, -1, +1) + p_3(+1, -1, -1) \quad (4.20)$$

and finally the proportion $p_2(A' = +1, A'' = -1)$ is given by:

$$p_2(A' = +1, A'' = -1) = p_3(+1, +1, -1) + p_3(-1, +1, -1) \quad (4.21)$$

Since the proportions are ratios between numbers that are positive or zero, they are themselves positive or zero. The preceding equalities then provide:

$$p_2(A = +1, A'' = -1) \leq p_2(A = +1, A' = -1) + p_2(A' = +1, A'' = -1) \quad (4.22)$$

an inequality which puts limits on the possible proportions when the two settings are different.

In quantum mechanics, this inequality may be violated. This is because, since result B'' is the opposite of A'' , equation (4.3) provides:

$$p_2(A = +1, A'' = -1) = p_2(A = +1, B'' = +1) = \frac{1}{2} \sin^2 \frac{\theta_{ab''}}{2} \quad (4.23)$$

where $\theta_{ab''}$ is the angle between directions a and b'' (or a'' , since we have assumed that $b = a$, $b' = a'$, and $b'' = a''$). Inequality (4.22) then becomes:

$$\sin^2 \frac{\theta_{aa''}}{2} \leq \sin^2 \frac{\theta_{aa'}}{2} + \sin^2 \frac{\theta_{a'a''}}{2} \quad (4.24)$$

In the particular case where a , a' , and a'' are equally distributed in a plane and in this order, we have $\theta_{aa'} = \theta_{a'a''}$ and $\theta_{aa''} = 2\theta_{aa'}$, and the preceding inequality becomes:

$$2\cos^2 \frac{\theta_{aa'}}{2} \leq 1 \quad (4.25)$$

which is violated in interval $-\pi/2 < \theta_{aa'} < +\pi/2$. We have therefore another case where the predictions of quantum mechanics contradict those of local realism.

4.2.1.c Mermin inequality

Another inequality, proposed by Mermin [126], can be obtained with the same experimental configuration, but different angles than those considered by Wigner. In footnote 11 of Chapter 3, we mentioned Schrödinger's comparison to illustrate the local realist reasoning: he represents the particles by schoolboys passing exams, the measurement apparatuses by teachers asking them questions, and the results of measurements by the answers of the schoolboys, assumed to be binary answers (+1 for yes, -1 for no). Extending this comparison, we assume that two schoolboys, Albert and Bernard, are each asked a question by two different teachers in separate rooms; each teacher chooses his question at random among three possible questions, with a probability 1/3 each (these three questions correspond to the three possible values of the settings a or b). The plan of the schoolboys is to give answers to the question that reproduce the predictions of quantum mechanics as well as possible; they know in advance the nature of the three questions, but not which precise question they will have to answer; moreover, when they meet one teacher, they do not know what question has been selected by the other (since the rooms are separate). However, before the examination, they are free to elaborate together any common strategy and to decide what answer each of them will give to each question. The total number of possible strategies is $2^3 \times 2^3 = 64$.

As above, in order to reproduce the quantum predictions, every time the questions are the same the schoolboys have to make opposite answers. Obtaining this result is not very difficult: they decide in advance which answer Albert will make to every possible question, and agree that Bernard will choose the opposite answer to the same question. This brings back the number of possible strategies to eight; each of them can be noted (A, A', A'') , where the three numbers between the brackets are equal to ± 1 and give Albert's answers to the three possible questions. Among these eight strategies, two are $(+1, +1, +1)$ and $(-1, -1, -1)$: the answers of the first schoolboy are always the same, whatever question is asked; for the 6 other strategies $(+1, +1, -1)$, $(+1, -1, -1)$, etc. two answers are of one sort and the third is opposite.

We now assume that the experiment is repeated a large number of times (the schoolboys may change their strategy each time if they wish); one retains only the sub-ensemble of the realizations where the two questions have been different. For each such realization, two cases may occur:

- (i) Either the selected strategy is one among the two where the three answers of the first schoolboy are the same; since those of the second are always opposite, the two results are then necessarily opposite.
- (ii) Or one of the 6 other strategies has been chosen; in this case, since the three questions are random with equal probabilities $1/3$, there are two chances out of three that the two selected questions correspond to cases where the first schoolboy has planned to give opposite answers, which means that there is also a probability $2/3$ that the two teachers will get the same answer; there is a probability $1/3$ that they will get opposite answers.

Therefore, whatever strategies are chosen, the probability to get opposite answers is some number between $1/3$ and 1 , but cannot be smaller than $1/3$. In terms of classical probabilities, this provides:

$$\mathcal{P}(+1, -1) + \mathcal{P}(-1, +1) \geq \frac{1}{3} \quad (4.26)$$

Let us now return to the quantum problem with two spins in a singlet state, and assume that the three a , a' , and a'' are coplanar directions, all making 120° with each other (Figure 4.6), directions b , b' , and b'' being the same as above. Relations (4.4) show that the probability to obtain different results is:

$$\mathcal{P}(+,-) + \mathcal{P}(-,+) = \frac{1}{2} \left(\frac{1}{4} + \frac{1}{4} \right) = \frac{1}{4} \quad (4.27)$$

But this probability is smaller than $1/3$, the minimum possible value according to (4.26); we therefore obtain one more case where the predictions of quantum mechanics are incompatible with those of local realism.

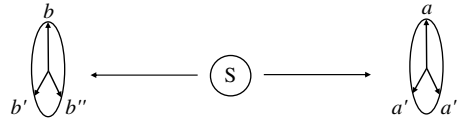


Figure 4.6. Scheme of a quantum experiment providing a large violation of inequality (4.26). The source S in the center emits two spins in a singlet state propagating to two apparatuses. Measurements of the spin components along three directions a , a' , and a'' can be performed with one particle, and along three possible directions b , b' , and b'' with the other.

4.2.2 Other sets of assumptions

We have given a derivation of the Bell theorem from one possible set of general assumptions, listed in §4.1.4. The only condition that is necessary for the derivation of the inequality is the existence of four numbers A , B , A' , and B' , all equal to ± 1 , which have to be well defined (but not necessarily known). It is therefore not surprising that other sets of assumptions could also be used to obtain the theorem [127]. This actually increases the strength of the result: a violation of the Bell inequalities means that, within each of all these sets, at least one of the assumptions should be rejected, so that the discussion of §4.1.4 can be generalized. To provide a few examples, we now list some possibilities; we do not mention the “free will assumption” (§4.5.1.c) within each set of assumptions, since it is common to all sets.

(i) We have already mentioned in the introduction of this chapter that the Bell theorem can be seen as a theorem applying specifically to theories with hidden variables. In this case, the λ arise, not from the EPR reasoning and their notion of realism, but from the a priori assumed existence of these variables. The conclusion is then that, if these variables evolve locally, the results of measurements must obey the Bell inequalities. Conversely, a violation of the inequalities means, either that the hidden variables do not exist, or that they evolve non-locally (or that there is no free will, meaning in this case that a and b are functions of additional components of λ). This point of view is less general than that we have used, but also simpler, which probably explains why it is rather popular. For instance, in one of his celebrated books on quantum mechanics [50], Jammer introduces the Bell theorem within a chapter treating specifically hidden variable theories.

This is nevertheless not the point of view taken by Bell in his article [103], where he clearly introduces his reasoning as a continuation of the EPR argument. The title is “On the Einstein–Podolsky–Rosen paradox” and the first sentences of the introduction are “The paradox of Einstein, Podolsky and Rosen was advanced as an argument that quantum mechanics could not be a complete theory but should be supplemented by additional variables. These additional variables were to restore to the theory causality and locality. In this note *that idea will be formulated mathematically* (our emphasis) and shown to be incompatible with the statistical predictions of quantum mechanics. It is the requirement of locality . . . that creates the essential difficulty”. In other words, the additional variable λ he introduces is a mathematical object associated with the EPR elements of reality. Bell therefore distinguishes between “additional variables” and “hidden variables”, which he then proceeds to discuss in the special case of Bohmian theory: “There have been attempts to show that even without such a separability or locality requirement no ‘hidden variable’ interpretation of quantum mechanics is possible. These attempts have

been examined elsewhere and found wanting. Moreover, a hidden variable interpretation of elementary quantum theory has been explicitly constructed”. Bell’s point of view is therefore not ambiguous.

(ii) Assuming counterfactuality is another way to derive the Bell theorem: the pre-existence of all the possible results of experiments allows one to derive BCHSH and other inequalities. This is because, as soon as the four quantities A , A' , B and B' (all equal to ± 1) can be defined for each realization of the experiment, the method of §4.1.2 immediately leads to (4.8) and provides a proof of the inequality. Counterfactuality is not necessarily related to the notion of space or to locality¹⁰, so that this view provides another independent logical frame for reasoning. Peres’s quotation in §4.1.3 [108] gives the conclusion to draw from a violation of the inequalities.

(iii) Assuming non-contextuality is still another possibility; this will be discussed in more detail in §5.5. Again, non-contextuality can be seen as a natural consequence of the spatial separation between Alice’s and Bob’s laboratories, but also as an independent and even more general principle (obeyed for instance in classical physics, local or non-local).

These are only a few examples of situations covered by the Bell theorem. In fact, one may combine various assumptions, such as separability, the existence of causes within relativistic past light cones, as in Figure 4.2, etc. and obtain the Bell theorem in a whole series of logical contexts. All these derivations are interesting since they extend the list of assumptions that are incompatible with the predictions of quantum mechanics, and therefore extend the range of the theorem.

4.2.3 Generalizations of the theorem, role of locality

We have already mentioned that several generalizations of the Bell theorem are possible. For instance, in some of these generalizations, it is assumed that the result of an experiment becomes a function of several fluctuating causes: the fluctuations taking place in the source, but also fluctuations in the measurement apparatuses [128], and/or perturbations acting on the particles during their motion towards the apparatuses, etc. Actually, even fundamentally indeterministic (but local) processes may influence the results, without changing the Bell inequality [99]; the theorem remains valid within local non-determinism, in situations such as that illustrated by Figure 4.7. The two former cases are almost trivial since they just require the

¹⁰ Counterfactuality can be postulated *ab initio*, without any particular reference to locality. Conversely, if one assumes locality and realism, the EPR reasoning shows the existence of elements of reality, which play the role of counterfactuals. The EPR theorem nevertheless requires more than locality only, since realism is also assumed.

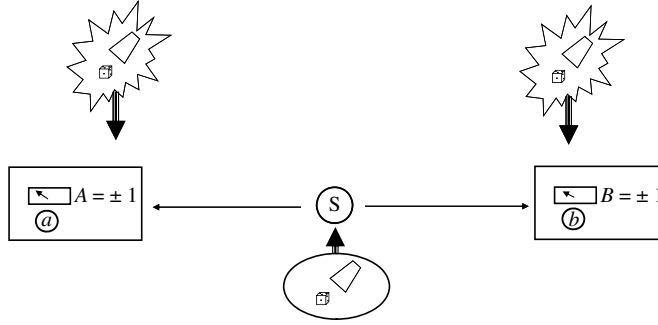


Figure 4.7. Scheme of a fundamentally indeterministic local theory. As in the preceding figures, fundamentally random processes are symbolized by flashes above the events, and the effect of uncontrolled local perturbations are represented by ovals below the events. Actually, for the present discussion, the nature of the stochastic process at the center of the figure, fundamentally indeterministic or not, is not irrelevant as long as it remains local; one may replace the central oval by a flash above the source acting only on it without changing the results. This scheme may be compared with Figure 4.5 associated with quantum mechanics, where the fundamentally random process is delocalized in space. A more detailed discussion of this type of indeterministic, but local, theory is given in Appendix C; such theories lead to predictions that necessarily obey the Bell inequalities.

addition of more dimensions to the vector variable λ ; the latter requires replacing the deterministic functions A and B by probabilities, but this is also relatively straightforward [105] (see also footnote 10 in [128] and Appendix B). Moreover, one should realize that the role of the A and B functions is just to relate the conditions of production of a pair of particles (or of their propagation) to their behavior when they reach the measurement apparatuses, and to the effects that they produce on them. They are, so to say, solutions of the equations of motion, which are assumed to be well-defined but do not need to be specified. They may perfectly include, in a condensed notation, a large variety of physical phenomena: propagation of point particles; propagation of one or several fields from the source to the detectors (see for instance the discussion in §4 of [71]); particles and fields in interaction; or whatever process one may have in mind (even random propagations can be included) – as long as they remain independent of the other setting. The exact mathematical form of the equations of propagation is irrelevant; the essential thing is that the functions exist, even if they are too complicated to be written down explicitly.

Locality, expressed mathematically in terms of a and b , is the crucial ingredient. What really matters for the proof of the Bell theorem is the dependence with respect to the settings a and b : the function A must depend on a only, while B must depend

on b only¹¹. For instance, if we wished, we could have assumed that the result A of one measurement is also a function of fluctuating random variables attached to the other apparatus, which introduces a non-local process; but this does not create any mathematical problem in the proof, as long as these variables are not affected by setting b . On the other hand, if A becomes a function of a and b (and/or the same for B), it is easy to see that the situation is radically changed: in the reasoning of §4.1.2, we must now associate eight numbers to each pair (since there are two results to specify for each of the four different combinations of settings), instead of four, so that the proof immediately collapses. Appendix C gives another concrete illustration showing that it is locality, not determinism, which is at stake; see also the appendix of [105].

An interesting generalization of the Bell theorem, where time replaces the settings, has been proposed by Franson [129] and implemented in experiments for an observation of a violation of the Bell inequalities (see for instance [130]). Another generalization shows that a violation of the Bell inequalities is not limited to a few quantum states (singlet for instance), but occurs for all states that are not products [131–134]. With statistical mixtures of states, violations of the BCHSH inequalities are also possible [135], although, conversely, entangled mixtures do not necessarily lead to violations [136]; one then has to use more elaborate sequences of measurements to reveal non-locality [137, 138]. Two independent sources can be used for observing violations, provided that appropriate interference measurements are performed [139]. The generalization is possible to an arbitrary number of particles, see [140]; for macroscopic systems, Leggett and Garg [141] have shown that BCHSH type inequalities exist for the same physical quantity considered at different times, provided two postulates are assumed: macroscopic realism, and non-invasive measurability at the macroscopic level. At the other extreme, schemes have been proposed to observe non-locality even with a single photon [142–144] (but, in the measurement process, more than one photon is involved, for example photons from local oscillators). For a general discussion of the conceptual impact of a violation of the inequalities, we refer to the book collecting Bell’s original articles [6].

¹¹ The fact that function A does not depend on variable b (and that function B does not depend on variable a) does not mean that the results observed on one side are uncorrelated with the results at the other side. In general, they are actually correlated through the fluctuations of variable λ , in a way that depends on the statistical distribution $\rho(\lambda)$ of this variable. In other words, strong correlations may well be introduced by fluctuations of a common cause in the past, as in any usual statistical theory.

In §4.4.2.b, we discuss stochastic generalizations of the theorem and the role of both the “setting dependence” and of the “outcome independence” (the factorization of the probability for a given value of λ , when the fluctuations of common causes are frozen).

4.2.4 Status of the theorem; attempts to bypass it

Provided locality is assumed, the Bell theorem is very general; it is therefore difficult to build a reasonable theory that can violate the inequalities, more difficult than one might think. All potential authors who believe that they have found a simple explanation for the observed violations should think twice before taking their pen and sending a manuscript to a physics journal. Every year a large number of such texts is submitted, with the purpose of introducing “new” ways to escape the constraints of the Bell theorem, and to “explain” why the experiments have provided results that are in contradiction with the inequalities. For instance, the violations could originate from some new sort of statistics, or from perturbations created by cosmic rays, gas collisions with fluctuating impact parameters, random effects of gravity, etc. Imagination is the only limit to the variety of the processes that can be invoked. Nevertheless, we know from the beginning that all attempts to obtain violations within classical local theories are doomed to failure, however elaborate these theories are.

The situation is somewhat reminiscent of the attempts in past centuries to invent “perpetuum mobile” devices: even if some of these inventions were extremely clever, and if it is sometimes difficult to find the exact reason why they should not work, it remains true that the law of energy conservation allows us to know immediately that they cannot function. In the same way, some of these statistical “Bell beating schemes” may be extremely clever, but we know that the theorem is a very general result in statistics: in all situations that can be accommodated by the mathematics associated with the variables λ and the A and B functions (and there are many!), it is impossible to escape the inequalities. No, non-local correlations cannot be explained cheaply; yes, a violation of the inequalities is a very, very, rare situation. In fact, until now, it has never been observed, except of course in experiments designed precisely for this purpose. In other words, if we wanted to build automata including arbitrarily complex mechanical systems and computers, we could never mimic the results predicted by quantum mechanics (at least for remote measurements) without allowing communication between them; it is even possible to calculate the minimum amount of information that must be exchanged to simulate the quantum correlations [145]. This will remain impossible forever – or at least until completely different computers working on purely quantum principles are built¹².

The only way to bypass the Bell theorem is to explicitly give up at least one of the assumptions listed in §4.1.4; in §4.5.1 we come back to one of them, the free will assumption.

¹² In terms of the Mendel parable: an observation of a violation of the Bell inequalities would imply that something inside both peas (maybe a pair of DNA molecules?) remains in a coherent quantum superposition, without decoherence, even if the distance between the peas is large.

4.3 Cirelson's theorem

In §4.1.2, we have introduced a combination $\langle M \rangle$ of four averages that can never exceed 2 within to local realism (BCHSH inequality). In §4.1.3 we have seen that, in quantum mechanics and with two spins in a singlet state, the corresponding combination of averages $\langle Q \rangle$ can reach $2\sqrt{2}$. This is of course already a large violation of the BCHSH inequality, but one could nevertheless hope to find other quantum systems or other measurements that lead to even larger violations. After all, each of the four average values $\langle AB \rangle$, $\langle A'B \rangle$, etc. appearing in $\langle Q \rangle$ is bounded between -1 and $+1$; if these averages were independent variables, mathematically the maximal value accessible to $\langle Q \rangle$ would be 4, larger than $2\sqrt{2}$. One could then wonder if better quantum situations could be found where the violation is larger than what is obtained with two spins in a singlet state.

Within quantum mechanics, Cirelson's theorem [146, 147] shows that, with a system made of two sub-systems, it is actually impossible to go beyond this value of $2\sqrt{2}$, whatever series of measures is made on the two sub-systems and whatever the initial state of the whole system is.

4.3.1 Measurements on two-level sub-systems

Consider a physical system made of two quantum sub-systems on which separate measurements can be performed, each leading to two possible results, $+1$ and -1 . Each of the sub-systems has a space of states that is spanned by two levels which, without loss of generality, we can consider as the two eigenstates of the O_z components of a (pseudo) spin $1/2$. We denote $\sigma_x(1)$, $\sigma_y(1)$, and $\sigma_z(1)$ the three components for the first sub-system – cf. equation (11.52) of Chapter 11 – which we group into a vectorial operator $\vec{\sigma}(1)$. A similar notation $\sigma_x(2)$, $\sigma_y(2)$, $\sigma_z(2)$, and $\vec{\sigma}(2)$ is taken for the second sub-system. Any measurement performed on the first sub-system corresponds to an Hermitian operator $O(1)$ acting in its space of states, which can be expressed as a linear combination of the three components of $\vec{\sigma}(1)$ and of the unit operator $\hat{I}(1)$ in the form:

$$O(1) = \alpha \hat{I}(1) + \vec{a} \cdot \vec{\sigma}(1)$$

where α and the three components of $\vec{a}$ are real scalar parameters. Nevertheless, since the results of the measurement can only be ± 1 , the two eigenvalues must have the same values, so that $\alpha = 0$ and $|\vec{a}| = 1$. We then denote $\sigma_a(1)$ and $\sigma_{a'}(1)$ the spin operators associated with two kinds of measurement performed on the first sub-system with the settings $\vec{a}$ and $\vec{a}'$:

$$\sigma_a(1) = \vec{a} \cdot \vec{\sigma}(1) \quad \sigma_{a'}(1) = \vec{a}' \cdot \vec{\sigma}(1) \quad (4.28)$$

and similarly, for the second sub-system:

$$\sigma_b(2) = \vec{b} \cdot \vec{\sigma}(2) \quad \sigma_{b'}(2) = \vec{b}' \cdot \vec{\sigma}(2) \quad (4.29)$$

(the norm of vectors $\vec{a}$, $\vec{a}'$, $\vec{b}$, and $\vec{b}'$ are all 1). Since the square of any Pauli matrix gives the identity operator $\hat{I}$, we have:

$$[\sigma_a(1)]^2 = [\sigma_{a'}(1)]^2 = \hat{I}(1) \quad ; \quad [\sigma_b(2)]^2 = [\sigma_{b'}(2)]^2 = \hat{I}(2) \quad (4.30)$$

The quantum average $\langle Q \rangle$ that generalizes (4.10) is given by $\langle \Psi | Q | \Psi \rangle$, where $|\Psi\rangle$ is any normalized state of the two-particle system, and where Q is the operator given by:

$$Q = [\sigma_a(1)] [\sigma_b(2)] - [\sigma_a(1)] [\sigma_{b'}(2)] + [\sigma_{a'}(1)] [\sigma_b(2)] + [\sigma_{a'}(1)] [\sigma_{b'}(2)] \quad (4.31)$$

Our purpose is to show that the modulus of the average value of this operator cannot exceed $2\sqrt{2}$, whatever choice is made for $|\Psi\rangle$ and the four vectors $\vec{a}$, $\vec{b}$, $\vec{a}'$, and $\vec{b}'$.

4.3.2 Maximal quantum violation

We first calculate the square of this operator, which contains three kinds of terms: the squares of each of the four operators appearing in (4.31), the crossed terms where one of the operators between brackets occurs twice, and finally the crossed terms where all σ operators are different. Since the squares of the Pauli matrices are $\hat{I}$, the contribution of the first kind of terms is:

$$4 \times \hat{I} \quad (4.32)$$

As for the terms implying one operator σ twice, they give:

$$\begin{aligned} & \left[-[\sigma_a(1)]^2 + [\sigma_{a'}(1)]^2 \right] \left[\sigma_b(2)\sigma_{b'}(2) + \sigma_{b'}(2)\sigma_b(2) \right] \\ & + \left[[\sigma_b(1)]^2 - [\sigma_{b'}(1)]^2 \right] \left[\sigma_a(2)\sigma_{a'}(2) + \sigma_{a'}(2)\sigma_a(2) \right] \end{aligned} \quad (4.33)$$

which vanishes since the squares of all Pauli matrices are equal to the unit matrix (§11.1.6). The crossed terms containing four different σ operators give:

$$\begin{aligned} & \sigma_a(1)\sigma_{a'}(1)\sigma_b(2)\sigma_{b'}(2) + \sigma_{a'}(1)\sigma_a(1)\sigma_{b'}(2)\sigma_b(2) \\ & - \sigma_a(1)\sigma_{a'}(1)\sigma_{b'}(2)\sigma_b(2) - \sigma_{a'}(1)\sigma_a(1)\sigma_b(2)\sigma_{b'}(2) \end{aligned} \quad (4.34)$$

which is nothing but the product of two commutators:

$$[\sigma_a(1), \sigma_{a'}(1)][\sigma_b(2), \sigma_{b'}(2)] \quad (4.35)$$

Collecting all these results, we obtain the equality [147]:

$$Q^2 = 4 \times \hat{I} + [\sigma_a(1), \sigma_{a'}(1)][\sigma_b(2), \sigma_{b'}(2)] \quad (4.36)$$

The commutation relation of the Pauli matrices:

$$[(\vec{a} \cdot \vec{\sigma}), (\vec{a}' \cdot \vec{\sigma})] = 2i (\vec{a} \times \vec{a}') \cdot \vec{\sigma}$$

can then be used to arrive at:

$$Q^2 = 4 \times \hat{I} - 4 \left[\left((\vec{a} \times \vec{a}') \cdot \vec{\sigma} \right) (1) \right] \left[\left((\vec{b} \times \vec{b}') \cdot \vec{\sigma} \right) (2) \right] \quad (4.37)$$

The operator $(\vec{a} \times \vec{a}') \cdot \vec{\sigma} (1)$ has eigenvalues $\pm |\vec{a} \times \vec{a}'|$, which in general have a modulus smaller than 1 since the length of vector $\vec{a} \times \vec{a}'$ cannot exceed 1; similarly, operator $(\vec{b} \times \vec{b}') \cdot \vec{\sigma} (2)$ has eigenvalues with modulus equal or smaller than 1. Therefore, the modulus of average value of this product of operators cannot exceed 1, so that:

$$\langle Q^2 \rangle = \langle \Psi | Q^2 | \Psi \rangle \leq 8 \quad (4.38)$$

But the square of the average $\langle Q \rangle^2$ of an Hermitian operator is always smaller¹³ than the average value of its square $\langle Q^2 \rangle$. We therefore always have:

$$-2\sqrt{2} \leq \langle Q \rangle \leq 2\sqrt{2} \quad (4.39)$$

This inequality is Cirelson's theorem.

Remarks:

(i) the proof of this inequality is useful to predict conditions under which the bound $\pm 2\sqrt{2}$ can be reached. We notice in particular that the vectors $\vec{a} \times \vec{a}'$ and $\vec{b} \times \vec{b}'$ should have unit length, which requires that $\vec{a}$ and $\vec{a}'$ should be perpendicular, as well as $\vec{b}$ and $\vec{b}'$ (cf. note 3). Moreover state $|\Psi\rangle$ should be an eigenvector with eigenvalue -1 of the product of the component of the first spin along axis $\vec{a} \times \vec{a}'$ by the component of the second spin along axis $\vec{b} \times \vec{b}'$; in the coplanar configuration of the four vectors, this corresponds to spin components along the same axis, which we may call Oz . Indeed, a singlet state is such an eigenstate of $\sigma_z(1)\sigma_z(2)$ – but this is also the case of a triplet state with zero

¹³ This Schwartz inequality is obtained by writing that the average value of $[Q - \langle Q \rangle]^2$ is positive; it appears in the definition of the square of the mean square value ΔQ .

component along Oz , for which one can check that the same violation is obtained as with the singlet state.

(ii) In §5.3, we study a case where, by considering three sub-systems instead of two, one can find cases where quantum mechanics predicts that $|\langle Q \rangle|$ exceeds $2\sqrt{2}$ and reaches the mathematical limit 4.

4.4 No instantaneous signaling

The theory of relativity requires that it should be fundamentally impossible to transmit signals containing information between two distant points faster than the speed of light (relativistic causality); suppressing this absolute limit would lead to serious internal inconsistencies in the theory. But one could then wonder if a violation of the BCHSH inequalities does not imply the possibility of a communication at an arbitrary velocity, since the distance between the two operators Alice and Bob has no influence on the appearance of correlations between the results of distant measurements. Indeed, from a local realist point of view, we have seen that the only possibility to reproduce the predictions of quantum mechanics is to include an instantaneous influence of the value of a local experimental setting on the result at the other end. Moreover, even within standard quantum mechanics, the instant determination of the state of the second particle by a measurement performed on the first seems to indicate an influence at an arbitrary distance. We now check that quantum mechanics does not lead to such a possibility of instantaneous signal transmission.

What methods of transmission could be envisaged? The first idea that comes to mind is to imagine a system having some analogy with the Morse telegraph, where the results $+1$ and -1 are used by Alice to code a message to be sent to Bob (for the sake of simplicity, here we assume that each of the measurements can only provide two results, but a generalization is possible). But, obviously, such a system would not work since the results of measurements are completely random, so that Alice has no control of them to code any message. A telegraphic system can work only if what is used to code a message is chosen by the experimenters (for instance, the experimental settings instead of the results). Is it then possible for Bob to determine the value a or a' of the setting that Alice has chosen, knowing that Bob can only make local observations in his laboratory, and that he is completely free to choose whatever value for his measurement setting b ? This signaling technique would not imply any particular delay for the transmission of information.

We now discuss under which conditions a conflict with relativity can be avoided. For a deterministic theory, we will see that this amounts to a condition of “setting independence”; for a stochastic theory, this also introduces so called

“non-signalling” conditions, or if more stringent conditions of locality are assumed, a condition of “outcome independence”. See [148] and [149] for a discussion in terms of “strong locality” and “predictive completeness” (or “parameter independence” and of “outcome independence” in [150]).

4.4.1 Non-signaling (NS) conditions

From a general point of view, and without restricting the discussion to quantum mechanics, what are the general mathematical conditions ensuring that a theory is “non-signaling” (NS conditions), meaning that it does not allow Alice to transmit instantaneous signals to Bob (and conversely)?

For a deterministic theory, these conditions are very simple: the result $A = \pm 1$ of each measurement performed by Alice must depend only on the measurement setting a , while the result $B = \pm 1$ of each measurement performed by Bob must depend only on b . This is the condition of “setting independence”.

For a stochastic theory, and for each well-defined experimental setup (as meant by Bohr) corresponding to given measurements settings a and b , four probabilities $\mathcal{P}(A, B|a, b)$ are associated with the four events $A = \pm 1$ and $B = \pm 1$, with a unit sum 1:

$$\sum_{A, B} \mathcal{P}(A, B|a, b) = 1 \quad \text{for every couple } a, b \quad (4.40)$$

In what follows, we will assume that the first setting can take only two values a and a' , while the second can take only two values b and b' , so that four experimental setups are possible for the whole experiment. We then have 16 probabilities obeying 4 sum relations similar to (4.40); the most general probabilistic model therefore depends on 12 free parameters.

When the experiment is repeated, since Bob does not have access to Alice’s results, the only thing that he can measure is the occurrence frequency of his own results. This corresponds to the probabilities obtained by a summation over A of the preceding probabilities (sum of probabilities associated with exclusive events):

$$\sum_A \mathcal{P}(A, B|a, b) \quad (4.41)$$

The NS condition amounts to assuming that this probability is independent of a ; we therefore obtain the condition:

$$\sum_A \mathcal{P}(A, B|a, b) = \sum_A \mathcal{P}(A, B|a', b) \quad \text{for any value of } b \quad (4.42)$$

One could write this relation for any value of B but, when only two results can occur, the sum rule (4.40) implies that the two conditions for different values of B are not independent (their sum provides $1 = 1$); we can therefore retain only one condition, for one value of B . Since we have assumed two different values of b , the NS condition relative to the Alice-to-Bob communication channel then implies two linear relations (4.42) between the probabilities. For the reverse communication channel, one gets two other relations:

$$\sum_B \mathcal{P}(A, B|a, b) = \sum_B \mathcal{P}(A, B|a, b') \quad \text{for any value of } a \quad (4.43)$$

Altogether, we therefore have four NS conditions.

These conditions are indeed satisfied in quantum mechanics, when Alice and Bob perform local experiments in two remote laboratories, without any possible influence of one of the sub-systems on the other. For the singlet state studied in §4.1.1, we immediately obtain $\mathcal{P}_{(+,\pm)} + \mathcal{P}_{(-,\pm)} = 1/2$ and $\mathcal{P}_{(\pm,+)} + \mathcal{P}_{(\pm,-)} = 1/2$, which shows that relations (4.42) and (4.43) are obeyed. The general proof is given in Appendix F, whatever entangled state is chosen for the quantum system, and whatever measurements are performed by Alice and Bob. This property of quantum mechanics avoids a frontal conflict with relativistic causality.

4.4.2 Logical boxes

Following Popescu and Rohrlich [151] (see also [152]), we define logical “boxes” as devices with which Alice and Bob can arbitrarily choose a binary value of “input variables”, $a = \pm 1$ for Alice and $b = \pm 1$ for Bob; the box returns to them binary “output variables” $A = \pm 1$ and $B = \pm 1$ (left part of Figure 4.8). We will distinguish between deterministic boxes, where A and B are given functions of a and b , and stochastic boxes defined by probabilities for the output variables that depend on a and b .

4.4.2.a Deterministic boxes

When the input variables can take only two values, 4 distinct experimental setups are possible for the whole system; when the output variables can take only two values, each setup can provide 4 different pairs of results. The number of distinct deterministic boxes is therefore $4^4 = 256$. But, if the NS condition is satisfied, output A must be a function of a only, so that 4 different functions $A(a)$ are possible for Alice (since she uses two values of her variable a), 4 also for Bob; the total number of deterministic boxes is only 16.

4.4.2.b Stochastic boxes

The number of stochastic boxes is not finite, since they depend on continuous parameters. A stochastic box associates to each experimental setup 4 probabilities

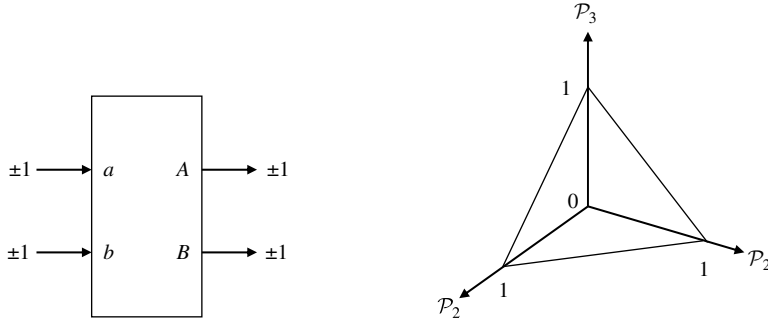


Figure 4.8. The left part of the figure gives a schematic representation of a logical box which, from the values ± 1 of the input variables a and b , provides the values of the output variables $A = \pm 1$ and $B = \pm 1$. The box is deterministic if A and B are given functions of a and b , stochastic if A and B are determined by a and b dependent probabilities.

The right part of the figure illustrates how a stochastic logical box is characterized by a point belonging to a polytope in probability space $\mathcal{P}_i(a, b)$ ($j = 1, \dots, 4$). For any given value of the input variables a and b , the sum of probabilities is equal to one, so that it is sufficient to plot the three first probabilities on the three axes; the point associated with the logical box belongs to the inside (or surface) of a tetrahedron with unit sides. Since four values are possible for the (a, b) couple, the polytope is the product of four tetrahedra. The deterministic boxes correspond to points at the corners.

$\mathcal{P}_i$ ($j = 1, \dots, 4$) with a unit sum. If we choose three axes $\mathcal{P}_1, \mathcal{P}_2, \mathcal{P}_3$, each box is associated with a point in the corresponding three-dimensional space, while $\mathcal{P}_4$ is given by $1 - \mathcal{P}_1 - \mathcal{P}_2 - \mathcal{P}_3$. Since all probabilities $\mathcal{P}_i$ remain non-negative, the corresponding point therefore lies inside (or at the surface) of a tetrahedron having a corner at the origin and the three other corners on the axes at a unit distance from the origin (right part of Figure 4.8). Each of the corners corresponds to a deterministic box. Since the box characterizes in fact the output values A and B for four different values of the pairs of input variables a and b , it is actually defined by four independent points inside (or at the surface of) four tetrahedra. We can group these four points into one single point G in a space with $4 \times 3 = 12$ dimensions. The space that is accessible to G is a polytope, with boundaries corresponding to one vanishing probability, and corners associated with deterministic situations. A general discussion of such polytopes is given in [152].

For a NS box, the four points are no longer independent, since they must satisfy the four relations (4.42) and (4.43); the stochastic NS boxes therefore depend on eight parameters, and the new polytope accessible to the parameters is smaller than before; in particular, the corners are limited to those associated with deterministic NS boxes.

4.4.2.c Local stochastic boxes

A sub-category of the stochastic NS boxes is given by the local stochastic boxes, which have the property of “outcome independence”¹⁴. Among them the simplest are those with probabilities that factorize:

$$\mathcal{P}(A, B|a, b) = p(A|a) \times p(B|b) \quad (4.44)$$

where $p(A|a)$ and $p(B|b)$ are two local probabilities between 0 and 1 obeying two separate normalization conditions:

$$\sum_A p(A|a) = 1 \quad \sum_B p(B|b) = 1 \quad (4.45)$$

The corresponding boxes depend on $2 \times 2 = 4$ continuous parameters only.

But, in the spirit of the Bell theorem, it is possible to generalize them by assuming the presence of fluctuating causes characterized by a random variable λ and a normalized distribution $\rho(\lambda)$:

$$\int d\lambda \rho(\lambda) = 1 \quad (4.46)$$

One then assumes that, for a given value of λ (which may be a multidimensional parameter), the conditions of the box are sufficiently well defined to obtain a factorization from a locality argument:

$$\mathcal{P}_\lambda(A, B|a, b; \lambda) = p_\lambda(A|a; \lambda) \times p_\lambda(B|b; \lambda) \quad (4.47)$$

Now the local probabilities p_λ and q_λ depend on the statistical parameter λ , while satisfying (4.45) for each value of λ . When λ fluctuates, the probabilities become:

$$\mathcal{P}(A, B|a, b) = \int d\lambda \rho(\lambda) p_\lambda(A|a) \times p_\lambda(B|b) \quad (4.48)$$

In general, this probability does not factorize, as opposed to (4.44) and (4.47). These boxes are classical, and can be seen as a natural continuation of local realism as introduced by EPR and Bell; they obey the Bell theorem. It is easy to check that these boxes are NS since:

$$\sum_A \mathcal{P}(A, B|a, b) = \int d\lambda \rho(\lambda) p_\lambda(B|b) \quad (4.49)$$

which is indeed independent of a (the proof is similar for the B summation).

¹⁴ The Bayes theorem provides the general relation $\mathcal{P}(A, B|a, b) = p(A|a, b) \times p(B|A, a, b)$, where $p(B|A, a, b)$ is the conditional probability to obtain B if result A has been obtained. Assuming “setting independence” amounts to replacing $p(A|a, b)$ by $p(A|a)$ as well as $p(B|A, a, b)$ by $p(B|A, b)$. Assuming “outcome independence” amounts to writing that this conditional probability is independent of the outcome A , and therefore leads to (4.44). The same assumptions can be made when one has to take into account the fluctuating variable λ , and then lead to (4.47).

Quantum mechanics can be used to build stochastic NS boxes, since it obeys relations (4.42) and (4.43), but which do not necessarily belong to the category of local stochastic boxes with outcome independence – otherwise it could not lead to violation of the Bell inequalities.

4.4.3 Popescu–Rohrlich boxes

Any quantum scheme with experiments depending on two experimental settings a and b and leading to binary results can be used to calculate probabilities and then define a logical box, called a quantum box. We mentioned above that conditions (4.42) and (4.43) are always satisfied; all quantum boxes are therefore NS. One can then ask the reverse question: can any logical NS box be a quantum box? Is it possible, by starting from an appropriate quantum state for two sub-systems and performing adequate measurements on them, to reconstruct any NS box? We will see that the answer to this question is no: the category of theories leading to logical boxes that remain compatible with relativity is actually larger than quantum mechanics.

Introducing the “Popescu–Rohrlich box” [151] is a way to prove this result by producing an example, a box that is sometimes called a “PR box”. The idea is simple: in expression (4.10), we obtained a quantum violation by considering values of the input variables (settings) for which the three cosines with the same sign have value $1/\sqrt{2}$, while that with opposite sign has value $-1/\sqrt{2}$, so that the sum reaches $2\sqrt{2}$; we can try to increase this sum even more by building a model where the correlation rates are enhanced to their maximal value 1. For the 3 pairs of values (a, b) , (a', b) and (a', b') of the input variables, we then assume that the output variables are perfectly correlated, and therefore always equal; moreover, to preserve the NS conditions at best, we minimize the transmission of information by attributing equal probabilities to the two pairs of possible results $(-1, -1)$ and $(+1, +1)$

$$\mathcal{P}(-1, -1) = \mathcal{P}(+1, +1) = 1/2 \quad (4.50)$$

so that the probability of a different value vanishes:

$$\mathcal{P}(-1, +1) = \mathcal{P}(+1, -1) = 0 \quad (4.51)$$

For the couple (a, b') , the situation is the opposite, and the output variables are never equal:

$$\begin{aligned} \mathcal{P}(-1, -1) &= \mathcal{P}(+1, +1) = 0 \\ \mathcal{P}(-1, +1) &= \mathcal{P}(+1, -1) = 1/2 \end{aligned} \quad (4.52)$$

We now check that this box is NS. In both the right- and left-hand sides of (4.42), only one term contributes in the summation over A if the probabilities (4.50), (4.51), and (4.52) are assumed. This is because the same value of B is never associated to two different values of A , for any experimental setting; the two sides have therefore value $1/2$ and the equality holds. The same is true of (4.43). The Popescu–Rohrlich box thus does not permit instantaneous transmission of signals; it is compatible with relativity¹⁵.

Let us now calculate the value obtained for the combination of products of results appearing in (4.7):

$$\langle M \rangle = \langle AB \rangle(a, b) - \langle AB \rangle(a, b') + \langle AB \rangle(a', b) + \langle AB \rangle(a', b') \quad (4.53)$$

The calculation is very simple since the probabilities we have chosen imply that the product AB is always $+1$ for the three pairs of inputs (a, b) , (a', b) , and (a', b') , while relation (4.52) implies that this product is always -1 for (a, b') . Therefore:

$$\langle M \rangle = 4 \quad (4.54)$$

which shows that the mathematical limit for $\langle M \rangle$ is indeed saturated.

The fact that the Cirelson bond (§4.3) is exceeded immediately signals an incompatibility with quantum mechanics: no quantum setup with two sub-systems undergoing measurements, no initial quantum state can reproduce a Popescu–Rohrlich, even if such a box forbids instantaneous transmission of signals. This box¹⁶ therefore provides an example of “superquantum” correlations.

It is also possible to simulate such a box by using post-selection of detected events, which amounts to introducing a bias in the sampling of detected pairs (we come back to this notion in more detail in §4.5.1.a). With a sufficient selection, one can then not only exceed the Cirelson limit $2\sqrt{2}$ but virtually reach the mathematical limit of 4 [153–155].

4.4.4 How to characterize quantum mechanics?

Popescu and Rohrlich propose new axioms to introduce quantum mechanics [151]. Instead of considering indeterminism as a fundamental postulate, as one usually

¹⁵ A Popescu–Rohrlich box also obeys the no-cloning theorem (§7.2.1).

¹⁶ Physically, it may seem more plausible to consider continuous input variables, instead of binary variables. This is because they correspond to experimental settings (such as angles of measurement for instance) which, in most cases, are continuous. Assuming relations (4.50) to (4.52) for all values of the input variables would then lead to discontinuities. But, in quantum mechanics, the maximal violation $2\sqrt{2}$ is obtained only for some values of the variables. Similarly, here, one can assume that these probability relations are satisfied only for some values of the parameters, and continuously interpolate between them in a second step; the violations are then limited to some range of the variables, but this does not change the proof of existence of superquantum correlations.

does in quantum mechanics, they suggest choosing two different axioms: the theory is non-local (meaning that it sometimes predicts violations of the BCHSH inequalities), but nevertheless remains compatible with relativistic causality. Non-determinism is then a consequence of these postulates, as a “reductio ad absurdum” reasoning shows. Indeed, if the theory was deterministic, it would provide results A and B as functions of the experimental parameters a and b , and the compatibility with relativistic causality would immediately imply that A depends on a only, and B on b only; the proof of §4.1.2 would then apply, forbidding any violation of BCHSH inequalities. In other words, within determinism, a BCHSH violation would immediately create another violation of relativistic causality; it is quantum indeterminism that avoids this conflict.

Nevertheless, the example of the Popescu–Rohrlich box shows that these axioms do not define the theory in a unique way. What is defined is a broader ensemble of theories, among which quantum mechanics with the Cirelson bound $2\sqrt{2}$, while other theories can reach the mathematical limit 4. A general discussion of the properties of non-signaling theories is given in [156]: intrinsic randomness, impossibility of perfect cloning, monogamy, etc. (we come back to the meaning of these terms in §§6.2.4 and 7.2.1). One can nevertheless take a point of view where the BCHSH inequalities and the degree of their violation are more considered as indicators of the degree of correlations permitted by a theory than as a proof of non-locality. The question is then: why are the correlations predicted by quantum mechanics not maximal among those that remain compatible with relativistic causality, but more restricted? What is the additional physical principle obeyed by quantum mechanics that selects the value $2\sqrt{2}$?

A partial answer to this question is given in [157] by remarking that stronger correlations would result in a world in which communication complexity is not trivial¹⁷. It has been proposed in [158] that the principle in question is a generalization of the principle of non-instantaneous communication: for any quantum system shared by Alice and Bob, if she sends to him m bits through a classical channel, whatever local operations and measurements Bob makes (these operations may depend on the received bits) he cannot obtain information that exceeds m .

In this scheme, Alice receives a chain of N binary numbers a_i ($i = 0, 1, \dots, N-1$); her objective is to transmit as much information as possible concerning this chain to Bob, knowing that she can only send m bits to him through a classical channel ($m < N$). She chooses some of the numbers of her chain as values for the measurement parameters she performs (inputs for her side of the box), and obtains

¹⁷ Alice and Bob wish to compute some Boolean function $F(a, b)$ of variable a , chosen by Alice and known to her only, and of variable b chosen by Bob (and known to him only). The communication complexity of F is said to be trivial if the operation can be performed by the transmission of a single classical bit of communication.

outputs (results); the message she sends to Bob can be built by combining her inputs and outputs in an arbitrary way. Bob, when he receives the message, is free to use its bits to perform any operation with his own logical system, in order to obtain the maximum possible number of values of the a_i .

- (i) We first assume that Alice and Bob share a logical system made of Popescu–Rohrlich boxes. One can then show [158] that such a device allows Bob to exactly determine the value of any series of m different a_i , and to choose which ones – but never to determine more than m values. All bits in Alice’s data base are accessible to Bob, but the number of bits he can acquire is strictly limited to m .
- (ii) Now assume that the system initially shared by Alice and Bob is a quantum system in any state. It is shown in [158] that there exists a relation between the rate of violation of the BCHSH inequality and the maximum amount of information obtained by Bob on the values of the series of a_i . It defines the information causality principle as: “the information gain that Bob can reach about a previously unknown to him data set of Alice, by using all his local resources and m classical bits communicated by Alice, is at most m bits”. It then shows that the information causality principle is violated exactly when the Cirelson limit $2\sqrt{2}$ is exceeded.

The information causality principle is, so to say, a principle of non-amplification of classical information. For $m = 0$, one recovers the NS condition of relativistic causality: no information at all can pass from Alice to Bob through simple local measurements that they can perform, each in his/her own laboratory. For larger values of m , the new principle posits that, for any local measurements performed by the two partners, the amount of transmitted information is not increased; these measurements are, in a way, useless. This provides an explanation of the Cirelson bond. It then becomes natural to consider that the information causality principle belongs to those defining quantum mechanics, or even to conjecture that it may be used to perfectly define the theory.

The authors of [159] propose a different approach to the characterization of quantum non-local correlations. Assuming that quantum mechanics holds locally, they show that the no-signalling condition then implies that all possible correlations between the two distant parties should be those of quantum mechanics as well. In other words, if any experiment provided non-local correlations beyond those of quantum mechanics, relativity would imply that quantum mechanics could not be valid locally either. The reference [160] generalizes the characterization of quantum correlations to those observed between N partners, instead of two. It discusses a non-local game where each partner is

given an input variable x_i and where, by using quantum correlations provided by a common quantum system in an entangled state $|\Psi\rangle$, each partner tries to guess the x_j variable of his neighbor. It turns out that, in this case, quantum correlations do not perform better than classical correlations. This result suggests that quantum mechanics might be characterized by a multi-partner no-improvement criterion, a generalization of the non-signaling principle.

4.5 Impact of the theorem: where do we stand now?

In view of the experimental results mentioned in §4.1.5, which were not available when the orthodox interpretation of quantum mechanics was invented, and which are in complete agreement with the predictions of quantum mechanics, some physicists conclude triumphantly: “Bohr was right!”. Others claim with the same enthusiasm “Einstein was right!” and will emphasize his precursor role in one more domain of physics. Both these opinions make sense, depending on what aspect of the debate one favors. Whether one personally feels closer to the orthodox quantum camp or to local realism, it remains clear that the line initiated by Einstein and Bell had the decisive role to play over the last 50 years. They are the ones who pointed out the crucial importance of the notion of locality in all these discussions, which resulted in much more progress and understanding than the simple re-statement of the orthodox position. For instance, even now, the introduction of the reduction of the state vector is sometimes “explained” by invoking the “unavoidable perturbations that the measurement apparatus brings to the measured system” – see for instance the traditional discussion of the Heisenberg microscope which still appears in textbooks! But, precisely, the EPR–Bell argument shows us that this is a cheap explanation: in fact, the quantum description of a particle can be modified without any mechanical perturbation acting on it, provided the particle in question was previously correlated with another particle. So, a trivial effect such as a classical recoil effect in a photon–electron collision cannot be the real explanation of the deep nature of the state vector reduction! It is much more fundamentally quantum and may involve non-local effects.

Another lesson is that, even if quantum mechanics and relativity are not incompatible, they do not fit very well together: the notion of events in relativity, which are supposed to be point-like in space-time, or the idea of causality are still basic notions, but not as universal as one could have thought before the Bell theorem. Indeed, quantum mechanics teaches us to take these notions “with a little grain of salt”. Still another aspect is related to the incredible progress that experiments have made in the twentieth century, whether or not stimulated by fundamental quantum mechanics. One gets the impression that this progress is such that it will

allow us to have access to objects at all levels of scale, ranging from the macroscopic to the microscopic. Therefore, while in Bohr's time one could argue that the precise definition of the border line between the macroscopic world of measurement apparatuses and microscopic objects was not crucial, or even academic, the question may become of real importance. In §3.3.4, we have given one example (macroscopic systems in Fock states) but many other possibilities may emerge and, hopefully, even give rise to experiments soon. All these changes, together, give the impression that the final stage of the theory is not necessarily reached and that conceptual revolutions are still possible, even if for the moment no precise new result has weakened the orthodox interpretation in any way.

4.5.1 Loopholes, conspiracies

One sometimes hears that the experiments that have been performed so far to test the Bell inequalities against quantum mechanics (§4.1.5) are not perfectly convincing; therefore, local realism has not been disproved. Strictly speaking, this is true: there are indeed logical possibilities, traditionally called “loopholes”, which are still open for those who wish to keep strict local realism in physics. There are several ways in which one can deny the existence of any real conflict between the experimental results and the Bell inequalities.

First of all, of course, one can always invoke trivial errors, such as very unlikely statistical fluctuations, to explain why the experiments seem to “mimic” quantum mechanics so well; for instance, some authors have introduced *ad hoc* fluctuations of the background noise of photomultipliers, which would magically correct the results in a way that would give the impression of exact agreement with quantum mechanics. One could as well even assume that all experimental results in physics are erroneous fluctuations! But the number and variety of Bell-type experiments supporting quantum mechanics with excellent accuracy is now large; in view of the results, few physicists take this explanation very seriously.

One could also think of more complicated scenarios: for instance, some local unknown physical variables may couple together in a way that will give the (false) impression of non-local results, while the mechanism behind them remains local. We now discuss some of these scenarios.

4.5.1.a Pair selection loophole (efficiency loophole)

In the proof of the Bell theorem, we have assumed that all pairs are detected and provide results A and B equal to ± 1 , whatever values for a and b are chosen. While keeping a local realist point of view, one can give up this assumption and consider that, somehow, the polarization analyzers select a subclass of pairs that depend on

their settings. For each choice (a, b) , only a fraction of the emitted pairs is then detected; one can even assume that, as soon as the orientation of one (or both) of the analyzers is changed by a few degrees, a completely different category of physical systems is selected. Then, depending on the choice of a made on one side, a different category of particles will reach the detector; it then becomes possible to ascribe to this (a dependent) category whatever *ad hoc* physical properties are needed to reproduce any result, including those of quantum mechanics. Clearly, the Bell limits do not apply to such situations. In other words, if on each side one counts only events that are correlated with a detection on the other side, one actually counts particles that may have properties that depend on the other remote setting, and locality can no longer be expressed in simple terms of a and b (see Appendix E for more details).

The Bell inequalities therefore apply only if the ensemble of detected pairs can be considered as independent of a and b ; this important point was realized before the first experiments were performed [99, 161, 162]. Therefore, any experiment leading to a violation of the Bell inequalities, but without any direct experimental proof of this independence, can in principle be interpreted within local realism by invoking what is often called the “pair selection loophole” (also called “detection loophole”, “efficiency loophole” or “(un)fair sampling loophole”, etc.).

Definition of the ensemble A crucial point is therefore the existence of averages over a non-biased, well-defined, ensemble of pairs (completely independent of the settings a and b). In terms of the probability distribution $\varrho(\lambda)$ of pairs, which defines the ensemble on which the averages are taken, the non-biasing assumption is equivalent to considering ρ as independent of a and b ; conversely, if ρ becomes a function of a and b , it is easy to see that the proof of the Bell inequalities is no longer valid. In the reasoning of §4.1.2, no function ρ was introduced, but what we have assumed is that the four numbers A , A' , B , and B' are all attached to the same pair. It then makes sense to obtain the ensemble average $\langle M \rangle$ from successive measurements of four average values $\langle AB \rangle$, $\langle AB' \rangle$, etc. But, if M is built from more numbers, such as numbers associated to different pairs, the algebra no longer holds¹⁸, and the rest of the proof of the inequality collapses.

Experimentally, it is not especially difficult to start with an ensemble of pairs that is independent of the settings, but the difficulty is to ensure that all pairs are

¹⁸ The pairs that are detected on both sides with orientations a and b of the analyzers belong to a first sub-ensemble, those detected with orientations a and b' to another sub-ensemble. If the choice of the first orientation is a , the result observed locally can then be written either A_b , if the pairs belong to the first sub-ensemble, or $A_{b'}$ if it belongs to the second. The number M is then formed from eight different numbers ± 1 instead of four, and the proof of the Bell limit is not possible.

detected and provide two results ± 1 (independently of the fact that the choice of a and b is made before or after the emission of the pair). If some pairs are undetected, one cannot be completely sure that the detection process has not selected an a and b dependent sub-ensemble of pairs. The statistics of counted pairs can then differ from the statistics of emitted pairs; rigorously speaking, there is no reason why the Bell limit should still be a consequence of local realism – see for instance the example given by Pearle in [161] and Appendix E.

An ideal situation would be provided by a device with a triggering button that could be used by an experimentalist, who could at will launch a pair of particles (with certainty); if the pair in question was always analyzed and detected with 100% efficiency, the loophole would be closed. But one could also obtain the same result without capturing all pairs, but by redefining the ensemble in relation to detection; the only condition is to make sure that the sample of counted events is independent of their settings a and b (unbiased sample). When discussing thought experiments, Bell introduced in some of his talks the notion of “preliminary detectors” [163], devices which he sketched as cylinders through which any pair of particles would have to propagate before reaching both ends of the experiment (where the a and b dependent measurement apparatuses sit); the preliminary detectors should signal the presence of pairs that, later, would always be detected at both ends, whatever choice of a and b is made. The role of these cylinders was therefore to make the definition of the sample more precise, even if initially the pairs were emitted by the source in all directions. Such systems, which allow a definition of an ensemble that is indeed totally independent of a and b , are sometimes called “event-ready detectors”. See also reference [164] where Bell imagines a combination of veto and go detectors associated with the first detected particles in a ternary emission, precisely for the purpose of better sample definition.

It is also possible to take the opposite point of view and to deliberately choose, as we saw in §4.4.3, to bias the sample of detected pairs with a post-selection process. One can then, not only violate Bell inequalities without non-locality (Appendix E, §E.2), but also exceed the Cirelson bound by a large factor [153–155]. This illustrates the importance of the fair sampling condition for obtaining clear experimental violations of local realism.

Imperfections of the experiments Needless to say, in practice, the experimental situation is never ideal! First, one should realize that, in all experiments performed until now, most pairs are simply missed by the detectors. There are several reasons for this situation: in photon experiments, the particles are emitted in all directions, while the analyzers collect only a small solid angle and, therefore, only a

tiny fraction of the pairs. This was especially true in the initial experiments using photon cascades; in more modern experiments [122], the use of parametric photon conversion processes introduces a strong correlation between the direction of propagation of the photons and a much better collection efficiency, but it still remains low. Moreover, the transmission of the analyzers is less than 1 (it is actually less than $1/2$ if ordinary photon polarization filters are used, but experiments have also been performed with birefringent two-channel analyzers [119], which are not limited to 50% efficiency). Finally, the quantum efficiency of particle detectors (photomultipliers for photons) is not 100% either, so that pairs of particles are lost at this stage too. The net result is that there is no independent way to determine the sample of detected pairs, since the detection process itself is obviously a and b dependent; as a consequence, all experimental results become useful only if they are interpreted within a “no-biasing” assumption, or “fair sampling hypothesis”, which amounts to considering that the settings of the analyzers does not bias the statistics of events.

On the other hand, we should also keep in mind that there is no known reason whatsoever why such a sample biasing should take place, so that the possibility remains speculative. The experimentalists are not in the dark concerning the detection efficiencies of their experimental setups. On the contrary, they can calculate them with accuracy, knowing the geometrical characteristics of the apparatuses, the quantum efficiency of the detectors, the rate of excitation of the source, etc.; all these parameters can be carefully controlled and allow them to check that the coincidence rates are exactly what one expects from theory. The rates are also compared with those of single-particle detection under the same experimental conditions; experimenters are not restricted to measurements of the relative variations of the coincidence rates as functions of a and b . The same care has been taken to make all possible checks, as in other important experiments in physics; the result of all this work is that everything fits very well with the predictions of quantum mechanics. It would be rather extraordinary if some mysterious physical effect existed that depended so crucially on the detection efficiency; for low efficiencies, it would mimic the results of quantum mechanics and fool the physicists, who would then erroneously think that the results of quantum mechanics are valid; but, for higher (not yet obtained) efficiencies, it would act completely differently and stop to reproduce quantum mechanics. This is probably why most physicists remain sceptical about this scenario and believe that, when more perfect experiments become possible, the same agreement with standard quantum mechanics will be obtained.

For a theoretical discussion of the necessary conditions to design a loophole-free experiment in terms of background level and counter efficiencies, see [165].

For proposals of really loophole-free experiments¹⁹, see [166, 167] and [168]. The latter reference proposes to use continuous variables that are quantized artificially afterwards (result +1 if the integral of the signal during a pulse is positive, −1 if it is negative). Especially interesting is an idea of “photon subtraction” that does two things at the same time: providing an event-ready detector (in order to close the pair selection loophole) and introducing the necessary quantum state to violate Bell’s inequalities. In this case, the predicted violation is only 1%, but there are more elaborate situations where much higher violations are expected [169].

In any case, there seems to be now a reasonable hope that this loophole will be closed by the experiments within a reasonable delay. An experiment with $^9\text{Be}^+$ ions was performed with high detection efficiency [170], so that the corresponding loophole was closed; nevertheless, the distance between the two ions was small (a few microns) so that the communication loophole (see below) could not simultaneously be closed. It has nevertheless been proposed to use the method of “entanglement swapping” (§6.3.2) where pairs of photons entangle remote ions in different traps [171]; the experiment was performed [172, 173] with $^{171}\text{Yb}^+$ ions at a distance of 1m. from each other, and led to a clear violation of the Bell inequalities (3 standard deviations). Recent experiments with solid-state Josephson junctions have given results violating the BCHSH inequalities by more than 200 standard deviations, with quantum systems that are macroscopic electrical circuits [174]. In this case, the problem of capturing pairs of particles does not occur. Nevertheless, since the distance between the circuits is small, the communication loophole remains open; no completely loophole-free experiment has been performed yet.

4.5.1.b Conspiracy (or communication) loophole

Other loopholes are also possible: even if experiments were done with 100% efficiency, one could also invoke some possibilities for local processes to artificially reproduce quantum mechanics. One of them is usually called the “conspiracy of

¹⁹ A perfect correlation between the existence of a detection event on each side (independently of the obtained results) would provide another possible scheme for a loophole free experiment – this, of course, would imply that two channel detectors with a 100% efficiency are used on both ends of the experiment. In itself, the fact that any click at one side is always correlated with a click at the other, independently of the settings a and b , is not sufficient to exclude a setting dependence of the ensemble of detected pairs. But, if one assumes locality at this stage also, a simple reasoning shows that a perfect detection correlation is sufficient to ensure the independence: if some selection occurs for the first particle, it is determined by the detection of the second particle with setting b , independently of the value chosen for a ; it is therefore independent of a . Symetrically, the detection of the second particle is independent of b . Therefore the detection of the pair is independent of a and b .

In other words, locality arguments may be used, not only for the results of the apparatuses (the functions A and B), but also in order to specify the ensemble of observed pairs (the distribution function ρ). Therefore, the observation (in some future experiment) of a violation of the Bell inequalities with a perfect detection correlation would be sufficient to exclude local theories, and therefore to close the loophole.

the polarizers”²⁰ (actually, “conspiracy of the analyzers” would be more appropriate), or also “communication loophole”. Assume that, by some unknown process, each analyzer becomes sensitive to the orientation of the other analyzer; it can then acquire a response function that depends on the other setting, and the function A may have a dependence on both a and b . Under these conditions, the proof of the Bell theorem is no longer possible. The same is true if the choice of a may somehow influence the propagation of the second particle between the source and the apparatus with setting b . Unknown physical processes could therefore conspire to reproduce violations of the inequalities without violating local realism, giving the false impression that the latter is not obeyed.

A way to exclude the possibility of these processes is to make use of relativistic causality. If the distances between the two analyzers and between the analyzers and the source are sufficiently large, and if the settings a and b are chosen at the very last moment, then no physical influence can propagate (at the maximum speed of light) between the different parts of the apparatus; unless of course they violate relativity, these unknown processes become impossible. The loophole is therefore closed. A first step in this direction was done by Aspect *et al.* in 1982 [120]. More recent experiments [175] with very fast changes of a and b have succeeded in excluding this possibility and completely closing this loophole. Quantum mechanics still works perfectly well under these more stringent time-dependent conditions.

4.5.1.c Fatalism versus free will

Along a similar line is what is sometimes called the “fatalistic loophole” (or also “superdeterminism”). The idea is to put into question an implicit assumption of the reasoning that leads to the Bell theorem: the completely arbitrary choice by the experimenters of the settings a and b . Usually, a and b are indeed considered as free variables: their values are not the consequence of any preliminary event that took place in the past, but those of free human choice, which may be made just before the measurement. On the other hand, it is true that there is always an overlap between the past cones of two events (here the choice of the settings). It is therefore always possible in theory to assume that they have a common cause; a and b are then no longer free parameters, but variables that can fluctuate (in particular, if this cause itself fluctuates) with all kinds of correlations. In this case, it is easy to see that the proof of the Bell theorem is no longer possible²¹.

²⁰ The word polarizer refers to the experiments performed with photons, where the spin orientation of the particles is measured with polarizing filters; but there is nothing specific on photons in the scenario, which can easily be transposed to massive spin 1/2 particles.

²¹ For instance, in the proof that makes use of a probability density $\rho(\lambda)$, if one assumes that a and b become two functions $a(\lambda)$ and $b(\lambda)$, it makes no sense to compare the average values for different fixed values of a and b .

What is then denied is the notion of free will of the experimenters, whose decisions are actually supposed always to be predetermined, even if they are not aware of it; the notion of arbitrary external parameters, which usually define the experimental conditions, no longer makes sense in this scheme. This price being paid, one can build a theory that remains at the same time realist, local, and (super)deterministic, includes a sort of physical theory of human decision, and can violate the Bell inequalities; see for instance [176]. This remains, of course, an unusual point of view, since the notion of arbitrary external parameters is generally accepted in physics; in the words of Bell [177]: “A respectable class of theories, including quantum theory as it is practiced, have free external variables in addition to those internal to and conditioned by the theory ... They are invoked to represent the experimental conditions. They also provide a point of leverage for free willed experimenters ...”.

In practice, when many values of the settings are chosen randomly in an experiment, they are not decided by a human being, but are automatically created by a random number generator. For instance, in the time-dependent experiment mentioned above [120], a double random generator was used to pilot the values of the settings at both ends of the experiment. Transposing the free will loophole to this situation leads to the question: are the number generators actually not random? Do they actually provide values that are actually consequences of the fluctuations of some common cause in the past? If so, a and b are functions of some variable λ , and the Bell theorem no longer applies. One could also imagine influences of the random number generator onto the source of particles, which would be emitted in a state that is correlated with the type of measurement they will undergo later; this also would block the proof of the theorem. It is then not totally impossible that λ fluctuates and influences a and b , and/or the source, in a way that reproduces the predictions of quantum mechanics, but from a totally different mechanism that may be local and realist. A method to reduce the plausibility of this possibility is to use very remote measurement apparatuses with independent random number generators, as was done in a recent experiment [124].

A fatalistic view of the physical universe leaves little room for experimental scientific disciplines. If the history of the universe is predetermined once and for all, including all experiments that have and will be performed within this universe, it is not possible to “ask questions to Nature” in various ways, since the questions themselves are already consequences of this unique history. The very notion of the scientific experimental method becomes somewhat dimmed, if it keeps any meaning at all. Needless to say, the fatalist attitude is not very common among scientists; many consider that eliminating the possibility of free observers in this way is too high a price to explain the difficulties of quantum mechanics.

It is nevertheless possible to take an intermediate point of view, where experimenters retain some free will, but not complete. It is shown in [178] that, if the experimenters just give up 14% of their decision independence, within local realism it becomes possible to reproduce the predictions of quantum mechanics for spin measurements with a singlet spin state.

4.5.1.d Credibility of loopholes

We should not conclude this section on loopholes without mentioning that, if they undoubtedly exist, they keep a somewhat “*ad hoc*” character. The explanations in question do not rest on any specific theory: no-one has precise ideas about the nature of the physical processes involved in the conspiracy, or of how pair selection would occur in a way that is sufficiently complex to perfectly mimic quantum mechanics. The only thing one can do (see Appendix E) is to build *ad hoc* models to reproduce quantum mechanics. But by what kind of mysterious process would experiments mimic quantum mechanics so perfectly at low collection efficiencies, and cease to do so at some higher value of efficiency? Bell himself was probably the one who should have most liked to see that his inequalities could indeed be used as a logical tool to find the limits of quantum mechanics; nevertheless, he found these explanations too unaesthetic to be really plausible. But in any case logic remains logic; yes, there still remains a slight possibility that, when the experiments reach a level of efficiency in pair collection where the loophole becomes closed, the results concerning the correlation rates will progressively deviate from those of quantum mechanics to reach values compatible with local realism. Who knows?

4.5.2 Is quantum mechanics itself non-local? Counterfactuality

One can find in the literature various attitudes concerning the exact relation between quantum mechanics and locality and realism. Some authors consider that the non-local character of quantum mechanics is a fact. For instance, in [179], Bell writes “Quantum mechanics is non locally causal” and adds “quantum mechanics cannot be embedded in a local causal theory”. As for Popescu and Rohrlich, as mentioned in §4.4.4, they even propose to consider non-locality as one of the axioms of quantum mechanics [151]. For others, quantum non-locality is just an artefact created by the introduction into quantum mechanics of notions that are foreign to it (typically the EPR elements of reality). They add that state vector reduction is not an essential component of quantum mechanics (see for instance §§1.2.2.b and 10.1.2.b) and that this theory never introduces any violation of relativistic causality. Their conclusion is then, in general, that quantum mechanics is contradictory with realism, but not locality.

Consider for instance the correlations of the results observed by Alice and Bob in distant galaxies. The proponents of a local non-realist theory remark that such distant correlations are unobservable directly; to be measured, they require that Alice and Bob should exchange messages, or travel to some common place, and then compare their results. Since neither they nor their messages can travel faster than the velocity of light, the observation of the correlations is indeed subject to the minimum time prescribed by relativity. Indeed, a perfectly relativistic space-time diagram accounting for the whole experiment can be drawn, from the emission of the entangled particles to the observation of the correlations. The proponents of non-locality then reply that a basic component of the scientific method is to trust the observation made by other scientists. Since Alice and Bob report that they observed the results long before they met at the end of the experiment, the correlation already existed when they made their measurement, and was therefore necessarily non-local. But the first group then remarks that this reasoning implies a definition of the reality in the past, a delicate concept that is not at all indispensable if realism is abandoned, etc. Lively discussions to decide whether or not quantum mechanics in itself is inherently non-local have taken place and are still active [180–182]; see also references [24] and [99, 183]. Delicate problems of logic are involved and we will not discuss the question in more detail here.

What is easier to grasp for the majority of physicists is the notion of “counterfactual statement” or, more generally, “counterfactuality” [184]. A counterfactual reasoning considers the results of possible experiments that can be envisaged for the future as well-defined quantities, and as valid mathematical functions to use in equations, even if they are still unknown – of course, in algebra one writes unknown quantities in equations all the time. This is very natural: as remarked by d’Espagnat [185, 186] and by Griffiths [187], “counterfactuals seem a necessary part of any realist version of quantum theory in which properties of microscopic systems are not created by the measurements”. One can also see the EPR criterion of reality as a statement of the existence of counterfactuals.

But it also remains true that, in practice, it is never possible to realize more than one of the four experiments that are necessary to obtain a violation of the BCHSH inequalities: for a given pair, one has to choose a single orientation of the analyzers for the measurement, so that all other orientations will remain forever in the domain of speculations. For instance, in the reasoning of §4.1.2, at least two of the numbers A , A' , B , and B' are counterfactuals. One could then conclude that counterfactuality is the notion to reject from quantum mechanics. We have already quoted a sentence by Peres [108], who wonderfully summarizes the situation in orthodox theory: “unperformed experiments have no results”; as Bell once regretfully remarked [177]: “it is a great inconvenience that the real world is given to us once only”!

But, after all, one can also accept counterfactuality and take a point of view that is still perfectly consistent with quantum mechanics, provided one then also accepts a more explicit appearance of non-locality. The Bell theorem is not an impossibility theorem, either for counterfactuality, or for hidden variables – a sort of (correct) version of the Von Neumann theorem. Accepting explicit non-locality is after all natural: why require that theories with counterfactuality/additional variables should be explicitly local at all stages, while this is not required from standard quantum mechanics? Indeed, in this theory, neither the state vector itself, nor the state vector reduction postulate – or the calculation of correlation of experimental results in the correlation point of view (§10.1.2) – corresponds to mathematically local calculations. As we have seen in §4.4.2.b, in terms of logical boxes, quantum mechanics does not lead to local stochastic boxes. In other words, even if one can discuss at a fundamental level whether or not quantum mechanics is local or not, it is perfectly clear that its formalism is not; it would therefore be just unfair to request a local formalism from a non-orthodox theory – especially when the theory in question is built in order to reproduce all results of quantum mechanics! As an illustration of this point, we quote Goldstein [18]: “in recent years it has been common to find physicists . . . failing to appreciate that what Bell demonstrated with his theorem was not the impossibility of Bohmian mechanics, but rather a more radical implication – namely non-locality – that is intrinsic to quantum theory itself”.

5

More theorems

The Bell theorem can take the form of several inequalities, as we have seen in §4.2. Moreover, since it was discovered, the theorem has stimulated the discovery of several other mathematical contradictions between the predictions of quantum mechanics and those of local realism. We review a few of them in this chapter: GHZ contradictions (§5.1) and their generalization (§5.2), Cabello’s inequality (§5.3), and Hardy’s impossibilities (§5.4). Finally, in §5.5, we discuss the notion of contextuality and introduce the BKS theorem.

5.1 GHZ contradiction

For many years, everyone thought that Bell had basically exhausted the subject by considering all really interesting situations, and that two-spin systems provided the most spectacular quantum violations of local realism. It therefore came as a surprise to many when in 1989 Greenberger, Horne, and Zeilinger (GHZ) showed that systems containing more than two correlated particles may actually exhibit even more dramatic violations of local realism [188, 189]. They involve a sign contradiction (100% violation) for perfect correlations, while the BCHSH inequalities are violated by about 40% (Cirelson bound) and deal with situations where the results of measurements are not completely correlated. In this section, we discuss three-particle systems, but generalizations to N particles are possible (§5.2).

5.1.1 Derivation

GHZ contradictions may occur in various systems, not necessarily involving spins. Initially, they were introduced in the context of entanglement swapping (§6.3.2) for four particles [188] or entanglement of three spinless particles [189]. Here, following Mermin [190], we will consider a system of three $1/2$ spins, since this simple example is sufficient to discuss the essence of the ideas. We assume that the

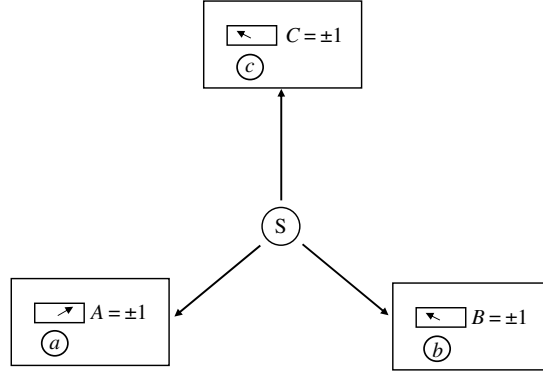


Figure 5.1. Scheme of a GHZ experiment, where three spins in state (5.1) undergo measurements in three different regions of space, and provide results $A = \pm 1$, $B = \pm 1$ and $C = \pm 1$.

three spins are described by the quantum state:

$$|\Psi\rangle = \frac{1}{\sqrt{2}} \left[|+, +, +\rangle + \eta |-, -, -\rangle \right] \quad (5.1)$$

where the $|\pm\rangle$ states are the eigenstates of the components of the spins along the Oz axis of an orthonormal frame $Oxyz$; in the three-particle kets, the first sign refers to the state of the first spin, the second to the state of the second spin, and, similarly, for the third spin; the number η is either $+1$ or -1 :

$$\eta = \pm 1 \quad (5.2)$$

We now calculate the quantum probabilities of measurements of the spins $\sigma_{1,2,3}$ of the three particles, either along direction Ox , or along a perpendicular direction Oy (Figure 5.1). We first consider a measurement of the product $\sigma_{1y} \times \sigma_{2y} \times \sigma_{3x}$; a straightforward calculation (made explicit in §5.2 in a more general case¹) then shows that $|\Psi\rangle$ is an eigenvector of this product with eigenvalue $-\eta$, so that $-\eta$ is the only possible result with state (5.1). Therefore the corresponding probability is:

$$\mathcal{P}(\sigma_{1y} \times \sigma_{2y} \times \sigma_{3x} \Longrightarrow -\eta) = 1 \quad (5.3)$$

while the probability $\mathcal{P}(\sigma_{1y} \times \sigma_{2y} \times \sigma_{3x} \Longrightarrow +\eta)$ of the other possible result vanishes. Similarly, we find that $|\Psi\rangle$ is an eigenvector of the two products

¹ With the notation of that section, here we have $\eta = e^{i\xi}$, $\varphi_1 = \varphi_2 = \pi/2$, and $\varphi_3 = 0$, so that $\zeta = e^{i(\xi - \varphi_1 - \varphi_2 - \varphi_3)} = -\eta$, which provides the eigenvalue. Similarly, for a measurement of the product $\sigma_{1x} \times \sigma_{2x} \times \sigma_{3x}$, we have $\zeta = e^{i\xi} = \eta$, and the eigenvalue is $+\eta$.

$\sigma_{1x} \times \sigma_{2y} \times \sigma_{3y}$ and $\sigma_{1y} \times \sigma_{2x} \times \sigma_{3y}$, with eigenvalues $-\eta$, so that the corresponding probabilities are:

$$\begin{aligned}\mathcal{P}(\sigma_{1x} \times \sigma_{2y} \times \sigma_{3y} \implies -\eta) &= 1 \\ \mathcal{P}(\sigma_{1y} \times \sigma_{2x} \times \sigma_{3y} \implies -\eta) &= 1\end{aligned}\tag{5.4}$$

All three products therefore take the value $-\eta$; the results are known with certainty before the measurement². Now, if we consider the product of three spin components along Ox , it is also easy to check (§5.2) that the same state vector is also an eigenstate of the product operator $\sigma_{1x} \times \sigma_{2x} \times \sigma_{3x}$, but now with eigenvalue $+\eta$, so that:

$$\mathcal{P}(\sigma_{1x} \times \sigma_{2x} \times \sigma_{3x} \implies +\eta) = 1\tag{5.5}$$

This time the result takes the value $+\eta$ with certainty.

Let us now investigate the predictions of an EPR local realist point of view in this kind of situation. Since the quantum calculation is so straightforward when the initial state is an eigenstate of all observables considered (all the results are perfectly certain), one could expect that nothing special will be found. But, actually, we will see that a complete contradiction emerges from this analysis! The local realist reasoning is a direct generalization of that given in §4.1.2. First, the perfect correlations imply that the result of measuring the component along Ox (or Oy) of the spin of any particle can be inferred from the results of measurements made on other particles, at arbitrarily large distances. Therefore the local realist EPR reasoning shows the existence of elements of reality attached to these two components, which we note $A_{x,y} = \pm 1$. This number is the result that the first spin will give for a measurement along Ox or Oy , independently of the type of measurement performed for the two other spins; similar letters B and C are used for the measurement on the two other spins. To reproduce the three equalities written in (5.3) and (5.4), we need to have:

$$\begin{aligned}A_y B_y C_x &= -\eta \\ A_x B_y C_y &= -\eta \\ A_y B_x C_y &= -\eta\end{aligned}\tag{5.6}$$

Now locality implies that the same values of A , B , and C can be used again for the experiment where the three Ox components are measured: the result is merely the product $A_x B_x C_x$. But, since the squares A_y^2 etc. are always equal to $+1$, we can

² The product is fixed, but each of the individual components still fluctuates between results $+1$ and -1 .

obtain this result by multiplying all three lines of equation (5.6), which provides:

$$A_x B_x C_x = -\eta \quad (5.7)$$

But equality (5.5) predicts that the measurement of $\sigma_{1x} \times \sigma_{2x} \times \sigma_{3x}$ will always give the result $+\eta$, which has the opposite sign! The contradiction between the predictions of local realism and those of quantum mechanics cannot be more pronounced.

5.1.2 Discussion

The GHZ contradiction looks even more dramatic than for the Bell inequalities since the quantum and local realist predictions do not differ by some significant fraction (about 40%), but are completely opposite. In a thought experiment, all fluctuations are eliminated since all of the results (the products of the three components) are perfectly known before measurement: the 100% contradiction is obtained with 100% certainty! Apart from this, what differentiates a GHZ situation from a usual Bell experiment with two spins? Different points of view are possible.

(i) One point of view assumes that the three spins are measured individually in each realization of the experiment. The three spins may be in different regions of space; when the spatial variables are included, (5.1) symbolizes a ket that can be written more explicitly as:

$$|\Psi\rangle = \frac{1}{\sqrt{2}} |1 : \varphi_a\rangle |2 : \varphi_b\rangle |3 : \varphi_c\rangle \otimes \left[|1 : +; 2 : +; 3 : +\rangle + \eta |1 : -; 2 : -; 3 : -\rangle \right] \quad (5.8)$$

where $|\varphi_{a,b,c}\rangle$ are three orbital states with non-overlapping wave functions. These functions can be entirely localized in separate boxes, where the spin measurements are performed, so that no particle is missed and each of them is addressed separately. The way to proceed is that, after choosing a component, O_x or O_y for each spin, one performs the three corresponding measurements, obtains three results $A_{x,y}$, $B_{x,y}$, and $C_{x,y}$, and then calculates their product. Averaging over many realizations of the experiment provides the average value $\langle A_{x,y} B_{x,y} C_{x,y} \rangle$. With this procedure, one first measures $\langle A_y B_y C_x \rangle$, $\langle A_x B_y C_y \rangle$, and $\langle A_y B_x C_y \rangle$ to check the perfect correlations predicted by quantum mechanics, from which the EPR reasoning leads to the existence of the six separate elements of reality. Then one measures $\langle A_x B_x C_x \rangle$ and, if quantum mechanics still gives correct predictions, obtains the opposite sign; the conclusion is then that local realism is violated. Equivalently, one can conclude that the value obtained by measuring, say, σ_{1x} depends on whether the spin component measured for the other particles is along O_x or O_y , even if the corresponding

operators commute with σ_{1x} . In other words, one arrives at the notion of “quantum contextuality”, which we will discuss in more detail in §5.5.

(ii) A different point of view is to consider that some experimental procedure has been found to measure directly the products of operators, without obtaining information on the three separate factors inside each product. All four product operators then commute with each other, which introduces an important conceptual difference with violations of Bell type inequalities, where the non-commutation is essential. Here, at least in principle, it is not impossible to measure all of them with a single setup; Bohr could not have invoked the incompatibility of experimental arrangements. Under these conditions, where is the contradiction between the local realist reasoning and quantum mechanics? While, within local realism, a measurement of the product of three operators is equivalent to a separate measurement of each of them, in quantum mechanics this is not the case. Within quantum mechanics, it is possible (in principle) to conceive an apparatus that measures all four products, but impossible to design a single experimental setup to have access to all six factors $A_{x,y}$, $B_{x,y}$, and $C_{x,y}$ (since, for instance, A_x and A_y correspond to the results of two incompatible measurements). What quantum mechanics then violates is the “product rule” discussed in §5.5.2.

An ideal GHZ experiment would therefore involve only measurements of commuting products, more precisely the simultaneous measurement of four products (without measuring separately each factor in the products). The beauty of this theoretical possibility is that loopholes such as the “biased sample loophole” (§4.5.1) would then automatically be closed³. On the other hand, measuring a product without measuring each factor may raise a real experimental challenge.

Even with separate measurements, experimental tests of the GHZ equality still require that three particles should be put in state (5.1), surely a non-trivial task. Nevertheless, with the help of elaborate techniques of quantum electronics, GHZ entanglement has been observed with three photons [191], leading to a successful experimental test of the GHZ non-locality [192]. Similar results have also been obtained with four entangled photons [193]. The techniques of NMR (nuclear magnetic resonance) have also been used to test the GHZ equality, at least with microscopic distances [194] – see also [195, 196]. All these experiments have confirmed the predictions of quantum mechanics.

When the measurements are imperfect, the GHZ equality can give rise to inequalities (as for the BCHSH theorem) see [197] (this reference also presents a generalization to an arbitrary number N of particles) and [198]. The use of particles originating from three independent sources with quantum interference effects that disprove local realism has been discussed in [199], and later generalized to an

³ One could not assume that different experimental setups select different groups of particles, since only one is used.

arbitrary large number of particles [140]. The transposition to continuous position and momentum variables is discussed in [200]. Two-particle interferometers can also lead to GHZ contradictions for spinless particles [201].

5.2 Generalizing GHZ (all or nothing states)

We now generalize the GHZ state (5.1) by introducing an N -particle state $|\Psi\rangle$ defined as:

$$|\Psi\rangle = \frac{1}{\sqrt{2}} \left[|+, +, \dots, +\rangle + e^{i\xi} |-, -, \dots, -\rangle \right] \quad (5.9)$$

where ξ is a real parameter defining the relative phase of the two components; $|\Psi\rangle$ is a coherent superposition of two N -particle states where all, or none, of the spins are in the individual state⁴ $|+\rangle$, hence the name “all-or-nothing state” that we use here⁵ – but “many-particle GHZ states” would be equally appropriate. One also sometimes finds the name “maximally entangled states”, or⁶ “NOON states”, in the literature.

The state we consider here is very special, since it is a superposition of two N -particle components where all particles are in different, orthogonal, individual states. This state should not be confused with a different state where many particles have each been prepared in a coherent superposition. For instance, one can prepare individually N spin 1/2 by sending them through a spin filter (polarization-dependent beam splitter, Stern–Gerlach analyzer, etc.), which prepares a coherent superposition $\alpha |+\rangle + \beta |-\rangle$ for each spin. The effect of the filter on the group of N particles is to put them into a state that is a product of single-particle coherent states, namely:

$$\begin{aligned} |\Psi'\rangle = & \left[\alpha |1 : +\rangle + \beta |1 : -\rangle \right] \otimes \left[\alpha |2 : +\rangle + \beta |2 : -\rangle \right] \\ & \otimes \dots \otimes \left[\alpha |N : +\rangle + \beta |N : -\rangle \right] \end{aligned} \quad (5.10)$$

This state contains components of the state vector where some of the spins are up, some down, in various proportions. In (5.10), each particle is in the same spin state, and the situation is somewhat analogous to a Bose–Einstein condensate where all

⁴ An all-or-nothing state does not necessarily require that all spins be up in the first component of the state vector, down in the second; what actually matters is that every spin changes to an orthogonal spin state when one goes from one N -particle component of $|\Psi\rangle$ to the other (in other words, the quantization axis may vary from spin to spin)

⁵ Strictly speaking, we should use the name “all and nothing state”, since both possibilities are simultaneously present in $|\Psi\rangle$.

⁶ These states can be written as a superposition of kets with occupation numbers $N, 0; 0, N$, hence the name.

particles are in the same coherent state – for instance a condensate located on both sides of a potential barrier coupled by tunnel effect, as in the Josephson effect. By contrast, in (5.9) the coherence is essentially an N -body coherence only, so that the entanglement is more subtle.

Consider now the component of spin j along a direction in plane xOy making angle φ_j with Ox ; it corresponds to the operator:

$$\sigma_j(\varphi_j) = \frac{1}{2} \left[e^{-i\varphi_j} \sigma_+(j) + e^{i\varphi_j} \sigma_-(j) \right] \quad (5.11)$$

where $\sigma_{\pm}$ denotes the operator $\sigma_x \pm i\sigma_y$ as usual, with the following action:

$$\sigma_{\pm} |\pm\rangle = 0 \quad \text{and} \quad \sigma_{\pm} |\mp\rangle = 2 |\pm\rangle \quad (5.12)$$

We now introduce the N -particle operator corresponding to the product of various components of the spins:

$$Q(\varphi_1, \varphi_2, \dots, \varphi_N) = \sigma_1(\varphi_1) \sigma_2(\varphi_2) \dots \sigma_N(\varphi_N) \quad (5.13)$$

which is equal to:

$$Q(\varphi_1, \varphi_2, \dots, \varphi_N) = \prod_{j=1}^N \frac{1}{2} \left[e^{-i\varphi_j} \sigma_+(j) + e^{i\varphi_j} \sigma_-(j) \right] \quad (5.14)$$

The action of this operator onto state (5.9) is easily obtained by using (5.12). If one takes the first component of $|\Psi\rangle$, the only way to obtain non-zero by application of operator (5.14) is to select the term in $e^{i\varphi_j} \sigma_-(j)$ in each bracket of the product over j ; therefore:

$$Q(\varphi_1, \varphi_2, \dots, \varphi_N) |+, +, \dots, +\rangle = e^{i(\varphi_1 + \varphi_2 + \dots + \varphi_N)} |-, -, \dots, -\rangle \quad (5.15)$$

In the same way:

$$Q(\varphi_1, \varphi_2, \dots, \varphi_N) |-, -, \dots, -\rangle = e^{-i(\varphi_1 + \varphi_2 + \dots + \varphi_N)} |+, +, \dots, +\rangle \quad (5.16)$$

Therefore, if we set:

$$\zeta = e^{i(\xi - \varphi_1 - \varphi_2 - \dots - \varphi_N)} \quad (5.17)$$

we obtain:

$$Q(\varphi_1, \varphi_2, \dots, \varphi_N) |\Psi\rangle = \frac{1}{\sqrt{2}} \left[\zeta |+, +, \dots, +\rangle + \frac{1}{\zeta} e^{i\xi} |-, -, \dots, -\rangle \right] \quad (5.18)$$

We then see that, if the condition:

$$\zeta = \frac{1}{\zeta} = \pm 1 \quad (5.19)$$

is fulfilled, the action of Q on $|\Psi\rangle$ reconstructs exactly $\zeta |\Psi\rangle$, which means that $|\Psi\rangle$ is an eigenstate of Q with eigenvalue ζ . In other words, provided the sum of all angles φ_j is equal to ξ (plus some integer multiple of π), the product of all operators $\sigma_j(\varphi_j)$ corresponds to a measurement that gives a result $\zeta = \pm 1$ with certainty.

We can calculate the quantum average of Q from (5.18) and (5.17); for any value of the angles, we obtain:

$$\langle Q \rangle = \cos(\xi - \varphi_1 - \varphi_2 - \dots - \varphi_N) \quad (5.20)$$

For instance, if all angles φ_j are equal to the same value φ , this formula predicts that the average value of Q oscillates rapidly as a function of φ if N is large. Whatever the value of N is, it turns out that it is completely impossible to reproduce the oscillations contained in (5.20) within local realism [202]. In the case $N = 2$, this of course reduces to the usual Bell theorem. But, as soon as N becomes 3 or takes a larger value, the contradiction becomes even more dramatic. In [202], it is assumed that a local probabilistic theory reproduces (5.20) only for some sets of particular value of the angles φ (those for which the result is certain). One can then show that the theory in question necessarily predicts that $\langle Q \rangle$ is independent of all angles φ : the average then keeps a perfectly constant value $+1$! Indeed, the very existence of the oscillation predicted by (5.20) is a purely quantum non-local effect.

As we have already remarked, the effect is essentially a coherent N body effect: if a single spin is missed in the measurement, in (5.14) the number of spin operators is no longer sufficient to transform the ket $|+, +, \dots, +\rangle$ into $|-, -, \dots, -\rangle$ as in (5.15). The result then becomes completely independent of the relative phase $e^{i\xi}$ of the two components; no quantum coherence effect then takes place. Actually, it is easy to check that the average value of any product of $N - 1$, $N - 2$, etc. components of the spins vanishes; the quantum interference effect leading to (5.20) takes place only if all N spins are measured.

This is not the only remarkable property of all-or-nothing states. For instance, it can be shown that, for large N , they lead to exponential violations of the bounds put by local realist theories [197]. In the context of a possible use of quantum correlations to reduce the quantum noise in spectroscopy [203], it has been pointed out [204] that these states (called “maximally correlated states”) have remarkable properties in terms of spectroscopic measurements: the frequency uncertainty of measurements decreases as $1/N$ for a given measurement time, therefore much faster than the usual $1/\sqrt{N}$ factor obtained with measurements on

independent particles⁷. A discussion of this method, as compared to others making use of Bose–Einstein condensates, is given in [205]. Such quantum states may also be the source of “quantum lithography”, using on a non-linear optics process (many photon absorption process) where the usual diffraction limit $\lambda/2$ obtained in classical physics is divided by N [206–208]. The quantum correlation of these states may turn out to be, one day, the source of improved accuracy in various experiments.

We have already mentioned in §5.1 that entanglement with $N = 3$ was reported in [191, 209] and used to test the GHZ equality [192]. Proposals for methods to generalize to larger values of N with ions in a trap were put forward by Mølmer *et al.* [210]; the idea exploits the motion dependence of resonance frequencies for a system of several ions in the same trap, as well as some partially destructive interference effects. The scheme was successfully put into practice in an experiment [211] where “all-or-nothing states” were created for $N = 2$ as well as $N = 4$ ions in a trap. GHZ states are not the only possibility to obtain violations of local realism with three particles; one can also use three particles originating from independent sources, and send them into an appropriate interference device [199].

5.3 Cabello’s inequality

We now discuss a scheme introduced by Cabello [212], which leads to violations of BCHSH type inequalities exceeding the Cirelson bound (§4.3) – in fact, reaching the maximal value 4 mathematically compatible with the definition of the sum of averages appearing in these inequalities. To obtain this result, we consider a system made of three sub-systems, and we cast the GHZ results into a form in which they appear as components of a BCHSH inequality; the idea is, so to say, to transform three-particle GHZ correlations into two-particle BCHSH correlations.

Again, we consider three spins in state (5.1) with $\eta = +1$ and measurements of the components of the three spins along Ox or Oy . More explicitly, we write the initial quantum state as:

$$|\Psi\rangle = \frac{1}{\sqrt{2}} |\varphi_a(1)\rangle |\varphi_b(2)\rangle |\varphi_c(3)\rangle \otimes \left[|1: +, 2: +, 3: +\rangle + |1: -, 2: -, 3: -\rangle \right] \quad (5.21)$$

⁷ One can give a simple physical interpretation of this sensitivity improvement. Suppose that the two individual spin states $|\pm\rangle$ represent single particles propagating in the two arms of a Mach–Zehnder interferometer. The ket (5.9) then describes a N particle state where all particles propagate together, either in one arm of the interferometer, or in the other. In other words, what propagates in the interferometric device is a cluster of N particles having N times the mass and N times the energy of a single particle. Since the de Broglie wavelength of a quantum object with mass Nm and energy Ne is $\hbar/N\sqrt{em}$, it is N times shorter than for a single particle, which improves the sensitivity of the interferometer by the same factor.

where $|\varphi_{a,b,c}\rangle$ are three orbital states located in three different regions of space A, B, and C, where the separate spin measurements take place (Alice, Bob, and Carol).

Assume for instance that we measure the three Ox components of all spins. We have shown that the product of the results is always $+1$, so that two cases are possible. If two results -1 are obtained, we call i and j the corresponding regions and spins (i being attributed to the first region in the alphabetical order A, B, C), and k the region where result $+1$ has been obtained. If the three results are equal⁸ to $+1$, we choose $i = A$, $j = B$, and $k = C$. We always have:

$$\langle \sigma_x(i) \sigma_x(j) \rangle = 1 \quad (5.22)$$

which is indeed a two-spin relation.

5.3.1 Local realist point of view

To each of the three regions of space we attach three numbers X , Y , and Z , each equal to ± 1 . The EPR local realist theorem and the existence of complete GHZ correlations between remote spins show the existence of nine numbers $X_{a,b,c}$, $Y_{a,b,c}$ and $Z_{a,b,c}$ for regions A, B, and C; these nine numbers give the result that a measurement of the spin component along Ox , Oy , or Oz in each region will provide. The transposition of the results of quantum mechanics implies that, for each realization of the experiment (each triplet of particles), the product $X_a X_b X_c$ is equal to 1; either two of the X are equal to -1 , or none.

This allows us to define three indices i , j , and k for all triplets of spins (not only those for which all Ox components of the spins are measured) as follows:

- if two of the X are equal to -1 , they get indices i and j (among these two regions, i is associated with the region that is first in alphabetical order), and k is for the third region; X_k is then always equal to $+1$.
- if all the X are equal to $+1$, then $i = a$, $j = b$, and $k = c$.

For each realization of the experiment, this defines a perfect correspondence between the regions and the three indices i , j , and k , but of course this correspondence may change from one experiment to the next. Nevertheless, because local realism has introduced the counterfactuality (§4.5.2) on which the numbering process is based, this correspondence is always possible whatever measurements are actually performed, which is of course not the case within quantum mechanics.

⁸ Expanding $|\Psi\rangle$ onto the eigenvectors of the Ox components of the three spins shows that this happens one time out of four.

In the proof of the BCHSH inequality for two spins in §4.1.2, we have considered in (4.7) a quantity of the form:

$$X(Y + mY') + nX'(Y - mY') \quad (5.23)$$

with:

$$m = \pm 1 \quad ; \quad n = \pm 1 \quad (5.24)$$

Expression (5.23) is always equal to ± 2 when the four numbers involved are ± 1 . Here, we introduce the following combination of the nine numbers:

$$X_i(X_j + mY_j) + nY_i(X_j - mY_j) \quad (5.25)$$

which is also always -2 or $+2$; this is because either $(X_j + mY_j)$ or $(X_j - mY_j)$ vanishes, and all numbers have modulus equal to 1. Since $Y_k = \pm 1$, we may choose $m = n = -Y_k$, so that this expression becomes:

$$X_i(X_j - Y_k Y_j) - Y_i Y_k(X_j + Y_k Y_j) \quad (5.26)$$

which provides:

$$X_i X_j - X_i Y_k Y_j - Y_i X_j Y_k - Y_i Y_j = \pm 2 \quad (5.27)$$

But since, by definition of index k , we have $X_k = 1$, equation (5.27) gives:

$$X_i X_j X_k - X_i Y_j Y_k - Y_i X_j Y_k - Y_i Y_j X_k = \pm 2 \quad (5.28)$$

Therefore, within local realism, the product of three measurements along Ox minus the three products of two Oy components with one Ox components is always equal to -2 or $+2$. The average value over many realizations is therefore always comprised between -2 and $+2$:

$$-2 \leq \langle X_i X_j X_k \rangle - \langle X_i Y_j Y_k \rangle - \langle Y_i X_j Y_k \rangle - \langle Y_i Y_j X_k \rangle \leq +2 \quad (5.29)$$

The average value of the product of the three different X , minus the three different combinations of one X and two Y , is always bounded by ± 2 .

5.3.2 Contradiction with quantum mechanics

In quantum mechanics, the average of the product of the three different Ox spin components, minus the three different combinations of one Ox and two Oy components, is:

$$\begin{aligned} Q = & \langle \sigma_x(1)\sigma_x(2)\sigma_x(3) \rangle - \langle \sigma_x(1)\sigma_y(2)\sigma_y(3) \rangle \\ & - \langle \sigma_y(1)\sigma_x(2)\sigma_y(3) \rangle - \langle \sigma_y(1)\sigma_y(2)\sigma_x(3) \rangle \end{aligned} \quad (5.30)$$

But we have seen above (§5.1) that the first term is $+1$, while all the others are -1 . Therefore quantum mechanics predicts $+4$ for this quantity, in strong violation of (5.29) – by a factor 2.

The conclusion is that, by combining elements of the BCHSH reasoning with others of the GHZ reasoning, it is possible with three particles to obtain violation of generalized BCHSH inequalities that saturate the absolute mathematical limit. Using a third particle as a “marker” so to say, allows one to select the appropriate BCHSH inequality for the two other particles and to exceed the Cirelson bound. The reference [213] proposes another scheme, involving pre- and post-selection, which also provides the maximum value 4 to the BCHSH term.

5.4 Hardy's impossibilities

Another scheme illustrating contradictions between local realism and the predictions of quantum mechanics was introduced by Hardy [59]. As the initial Bell theorem, it involves two correlated particles, but it is conceptually different: Hardy's theorem leads to conclusions concerning the very possibility (or impossibility) of occurrence for some type of events – instead of mathematical constraints on correlation rates. A general discussion of this interesting contradiction is given in [214]. As in §4.1.2, we assume that the first particle may undergo two kinds of measurements, characterized by two values a and a' of the first setting.

Within local realism, we call A and A' the corresponding results. Similar measurements can be performed on the second particle, and we call B and B' the results. Let us now consider three types of situations (Figure 5.2):

- (i) if the settings are a, b , the result $A = 1, B = 1$ is sometimes obtained.
- (ii) when “crossed measurements” are performed (either a, b' or a', b), we have certainties: if $A = 1$, the value of B' is always -1 ; if $B = 1$, the value of A'

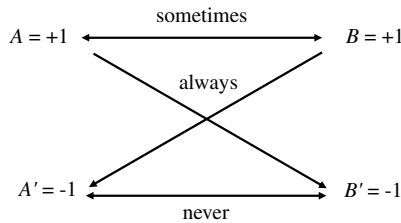


Figure 5.2. Logical scheme associated with the Hardy impossibilities. If results $A = B = 1$ are sometimes obtained, and if B' (resp. A') is always equal to -1 when A (resp. B) is equal to 1, then event $A' = B' = -1$ must sometimes occur. But one can construct a quantum state for which the first two conditions are obeyed while such events never occur, which creates a contradiction with local realism.

is always -1 . One can also express this as an exclusion: “double ones” never occur in crossed measurements, neither $A = 1, B' = 1$ nor $A' = 1, B = 1$ are ever obtained.

- (iii) when both settings are primed, we assume that “double minus one” is impossible: $A' = -1, B' = -1$ is never observed.

It turns out that these three assumptions are in fact incompatible. To see why, let us consider the logical scheme of Figure 5.2, where the upper part corresponds to the possibility opened by statement (i); statement (ii) then implies that, if $A = 1$, one necessarily has $B' = -1$, which explains the first diagonal in the figure; the second diagonal follows by symmetry. Then we see that all events corresponding to the results $A = B = 1$ also necessarily correspond to $A' = B' = -1$, so that a contradiction with statement (iii) appears. Indeed, the “sometimes” of (i) is contradictory with the “never” of the exclusion of proposition (iii).

But it turns out that quantum mechanics does allow a simultaneous realization of all three propositions! To see how, let us consider a two-spin state vector of the form:

$$|\Psi\rangle = \alpha |+, -\rangle + \beta |-, +\rangle + \gamma |+, +\rangle \quad (5.31)$$

where the kets $|\pm, \pm\rangle$ refer to the common eigenstates of the Pauli operator $\sigma_z(1)$ for spin 1 and the spin operator $\sigma_z(2)$ for spin 2, with a given quantization axis Oz . We assume that $A' = \sigma_z(1)$ and $B' = \sigma_z(2)$, so that:

$$|\pm, \pm\rangle = |A' = \pm, B' = \pm 1\rangle \quad (5.32)$$

The absence of any $|\Psi\rangle$ component on $|A' = -1, B' = -1\rangle$ ensures that proposition (iii) is automatically true. As for the measurements without prime, we assume that they are both performed on the spins along a direction in the plane xOz that makes an angle 2θ with Oz ; the single-spin eigenstate with eigenvalue $+1$ is then:

$$\cos\theta |+\rangle + \sin\theta |-\rangle \quad (5.33)$$

Proposition (ii) (diagonals in Figure 5.2) amounts to excluding the two-spin state:

$$\cos\theta |+, +\rangle + \sin\theta |-, +\rangle \quad (5.34)$$

as well as state:

$$\cos\theta |+, +\rangle + \sin\theta |+, -\rangle \quad (5.35)$$

The two exclusion conditions are obtained by cancelling the scalar product of $|\Psi\rangle$ and these two vectors. This is equivalent to the condition:

$$\alpha \sin\theta + \gamma \cos\theta = \beta \sin\theta + \gamma \cos\theta = 0 \quad (5.36)$$

or:

$$\alpha = \beta = -\gamma \cot \theta \quad (5.37)$$

Within an arbitrary coefficient, we can then write $|\Psi\rangle$ in the form:

$$|\Psi\rangle = -\cos \theta \left(|+, -\rangle + |-, +\rangle \right) + \sin \theta |+, +\rangle \quad (5.38)$$

The scalar product of this ket with that where the two spins are in the state (5.33) is:

$$-\sin \theta \cos^2 \theta \quad (5.39)$$

The final step is to divide the square of this result by the square of the norm of ket (5.38) in order to obtain the probability of the process considered in (iii); this is a straightforward calculation (Appendix D), but here we just need to point out that the probability is not zero; the precise value of its maximum with respect to θ is found in Appendix D to be about 9%. This proves that the pair of results considered in proposition (i) can sometimes be obtained together with (ii) and (iii): indeed, in 9% of the cases, the predictions of quantum mechanics are in complete contradiction with those of a local realist reasoning.

An interesting aspect of the above propositions is that they can be generalized to an arbitrary number of measurements [215]; it turns out that this permits a significant increase in the percentage of “impossible events” (impossible within local realism) predicted by quantum mechanics – from 9% to almost 50%! The generalization involves a chain (Figure 5.3), which keeps the two first lines (i) and (ii) unchanged, and iterates the second in a recurrent way by assuming that:

- (iii) for measurements of the type (a', b'') or (a'', b') , one never gets opposite results⁹;
- (iv) similarly, for measurements of the type (a'', b''') or (a''', b'') , one never gets opposite results,
etc ...
- (n) but, for measurement of the type (a^n, b^n) , one never gets -1 and -1 .

The incompatibility proof is very similar to that given above; it is summarized in Figure 5.3.

In both cases, the way to resolve the contradiction is the same as for the Bell theorem: in quantum mechanics, for a given pair of spins, it is not correct to reason on all four quantities A , A' , B , and B' , even considered as quantities that are

⁹ In fact, the reasoning just requires that the pair $-1, +1$ is never obtained, and does not require any statement about $+1, -1$.

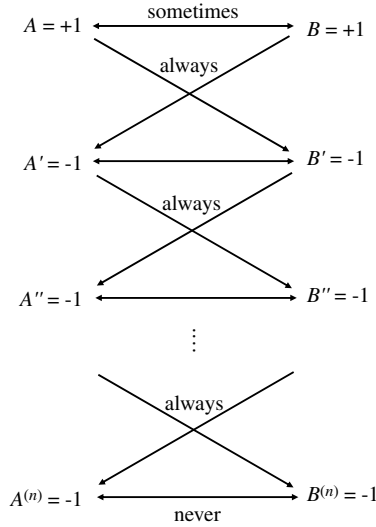


Figure 5.3. Logical scheme for iterating the Hardy impossibilities

unknown and that could be determined in a future experiment. This is simply because, with a given pair, it is obviously impossible to design an experiment that will measure all of them: the measurements are incompatible. For a discussion of non-local effects with other states, see [216].

GHZ contradictions, or Hardy impossibilities, may appear as a stronger contradiction with quantum mechanics than just a violation of statistical BCHSH inequalities. But a closer examination shows that, in fact, both are logically related and that all violations discussed so far involve inequalities at a deeper level [217].

5.5 Bell–Kochen–Specker theorem: contextuality

Another theorem was introduced also by Bell [5] as well as (independently and almost simultaneously) by Kochen and Specker [218], hence the name “BKS theorem” that is often used for it. This theorem is not particularly related to locality, but to another notion called “contextuality”. An additional variable attached to a physical system is called “contextual” if its value depends, not only on the physical quantity that it describes, but also on the other physical quantities that can be measured at the same time on the same system (therefore corresponding, in quantum mechanics, to commuting observables). If, on the other hand, its value is independent of all the other observables that the experimenter may decide to measure at the same time, the additional variable is called “non-contextual”. One can then say that it describes a property of the physical system only, and not a combined property of the system and the measurement apparatus; it may have pre-existed in

the system before any measurement. The notion of distance is no longer relevant in this context; for instance, the theorem applies to a single system with no extension in space.

5.5.1 A spin 1 particle

Let us first consider a spin 1 particle in quantum mechanics, with three quantum states $|-1\rangle$, $|0\rangle$, and $|+1\rangle$ as a basis of a state space of dimension 3. The three components S_x , S_y , and S_z do not commute (they obey the usual commutation relation for the angular momentum), but it is easy to show that all the squares of these three operators do commute; this is a specific property of angular momentum 1, and can be seen by an elementary matrix calculation with the usual operators $S_{\pm}$. Moreover, the sum of these squares is a constant (an operator proportional to the identity operator, sometimes called a “c-number”) since:

$$S_x^2 + S_y^2 + S_z^2 = 2\hbar^2 \quad (5.40)$$

It is not against any fundamental principle of quantum mechanics, therefore, to imagine a triple simultaneous measurement of the observables S_x^2 , S_y^2 , and S_z^2 ; we know that the sum of the three results will always be 2 (from now on we drop the factor $\hbar^2$, which plays no role in the discussion). Needless to say, the choice of the three orthogonal directions is completely arbitrary, and the compatibility is ensured for any choice of this triad, but not more than one: the measurements for different choices of the triad remain incompatible.

In passing, we note that the measurement of the square S_x^2 of one component cannot merely be seen as a measurement of S_x followed by a squaring calculation made afterwards by the experimentalist! Not obtaining some information (the sign) is not equivalent to ignoring the information after it has been obtained by a measurement (we come to this point in more detail, in terms of interferences and decoherence, at the end of §10.1.2). There is indeed less information in S_x^2 than in S_x itself, since the former has only two eigenvalues (1 and 0), while the latter has three (since result -1 is also possible). What is needed to measure S_x^2 is, for instance, a modified Stern–Gerlach system where the components of the wave function corresponding to results ± 1 are not separated, or where they are separated but subsequently grouped together in a way that makes them impossible to distinguish. Generally speaking, in quantum mechanics, measuring the square of an operator is certainly not the same physical process as measuring the operator itself!

Now, suppose that we try to attach to each individual spin an EPR element of reality (additional variable) that corresponds to the result of measurement of S_x^2 ; by symmetry, we will do the same for the two other components, so that each spin now gets three additional variables λ to which we may attribute values that determine

the possible results: 1 or 0. The results are described by functions of these variables, which we note $A_{x,y,z}$:

$$A_x = 0 \text{ or } 1; A_y = 0 \text{ or } 1; A_z = 0 \text{ or } 1 \quad (5.41)$$

At first sight, this seems to provide a total number of eight possibilities; but, if we want to preserve relation (5.40), we have to select among these eight possibilities only those three for which two A are 1, and one is 0. For this particular spin we then attribute colors to the three orthogonal directions Ox , Oy , and Oz : the two directions that get an $A = 1$ are painted in red, the last in blue [219].

The same operation can obviously be made for all possible choices of the triplet of directions $Oxyz$. A question which then naturally arises is: for an arbitrary direction Ox , can one attribute a given color (a given value for A_x) that remains independent of the context in which it was defined? Indeed, we did not define the value as a property of an Ox direction only, but in the context of two other directions Oy and Oz ; the possibility of a context independent coloring is therefore not obvious. Can we for instance fix Oz and rotate Ox and Oy around it, and still keep the same color for Oz ? We are now facing an amusing little problem of geometry that we might call “ternary coloring of all space directions”. Bell as well as Kochen and Specker showed that this is actually impossible; for a proof see either the original articles¹⁰, or the review [9] given by Mermin.

The conclusion is that any theory where measurements reveal a property that the system already had before measurement must be contextual to reproduce the predictions of quantum mechanics: the results of several compatible measurements performed at the same time must depend on the nature of all these measurements, otherwise contradictions appear. An important feature of this theorem is that it does not require the system to be in a precise quantum state, but applies to any state (we come back to state-independence in §5.5.3).

5.5.2 Two spin 1/2 particles, product rule

In a similar line, Peres [221] has shown that the results of quantum mechanics are incompatible with the two following propositions:

- (i) the results of the measurement of an operator A depend solely on A and on the system being measured (non-contextuality)
- (ii) if operators A and B commute, the result of a measurement of their product AB is the product of the results of separate measurements of A and B (product rule).

¹⁰ The original proof by Kochen and Specker involves projectors over 117 different directions in real space; since then proofs involving a smaller number have been given, for instance in [220] where 18 directions are used.

Peres shows the existence of the incompatibility by considering two spin 1/2 particles in a singlet state. Mermin has generalized the result [9, 222] to any state vector of the two spins, and we follow his proof to show that the two assumptions lead to a sign contradiction with the predictions of quantum mechanics¹¹. We consider two spin 1/2 particles and the following table of nine quantum variables (we use the same notation as in §5.1):

$$\begin{array}{ccc}
 \sigma_{1x} & \sigma_{2x} & \sigma_{1x}\sigma_{2x} \\
 \sigma_{2y} & \sigma_{1y} & \sigma_{1y}\sigma_{2y} \\
 \sigma_{1x}\sigma_{2y} & \sigma_{1y}\sigma_{2x} & \sigma_{1z}\sigma_{2z}
 \end{array} \tag{5.42}$$

All operators have eigenvalues ± 1 . The three operators belonging to the same line commute, as well as the three operators belonging to the same column (the products of two anti-commuting σ are commuting operators, since changing their order introduces two -1 signs with cancelling effects). Moreover, the products of all three operators is always $+1$, except the last column¹² for which it is -1 . Here, instead of an infinite number of triplets of directions in space, we have three groups of three operators, but the same question as above arises: can we attribute a value ± 1 to each of the nine elements of matrix (5.42) in a way that is consistent with the results of quantum mechanics? For this consistency to be satisfied, all lines and columns should contain either three $+1$ values or one $+1$ and two -1 , except the last column that must contain one or three -1 .

This little matrix problem is much simpler¹³ than the geometrical coloring problem mentioned in §5.5.1. One can calculate the product of all matrix elements, either as the product of the products inside each line or as the product of the products inside columns. The product of all lines has to be $(+1)^3 = +1$, while the product of all columns is $(+1)^2(-1)$, which is -1 ; a sign contradiction arises. It is therefore impossible to find nine numbers satisfying all conditions.

For another illustration of this sort of impossibility, see also §VI of [9] which deals with three 1/2 spins instead of two.

5.5.3 Contextuality versus local realism

From the discussions above, we can conclude that the predictions of quantum mechanics are incompatible with a non-contextual view on the EPR elements of

¹¹ We note in passing that the reasoning is very close to that of §5.1, which illustrates again the similarity between the GHZ theorem and this form of the BKS theorem.

¹² This can easily be checked from the well-known properties of the Pauli matrices; the minus sign for the third column comes from the product of the two i , arising from the relation $\sigma_x\sigma_y = i\sigma_z$; on the other hand, in the third line one gets $i \times (-i) = 1$ because of the change of order of the operators.

¹³ The complication of the geometrical problem of the original BKS theorem is entirely avoided by going to a space of states of dimension 4 instead of dimension 3.

reality/additional variables. Now, is this result more or less general than the Bell theorem, based on locality instead of contextuality? The Bell theorem, seen as a continuation of the EPR argument, does not assume additional variables: their existence is derived from the local realist EPR argument applied to perfect correlations. In the BKS theorem, such a derivation is not made; the existence of additional variables is assumed, as well as some specific properties. In this sense, the Bell theorem is more general. Moreover, as Bell noted [5], “the result of an observation may reasonably depend not only on the state of the system (including hidden/additional variables) but also on the complete disposition of the apparatus”. Mathematically, in a theory with additional variables λ , the function A giving the result of a measurement then depends on the settings of all the apparatuses interacting simultaneously with the measured quantum system; for instance, in the case studied in §5.5.1, the function should be written as $A(a, b, \lambda)$, where a and b define the two directions along which the squared measured components of the angular momentum are measured. One can also build a theory where additional variables are attributed to the apparatuses, and where both kinds of additional variables collaborate in order to determine the observed a and b dependent results; the results of quantum mechanics can then be recovered. By contrast, the Bell local realist theorem can handle variables associated with the measurement apparatuses – provided of course they are local – and maintain a contradiction with these results. Violations of the Bell theorem by quantum mechanics are therefore generally considered as more significant quantum manifestations than violations of the BKS theorem.

On the other hand, if one reasons from the beginning within hidden variable theories, and disconnects the Bell theorem from the EPR argument, the situation is different. Local hidden variables then appear as a special type of non-contextual hidden variables [223] (the Bell conditions of setting independence are a special case of non-contextuality). Seen in this way, the Bell theorem appears as less general than the BKS relations. Both theorems can be expressed in forms that are either state-dependent, or state-independent and therefore more general. Examples of state-dependent forms were given above when we assumed that the system should be put in a given quantum state, such as (4.1) or (5.1); by contrast, the reasonings used in §§5.5.1 and 5.5.2 are state-independent. For a general discussion of the status of the various “impossibility theorems” with emphasis on the BKS theorems, see [9, 222].

We have seen in Chapter 4 that local realism can give rise to several forms of the Bell theorem, leading to different equalities or inequalities. In the same way, non-contextual realism can be expressed in various mathematical ways. An all-or-nothing state-independent test involving two spins $1/2$ was proposed in [224]; extending the idea, [225] has proposed an all-or-nothing test that can be performed with single particles. This resulted, a few years later, in an experiment done with

single photons [226], which provided results in agreement with quantum mechanics, eliminating non-contextual hidden variables. Another inequality applying to three-level systems was introduced in [227] and [223], leading to experiments done with single photons propagating along three modes [228]; a violation of the inequality by more than 120 standard deviations was obtained, eliminating again non-contextuality. Experiments have also been performed with neutrons [229], trapped ions [230], and nuclear spins in a solid studied by nuclear magnetic resonance [231], providing other examples of good agreement with the predictions of quantum mechanics.

The notion of contextuality has been used to propose a realist interpretation of quantum mechanics, based on the notion of contextual objectivity [232]. In this formulation, realism is defined in a way that is reminiscent of the EPR criterion of reality, but emphasizes the role of the whole experimental setup (holism) rather than local properties of sub-systems or regions of space – in other words in a way that is in line with Bohr’s point of view. Under these conditions, clearly, neither the EPR incompleteness argument nor the Bell theorem apply, and contradictions with the standard formulation of quantum mechanics are avoided.

6

Quantum entanglement

In this chapter, we study the properties of quantum entanglement, and more generally the way correlations can appear in quantum mechanics. Quantum entanglement is an important notion that we have already discussed, for instance in the context of the Von Neumann chain or of the Schrödinger cat, but here we give more details on its properties.

In classical physics, the notion of correlation is well known. It hinges on the calculation of probabilities and on linear averages over a number of possibilities. A distribution gives the probability of having the first system in a some given state and the second in another state. If this distribution is not a product, the two systems are correlated. If it is a product, they are uncorrelated; measuring the properties of one system does not bring any information on the other. This is in particular the case if the state of each of the two systems is perfectly defined (which also defines the state of the whole system perfectly well). The notion of correlation between sub-systems therefore stems from the multiplicity of possible states of the whole system; fluctuations of this state are necessary to give its full meaning to the classical notion of correlation.

In quantum mechanics, the situation is different: as we have seen (in particular in Chapter 4), even a physical system that is perfectly defined by a given state vector already contains fluctuations. This leads naturally to another notion of correlation, independent of any fluctuation of the state. For instance, the components of two $1/2$ spins are strongly correlated in a singlet state (§4.1.1), which is a pure state. This is because the principle of linear superposition of quantum mechanics allows one to introduce superpositions directly inside the state vector; this is very different from superpositions of probabilities, which are quadratic functions of this state vector. In other words, in quantum mechanics one may introduce correlations directly at the level of probability amplitudes, one level “below” linear combinations of probabilities; this may create cross terms in the amplitudes resulting in quantum interference effects. Since it is also possible in quantum mechanics to assume that

the state of the whole system is imperfectly known and defined only by probabilities, the two levels may coexist, so that the quantum notion of correlation is much richer than in classical physics.

We come back in more detail to these points in §6.1, and to methods to characterize entanglement in §6.2. In §6.3, we discuss how quantum entanglement may be created in experiments, as well as an opposite process, decoherence, which tends to make entanglement disappear¹. The dynamics of a sub-system of a large quantum system and the corresponding master equation are studied in §6.4.

6.1 A purely quantum property

In quantum mechanics, the relation between parts and the whole is very special, and certainly non-intuitive. We have already mentioned that Schrödinger was the first to use the words “quantum entanglement” in 1935, when he wrote (page 555 of [233]): “When two systems, of which we know the states by their respective representatives, enter into temporary physical interaction due to known forces between them, and when after a time of mutual influence the systems separate again, then they can no longer be described in the same way as before, viz. by endowing each of them with a representative of its own. I would not call that *one* but rather *the* characteristic trait of quantum mechanics, the one that enforces its entire departure from classical lines of thought. By the interaction the two representatives [the quantum states] have become entangled . . . Another way of expressing the peculiar situation is: the best possible knowledge of a whole does not necessarily include the best possible knowledge of all its parts, even though they may be entirely separate and therefore virtually capable of being ‘best possibly known’, that is, of possessing, each of them, a representative of its own. The lack of knowledge is by no means due to the interaction being insufficiently known – at least not in the way that it could possibly be known more completely – it is due to the interaction itself.”

6.1.1 The part and the whole

In classical mechanics, the description of the part is simply contained in the description of the whole²: when a system is made of two sub-systems, 1 and 2, a full description of the whole system $1 + 2$ immediately provides a full description of

¹ More precisely, it tends to make it propagate further and further into the environment, so that in practice it rapidly becomes impossible to observe; for all practical purposes, one can consider that entanglement has disappeared.

² Heisenberg published in 1969 a book with the title *The Part and the Whole*, relating how quantum mechanics emerged from discussions between him and other physicists. Nevertheless this title does not seem to be specifically related to entanglement, but to a more abstract concept: “There is a fundamental error in separating the parts from the whole, the mistake of atomizing what should not be atomized. Unity and complementarity constitute reality” (sentence attributed to Heisenberg [234]).

each sub-system. Actually, the description of the dynamical state of the whole system is nothing but the sum of the descriptions of the sub-systems: if one specifies the values of the positions and the momenta of all particles (as well as the values in space of the fields), all dynamical variables associated to each sub-system are automatically known. For instance, a perfect description of the solar system contains a perfect description of all the positions and velocities of the planets and their satellites; a perfect description of the state of the electromagnetic field contains a perfect description of the electric field, as well as of the magnetic field, in all space.

Strangely enough, and as emphasized by Schrödinger, this is no longer true in quantum mechanics. If the whole system is described with the most accurate way accessible in quantum mechanics, a state vector, the sub-systems may be described less accurately; one can assign them just probabilities of occupying quantum states – in other words, as we see in more detail below, they are not in a pure state but in a statistical mixture, described mathematically by a density operator. Therefore, parts may be known only statistically, while the whole is perfectly known (with probability unity).

To see how this can happen in simple terms, suppose that a system is made of a first sub-system 1 with quantum states $|\varphi\rangle$, and of a second sub-system 2 with quantum states $|\chi\rangle$. If the whole system is described by a simple product quantum state³:

$$|\Psi\rangle = |1 : \varphi\rangle |2 : \chi\rangle \quad (6.1)$$

the two sub-systems are uncorrelated; the first is described by a quantum state $|\varphi\rangle$, while the second is described by $|\chi\rangle$. All three systems (whole and two sub-systems) are then associated with state vectors, the most accurate description that standard quantum mechanics can assign to any physical system.

But assume now that the whole system is in state:

$$|\Psi\rangle = \alpha |1 : \varphi_a\rangle |2 : \chi_a\rangle + \beta |1 : \varphi_b\rangle |2 : \chi_b\rangle \quad (6.2)$$

where $|1 : \varphi_{a,b}\rangle$ are two orthonormal states for the first system, and $|2 : \chi_{a,b}\rangle$ two other orthonormal states associated with the second; α and β are two (non-vanishing) complex numbers satisfying $|\alpha|^2 + |\beta|^2 = 1$. The whole system is still in a well defined quantum state, but the first sub-system has a probability $|\alpha|^2$ of being in state $|\varphi_a\rangle$ and a probability $|\beta|^2$ of being in state $|\varphi_b\rangle$, so that it is described by

³ As already mentioned in Chapter 11 – see equation (11.61) – we use three equivalent notations for the tensor product:

$$|1 : \varphi\rangle \otimes |2 : \chi\rangle \equiv |1 : \varphi\rangle |2 : \chi\rangle \equiv |1 : \varphi; 2 : \chi\rangle$$

Depending on the context of each calculation, one or the other may be more convenient and correspond better to the general use.

a statistical mixture instead of a pure state⁴ – we come back to this point below in terms of density operators. The sub-systems are then described less accurately than the whole system, a situation without any classical equivalent. The words “quantum entanglement” are used to describe such situations.

The formalism of the density operator (§11.1.5) and the use of partial traces (§11.2.3) is useful to express this property in a more quantitative way. The density operator ρ associated with state $|\Psi\rangle$ is:

$$\rho = |\Psi\rangle\langle\Psi| \quad (6.3)$$

which is merely the projector over state $|\Psi\rangle$, obeying $\rho^2 = \rho$ since:

$$\rho^2 = |\Psi\rangle\langle\Psi| |\Psi\rangle\langle\Psi| = |\Psi\rangle\langle\Psi| = \rho \quad (6.4)$$

One then says that ρ describes a “pure state”, a case characterized by the equality:

$$\text{Tr}\{\rho^2\} = 1 \quad (6.5)$$

(the trace of ρ itself is always 1 by definition). The sub-systems are described by partial traces of ρ ; for instance, the first sub-system is described by operator:

$$\rho_1 = \text{Tr}_2\{\rho\} \quad (6.6)$$

with the following matrix⁵ in basis $|\varphi_a\rangle, |\varphi_b\rangle$:

$$(\rho_1) = \begin{pmatrix} |\alpha|^2 & 0 \\ 0 & |\beta|^2 \end{pmatrix} \quad (6.7)$$

This diagonal expression shows that system 1 is described by a statistical mixture of states $|\varphi_a\rangle$ and $|\varphi_b\rangle$ with probabilities $|\alpha|^2$ and $|\beta|^2$.

We also have:

$$(\rho_1)^2 = \begin{pmatrix} |\alpha|^4 & 0 \\ 0 & |\beta|^4 \end{pmatrix} \quad (6.8)$$

⁴ Similarly, the second sub-system has a probability $|\alpha|^2$ of being in state $|\chi_a\rangle$ and a probability $|\beta|^2$ of being in state $|\chi_b\rangle$, and is therefore also associated with a statistical mixture.

⁵ According to equation (11.77) of Chapter 11 defining a partial trace, the matrix elements of ρ_A are given by:

$$\langle\varphi_i|\rho_A|\varphi_j\rangle = \sum_k \langle 1:\varphi_i; 2:\chi_k|\Psi\rangle\langle\Psi|1:\varphi_j; 2:\chi_k\rangle$$

(where i, j and k label one of the basis vectors, a or b). From expression (6.2) of $|\Psi\rangle$, we then obtain:

$$\langle\varphi_i|\rho_A|\varphi_j\rangle = (\alpha\delta_{i,a})(\alpha^*\delta_{j,a}) + (\beta\delta_{i,b})(\beta^*\delta_{j,b})$$

which is matrix (6.7).

and:

$$\begin{aligned} \text{Tr} \{(\rho_1)^2\} &= |\alpha|^4 + |\beta|^4 = [|\alpha|^2 + |\beta|^2]^2 - 2|\alpha|^2 |\beta|^2 \\ &= 1 - 2|\alpha|^2 |\beta|^2 \leq 1 \end{aligned} \quad (6.9)$$

If one of the coefficients α or β vanishes, we are back to the case where the sub-systems are uncorrelated as in (6.1); one then obtains $(\rho_1)^2 = \rho_1$ and the trace of $(\rho_1)^2$ is equal to 1, which indicates that the first sub-system is in pure state (as well as the second sub-system). But, when none of the coefficients vanishes, the sub-systems are correlated, we have $(\rho_1)^2 \neq \rho_1$, and the trace of $(\rho_1)^2$ is smaller than 1; this indicates that the sub-system is described with less accuracy than if it was in a pure state, by a statistical mixture.

Of course, one can add more than two terms in the sum of (6.2), which tends to make the knowledge on each sub-system less and less precise, while the whole system still remains perfectly defined. For instance, if $|\Psi\rangle$ contains a linear superposition of three terms, all containing orthogonal individual states, the density operator ρ_1 describing system 1 becomes a 3×3 diagonal operator and introduces three possible states for this system (eigenstates of ρ_1). Moreover, in general and as remarked by Schrödinger [233], the determination of these possible states is not necessarily unique; one can then not even list the states that are accessible to the sub-system without ambiguity⁶. One can even reach a situation where the description of each sub-system is completely random, so that no information at all is available on their states, all kets in their space of states being equally probable. In such cases of extreme entanglement, the whole system has some physical properties, those associated with its state, but nothing can be said about any specific property of the sub-systems. A simple example of extreme quantum entanglement is given by the singlet state of two spins 1/2 – cf. equation (4.1) – or its direct generalization:

$$|\Psi\rangle = \frac{1}{\sqrt{2}} \left[|+, -\rangle + e^{i\xi} |-, +\rangle \right] \quad (6.10)$$

(where $e^{i\xi}$ is an arbitrary phase factor). With this state, the whole system is perfectly defined, while no information at all is available on the state of each individual spin, which has equal probabilities to be in state $|+\rangle$ or $|-\rangle$.

The purely quantum character of entanglement is clear. In classical mechanics, if the properties of two physical systems are initially defined with the best possible accuracy, after mutual interactions each system is still defined as accurately as possible. By contrast, in quantum mechanics, starting from systems with the best possible definition compatible with their quantum nature, at the end of the process

⁶ This happens each time two probabilities (eigenvalues of the reduced density operator) are equal.

one reaches in general situations where they are defined with less accuracy. This is a signature of the quantum character of the entanglement created by the interaction. In terms of entropy of the two systems (§6.2.2), if initially they both have zero entropy, the same is true at the end of the process, while in quantum mechanics this is not the case.

6.1.2 Two possible origins of correlations

Entanglement and correlation are obviously closely related notions. As we have already mentioned in the introduction to this chapter, correlations may arise in quantum mechanics in two ways:

- (i) By a process that is the direct transposition of classical correlations: one attributes random correlated quantum state vectors to each of the two sub-systems. The statistical average is then performed linearly with respect to probabilities, and one then says that the whole system is described by a statistical mixture of product states (or by a density operator). This leads to classical correlations, which necessarily obey the Bell inequalities (§4.1 and Appendix C).
- (ii) By a purely quantum process involving quantum entanglement: the whole system is described by a state vector that is neither random nor a product. The linear superposition then occurs at the level of the state vector itself, which is a coherent sum of several components, with a relative phase that plays an important role. The correlations then have a purely quantum origin, and may violate the Bell inequalities.

The former process takes place at the level of events, properties of systems, results of measurement, etc. as in classical physics. The latter takes place “at a lower level” involving directly state vectors and probability amplitudes (instead of probabilities themselves). It can also result in a broader range of correlations, as illustrated by the Bell theorem.

Both possibilities may also be combined: one can assign to the whole system itself a statistical mixture of states that are not necessarily products. The formalism of the density operator (§11.1.5) makes it possible to include both into the same operator, which is very convenient for compact calculations; on the other hand, it loses the trace of the origin of correlations⁷, classical or quantum, which may also

⁷ The formalism of the density operator, or matrix, is elegant and compact, but precisely because it is compact it sometimes partially hides the physical origin of the mathematical terms. The density matrix allows one to treat in the same way classical probabilities (proper mixtures), arising from non-fundamental uncertainties and imperfect knowledge of a physical system, and purely quantum probabilities (improper mixtures) which are more fundamental and have nothing to do with imperfect knowledge. But mathematical analogies should not obscure conceptual difficulties!

sometimes be inconvenient. One sometimes calls “proper mixtures” the situations associated with usual quantum statistics, and “improper mixtures” those associated with purely quantum correlations.

6.2 Characterizing entanglement

We now study how quantum entanglement can be characterized more quantitatively. For this purpose, we come back to the properties of the density operator ρ of the whole system, when compared to those of the density operators ρ_1 and ρ_2 of sub-systems 1 and 2.

6.2.1 Schmidt decomposition of a pure state

We now generalize the discussion of §6.1.1. The quantum system containing sub-systems 1 and 2 is described by a normalized state vector $|\Psi\rangle$, or equivalently by the density operator ρ given by the projector onto $|\Psi\rangle$; relation (6.5) is obeyed, as expected for a pure state. Each of the sub-systems is described by density operators obtained by partial traces:

$$\rho_1 = \text{Tr}_2 \{\rho\} \quad ; \quad \rho_2 = \text{Tr}_1 \{\rho\} \quad (6.11)$$

These two operators are Hermitian and non-negative; their matrices can therefore be diagonalized, with eigenvalues that are real, comprised between 0 and 1. We call $|u_i\rangle$ the eigenvectors of ρ_1 (index i takes P different values, where P is the dimension of the space of states of sub-system 1) with eigenvalues q_i , all positive or zero (but not necessarily all distinct). Similarly, the kets $|v_j\rangle$ are the eigenvectors of ρ_2 (where j takes Q different values, Q being the dimension of the space of states of the second sub-system) and r_j the corresponding eigenvalues. The two partial density operators can then be written as:

$$\rho_1 = \sum_{i=1}^P q_i |u_i\rangle \langle u_i| \quad \rho_2 = \sum_{l=1}^Q r_l |v_l\rangle \langle v_l| \quad (6.12)$$

with $0 \leq q_i, r_j \leq 1$.

We can now expand $|\Psi\rangle$ onto the base of tensor products $\{|1 : u_i\rangle \otimes |2 : v_l\rangle\}$, for which we use the simplified notation $\{|u_i, v_l\rangle\}$; we therefore write:

$$|\Psi\rangle = \sum_{i,l} x_{i,l} |u_i, v_l\rangle \quad (6.13)$$

where the $x_{i,l}$ are the components of $|\Psi\rangle$ in this basis. If we introduce the ket $|\bar{w}_i\rangle$, which belongs to the space of states of sub-system 2, by:

$$|\bar{w}_i\rangle = \sum_l x_{i,l} |v_l\rangle \quad (6.14)$$

expression (6.13) becomes:

$$|\Psi\rangle = \sum_i |u_i, \bar{w}_i\rangle \quad (6.15)$$

From the definition (11.77) of partial traces⁸ in Chapter 11, we then obtain:

$$\rho_1 = Tr_2 \{ |\Psi\rangle \langle \Psi| \} = \sum_{i,j} |u_i\rangle \langle u_j| \times \langle \bar{w}_j | \bar{w}_i \rangle \quad (6.16)$$

But, by construction of the basis $\{|u_i\rangle\}$ we have used in the calculation, ρ_A is diagonal and given by expression (6.12); comparing with (6.16) shows that we necessarily have:

$$\langle \bar{w}_j | \bar{w}_i \rangle = \delta_{i,j} \times q_i \quad (6.17)$$

For all values of index i that correspond to non-zero eigenvalues q_i , this relation implies that we can define an ensemble of orthonormal kets $|w_i\rangle$ belonging to the space of states of sub-system 2 by:

$$|w_i\rangle = \frac{1}{\sqrt{q_i}} |\bar{w}_i\rangle \quad (6.18)$$

For all values of index i associated to zero eigenvalues q_i , the same relation shows that the kets $|\bar{w}_i\rangle$ must vanish.

Therefore, expression (6.15) becomes:

$$|\Psi\rangle = \sum_i \sqrt{q_i} |u_i, w_i\rangle \quad (6.19)$$

where the $|u_i\rangle$ are a set of orthonormal vectors in the space of states of the first system, and where the $|w_i\rangle$ are another set of orthonormal vectors in the second space of states. The sum over i may run, at choice, over all possible P values,

⁸ According to definition (11.77) of a partial trace, we have $\langle u_i | \rho_1 | u_{j'} \rangle = \sum_m \langle u_i, v_m | \Psi \rangle \langle \Psi | u_{j'}, v_m \rangle$, where we insert the expression of $|\Psi\rangle \langle \Psi|$ obtained from (6.15). One then obtains $|\Psi\rangle \langle \Psi| = \sum_{i',j'} |u_{i'}, \bar{w}_{i'}\rangle \langle u_{j'}, \bar{w}_{j'}|$. In the summation, the only relevant terms have $i' = i$ and $j' = j$, and we finally obtain $\langle u_i | \rho_1 | u_{j'} \rangle = \sum_m \langle v_m | \bar{w}_j \rangle \langle \bar{w}_{j'} | v_m \rangle = \langle \bar{w}_{j'} | \bar{w}_j \rangle$, which is equivalent to (6.16).

or only over those for which the eigenvalue q_i is non-zero. This expression is the Schmidt decomposition of a pure entangled state; it is also sometimes called the “bi-orthonormal decomposition”, and plays an important role in the modal interpretation of quantum mechanics (§10.7).

If we now come back to the partial traces ρ_1 and ρ_2 , a calculation from (6.19) provides two symmetrical expressions:

$$\rho_1 = \sum_i q_i |u_i\rangle \langle u_i| \quad (6.20)$$

(already known) and:

$$\rho_2 = \sum_i q_i |w_i\rangle \langle w_i| \quad (6.21)$$

This shows that the two partial density operators always have the same eigenvalues⁹, with a sum equal to 1, since the two operators have unit trace. In the special case where they are all zero except one, each of the sub-systems is in a pure state. But, in general, several eigenvalues are non-zero, and we then immediately see that $(\rho_1)^2$ is not equal to ρ_1 , and, similarly, for ρ_2 . We then find again a case where the two sub-systems are described by statistical mixtures while the whole system is in a pure state.

The number of non-zero eigenvalues q_i , in other words the number of effective terms in (6.19), is called the “Schmidt rank” R of $|\Psi\rangle$. If $R = 1$, the state of the whole system is not entangled, and the two sub-systems are in pure states. If $R = 2$, the situation is that discussed in the example of §6.1.1; if $R = 3$, the entanglement is more complicated, etc. Because R does not depend on the sub-system, 1 or 2, entanglement is shared by them in a symmetrical way; for instance, one of the sub-systems cannot be in a pure state while the other is in a statistical mixture. The dimension of the space of state of sub-system 2 with which 1 is entangled gives an upper bound to the number of independent kets $|w_i\rangle$, and therefore to the rank R ; actually, R cannot exceed the dimension of any of the spaces of states of the sub-systems: a complicated entanglement between them therefore requires that both should have spaces of states with sufficient dimensions.

6.2.2 Statistical entropies

We can associate a statistical entropy (§11.1.5.d) to each density operator ρ . We now compare the entropy of ρ to those associated with the two partial density operators ρ_1 and ρ_2 .

⁹ These properties are specific of pure states for the whole system; they are not necessarily true if it is described by a statistical mixture.

In the case studied in the preceding section, the whole physical system is in a pure state, its density operator is the projector (6.3) over a single state $|\Psi\rangle$, with a vanishing associated entropy $S = 0$. Nevertheless, the two sub-systems are not in general in pure states (except if $|\Psi\rangle$ is a tensor product, without any entanglement), so that:

$$\begin{aligned} S_1 &= -k_B \text{Tr} \{\rho_1 \ln \rho_1\} \geq 0 \\ S_2 &= -k_B \text{Tr} \{\rho_2 \ln \rho_2\} \geq 0 \end{aligned} \quad (6.22)$$

We then have:

$$S_1 + S_2 \geq S \quad (6.23)$$

The equality corresponds to the particular case where $|\Psi\rangle$ is a product, and where the Schmidt rank is equal to 1.

In a more general situation, the whole system is described by a density operator ρ that does not necessarily correspond to a pure state, so that its entropy S does not vanish either. One can nevertheless show that this entropy S always remains smaller than, or equal to, the sums of the two entropies of the sub-systems¹⁰; in other words, relation (6.23) remains valid in this case. The equality is obtained only when ρ is a product:

$$\rho = \rho_1 \otimes \rho_2 \quad (6.24)$$

which corresponds to a case where both sub-systems are described by statistical mixtures, but remain completely uncorrelated. The difference $S_1 + S_2 - S$ therefore gives a quantitative estimation of the loss of accuracy between the quantum description of the whole system and the separate quantum descriptions of the two sub-systems.

6.2.3 Measures of entanglement

If the whole system $S_1 + S_2$ is in a pure state $|\Psi\rangle$, the Schmidt decomposition shows that:

$$S_1 = S_2 = \sum_i q_i \ln q_i \geq 0 \quad (6.25)$$

It is then natural to choose S_1 as the definition of the entanglement of the two sub-systems. But one may also wish to define a precise measure of the degree of

¹⁰ One sometimes speaks of “sub-additivity of the entropy” to express the fact that the entropy of the whole system is less than the sum of the entropies of the two sub-systems.

entanglement of two quantum sub-systems, whether the whole system is in a pure state or not. To be physically acceptable, this measure must obey several criteria. For instance, it must obviously vanish if the whole system is in a state that is a tensor product where the two sub-systems are not correlated. It is also necessary that the degree of entanglement should remain invariant when Alice and Bob, each acting on two remote sub-systems, perform local operations on them [235]. For instance, maximally entangled states, such as those we have studied in §5.2, keep the same form if different unitary transformations are applied to the spins (footnote 4 of that section). Several such measures have been proposed, generally related to Von Neumann entropies; for a review, see for instance [236]. When the number of entangled systems is three or more, the definition of their degree of entanglement becomes more difficult; for the moment, there is no generally accepted definition of a degree of entanglement of a multipartite system.

One could also expect that a system giving rise to maximal violations of the BCHSH inequalities should reach the maximal degree of entanglement. Actually, it turns out that the notions of entanglement and of non-locality are not directly related, except in a few special cases (two spins 1/2 for instance). Curiously, maximally entangled states generally produce less non-local effects than non-maximally entangled states [237]. Non-locality and entanglement are therefore really different concepts.

6.2.4 Monogamy

A state such as (6.10) is a state where the two spins are strongly correlated, while a GHZ state such as (5.1) can be regarded as the equivalent for three spins. One might think that the latter state conserves the entanglement between the two first spins available in the former, while they have become entangled with the third. In fact, this is wrong: while going from (6.10) to (5.1), it is true that one entangles the third spin with the group of the two others, but at the same time one totally destroys the entanglement within the initial group of two spins. Actually, we have already noticed this property, when we remarked in §5.2 that these states exhibit strong correlations between the spins only if all of them are measured; if the measurements relate to two spins instead of three, no correlation between them appears whatsoever.

Under these conditions, what can we do to add an additional spin without destroying the correlation between the two initial spins? We can assume the following form of the three-spin state:

$$\begin{aligned} |\Psi\rangle &= \frac{1}{\sqrt{2}} \left[|1: +; 2: -\rangle + e^{i\xi} |1: -; 2: +\rangle \right] \otimes |3: \theta\rangle \\ &= \frac{1}{\sqrt{2}} \left[|1: +; 2: -; 3: \theta\rangle + e^{i\xi} |1: -; 2: +; 3: \theta\rangle \right] \end{aligned} \quad (6.26)$$

(where $|\theta\rangle$ is any normalized state for the third spin), with a factorized state for the additional spin; this choice retains the same entanglement between spins 1 and 2 as (6.10)¹¹. But then the third spin is totally uncorrelated with the two former!

A compromise between the two preceding attempts is:

$$\frac{1}{\sqrt{2}} \left[|1 : +; 2 : -; 3 : \theta_1\rangle + e^{i\xi} |1 : -; 2 : +; 3 : \theta_2\rangle \right] \quad (6.27)$$

If $|\theta_1\rangle = |\theta_2\rangle$, one recovers (6.26), and the additional spin is not entangled with the two first; if $|\theta_1\rangle$ and $|\theta_2\rangle$ are orthogonal, one recovers a GHZ state where none of the three pairs has any entanglement, this property being restricted to all three particles. When $|\theta_1\rangle$ and $|\theta_2\rangle$ are neither parallel nor orthogonal, we have an intermediate situation: the more parallel they are, the more entangled the two first spins remain (we will see in the next section that the coherent terms contain the scalar product $\langle \theta_1 | \theta_2 \rangle$), but the third then has little entanglement. Conversely, the more orthogonal they are, the less entangled the two initial spins are: they lose their correlations in favor of a three-spin correlation. The third spin actually plays for the two others a role that is similar to the role of the environment in decoherence (§6.3.3.a): indeed, the environment destroys the coherence of the initial state more efficiently when it correlates to its components with states that are almost orthogonal.

This is a general property: if two systems are maximally entangled, a principle of mutual exclusion makes it impossible to be entangled with a third system. Mathematically, the property can be expressed through the Coffman–Kundu–Wootters inequality¹² [238, 239]. It has no classical equivalent, since classically nothing forbids the correlation of a third system with two others without destroying the initial correlation. One often calls this property “entanglement monogamy”. It is for instance possible to show that, if two quantum sub-systems A and B are mutually entangled as well as to a third system C, and if two of them are sufficiently correlated to create violations of the BCHSH inequalities with measurements on these sub-systems, then the inequalities are necessarily satisfied for all measurements related to the two other pairs of sub-systems [240, 241].

6.2.5 Separability criterion for density operators

We have seen in §3.3.3.b that the EPR argument and the Bell theorem are strongly related to the notion of quantum non-separability; they are also related with the

¹¹ In the calculation of the effect of the coherent term (in $e^{i\xi}$) with the method used in §5.2, it is now sufficient to flip two spins, the third remaining always in state $|\theta\rangle$.

¹² Reference [238] makes use of a measure of entanglement called “concurrence”. It shows that the sum of the squares of the concurrence between A and B and of that between A and C cannot exceed the concurrence between A and the BC pair.

notion of quantum entanglement, since violations of the Bell inequalities require the presence of quantum entanglement. In terms of density operators, how is it possible to recognize quantum non-separability?

Consider a quantum system made of two sub-systems 1 and 2 and described by a density operator ρ . If ρ can be expanded in terms of a series of density operators ρ_1^n and ρ_2^n relative to each sub-system according to:

$$\rho = \sum_n w_n \rho_1^n \otimes \rho_2^n \quad (6.28)$$

with real positive coefficients w_n , one says that ρ is separable¹³ [136, 242]. When the traces of all density operators are normalized to 1, one obtains:

$$1 = \sum_n w_n \quad \text{and therefore} \quad 0 \leq w_n \leq 1 \quad (6.29)$$

which shows that the w_n can be interpreted as probabilities¹⁴; the event where first sub-system is described by ρ_1^n and the second by ρ_2^n has probability w_n . If the expansion (6.28) necessarily contains coefficients w_n that are not real positive numbers, one says that the density operator ρ is “non-separable” and contains quantum entanglement¹⁵.

If one performs separate measurements on sub-systems 1 and 2, and if the whole system has a separable density operator, the results always obey the Bell inequalities (nevertheless, the opposite is not necessarily true: non-separable density operators do not necessarily give rise to violations of the inequalities). In such a situation, even if each of the two sub-systems has strong quantum properties, the way correlations between them are introduced remains purely classical. A separable system cannot have quantum entanglement, even if each sub-system exhibits extreme quantum properties. The Peres–Horodecki criterion [242, 243] indicates that a necessary condition for a density matrix to correspond to a separable density operator ρ is that the matrix obtained by partial transposition (where only the indices relative to one of the sub-systems are transposed, not for the other) should have non-negative eigenvalues. The appearance of negative eigenvalues may then signal quantum entanglement, with a better sensitivity than Bell violations.

¹³ In his initial article, Werner uses the words “classically correlated” [136], but the word “separable” chosen by Peres [242] is more frequently used nowadays.

¹⁴ If one replaces the discrete variable n by a continuous variable λ , one obtains the case studied in Appendix C.

¹⁵ For two spins 1/2, one can easily show that the density operator associated with a singlet state, or its direct generalization (6.10), necessarily involves negative expansion coefficients. Two spins in a singlet state are therefore not separable, and possess quantum entanglement.

6.3 Creating and losing entanglement

Historically, at Schrödinger's time, quantum entanglement was seen as a rather rare phenomenon, playing a role mostly in thought experiments. Nowadays, many experimental methods have been invented to obtain entanglement. Actually, quantum entanglement has now become an essential part of quantum information, quantum cryptography, teleportation, etc. All experiments mentioned in §4.1.5 involve pairs of entangled particles, often photons with entangled polarization variables.

6.3.1 Entanglement created by local interactions

As Schrödinger had initially suggested (see quotation in §6.1), one way to obtain entanglement between physical sub-systems is to make use of local interactions between particles. An atom emitting two photons in succession may provide such a scheme, which in fact has been used in many experiments. Initially, the atom is in an excited state, then emits a first photon, and reaches an intermediate state that depends on the polarization characteristics of the emitted photon; at this stage, the atom plus photon system is described by an entangled state, with coherent components on several polarization states associated with intermediate states of the atom. Each of these components then gives rise to the emission of a second photon with different polarizations, while the atom itself reaches a ground state that is independent of the polarizations of the emitted photons. This corresponds to the case $|\theta_1\rangle = |\theta_2\rangle$ in (6.27), where the state of one of the three particles (here, the final state of the atom) factorizes; the atom leaves the quantum entanglement party, which allows the two photons to enter maximal entanglement.

An often cited example is that of the atomic cascade $J = 0 \rightarrow 1 \rightarrow 0$ of the Calcium atom. It provides, by successive spontaneous emissions, two photons in a state:

$$|\Psi\rangle = \frac{1}{\sqrt{2}} \left[|1 : H; 2 : H\rangle + |1 : V; 2 : V\rangle \right] \quad (6.30)$$

where $|H\rangle$ and $|V\rangle$ are two states of polarization (horizontal and vertical) – for photons¹⁶ these states are analogous to states $|+\rangle$ and $|-\rangle$ for a spin 1/2. We have also seen that the techniques of parametric down-conversion in non-linear optics provide similarly entangled photons, often in more favorable experimental conditions.

¹⁶ These polarizations are mutually perpendicular as well as perpendicular to their direction of propagation. For the sake of simplicity, we limit our discussion to the polarization variables of the photons, which are the essential variables in the discussion of the BCHSH inequalities. Nevertheless, photons have other variables, frequency (energy) and direction of polarization, which we consider as fixed here.

Quantum entanglement is not limited to photons, but can also be obtained with particles with rest mass. For instance, the experiment described in [117] studies the correlation between the spins of two protons after a low energy collision between a beam of protons and an hydrogen target; the reference [167] proposes using the dissociation of Mercury dimers (isotope 199 has a nuclear spin 1/2) to obtain atoms with correlated spin variables. Experiments have also been done where two atoms are entangled because they exchange a single photon confined in a cavity with extremely high quality factor [244].

Trapped ions provide other possibilities to obtain quantum entanglement, in a case where the particles are localized and may be observed for a long time. The reference [245] describes how such an entanglement has been obtained with Beryllium ions sitting in a radio-frequency Paul trap, according to the method proposed by Cirac and Zoller [246]. For a review of experiments with entangled ions, see [247]. Entanglement may also be observed in solid-state physics with superconducting currents [248] involving a very large number of electrons (macroscopic systems). We now discuss still another method, entanglement swapping, which produces entangled particles without any interaction between them, only under the effect of the process of quantum measurement acting on other particles.

6.3.2 Entanglement swapping

The “entanglement swapping” method can entangle two particles created by two independent remote sources (the particles then have no common past) under the effect of the process of quantum measurement [249, 250]. Assume that we have two independent sources S_{12} and S_{34} , each emitting a pair of entangled photons, 1 and 2 for the former, 3 and 4 for the latter (Figure 6.1). The state describing both pairs is a tensor product of terms that are similar to (6.30):

$$|\Psi\rangle = \frac{1}{2} [|1:H; 2:H\rangle + |1:V; 2:V\rangle] \otimes [|3:H; 4:H\rangle + |3:V; 4:V\rangle] \quad (6.31)$$

If we introduce the 4 Bell states relative to particles i, j defined as:

$$\begin{aligned} |\Phi_{i,j}^B\rangle_{(\pm)} &= \frac{1}{\sqrt{2}} [|i:H; j:H\rangle \pm |i:V; j:V\rangle] \\ |\Theta_{i,j}^B\rangle_{(\pm)} &= \frac{1}{\sqrt{2}} [|i:H; j:V\rangle \pm |i:V; j:H\rangle] \end{aligned} \quad (6.32)$$

we obtain an orthonormal basis in the space of states associated with particles i and j . Since:

$$|\Phi_{1,4}^B\rangle_{(+)} \otimes |\Phi_{2,3}^B\rangle_{(+)} + |\Phi_{1,4}^B\rangle_{(-)} \otimes |\Phi_{2,3}^B\rangle_{(-)} = [|HHHH\rangle + |VVVV\rangle]$$

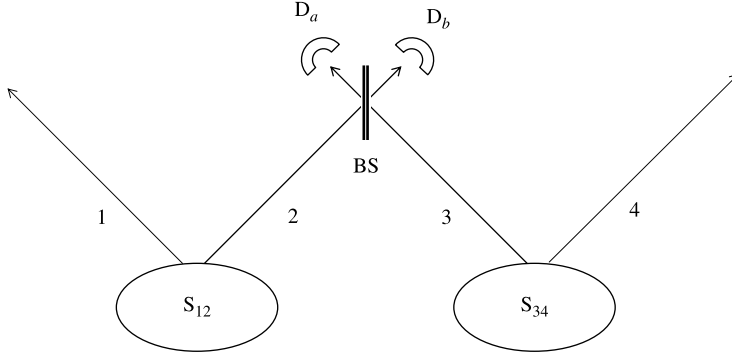


Figure 6.1. Schematic representation of entanglement swapping. Two independent sources S_{12} and S_{34} emit each a pair of particles, numbered 1 and 2 for the former, 3 and 4 for the latter. A half-reflecting beam splitter BS is inserted on the trajectory of particles 2 and 3, and two detectors D_a and D_b measure the number of particles in the two output channels. This measurement can project the state vector and put particles 1 and 4 into a completely entangled state, even if they have never interacted with each other.

(to simplify the notation, we implicitly assume that, in the right-hand side, the particles are always in the order 1, 2, 3 and 4) and:

$$|\Theta_{1,4}^B > (+)\rangle \otimes |\Theta_{2,3}^B > (+)\rangle + |\Theta_{1,4}^B > (-)\rangle \otimes |\Theta_{2,3}^B > (-)\rangle = [|HHVV\rangle + |VVHH\rangle] \quad (6.33)$$

we can write state (6.31) in the form:

$$|\Psi\rangle = \frac{1}{2} \left[|\Phi_{1,4}^B > (+)\rangle \otimes |\Phi_{2,3}^B > (+)\rangle + |\Phi_{1,4}^B > (-)\rangle \otimes |\Phi_{2,3}^B > (-)\rangle + |\Theta_{1,4}^B > (+)\rangle \otimes |\Theta_{2,3}^B > (+)\rangle + |\Theta_{1,4}^B > (-)\rangle \otimes |\Theta_{2,3}^B > (-)\rangle \right] \quad (6.34)$$

Suppose now that the experiment shown schematically in Figure 6.1 is performed: particles 2 and 3 undergo an interference experiment with eigenvectors that are the four Bell states for these two particles – this can be obtained by sending the two particles onto a semi-reflecting beam splitter and observing at which detectors, D_a or D_b , they are found in the output channels¹⁷. The projection onto one of the four Bell states for these particles projects the state of the two others onto the same Bell state. The two non-observed particles are then put into a completely entangled state. What is remarkable in the process is that, initially, the 1, 2 pair is

¹⁷ Among the four Bell states, the only one for which each output channel contains one particle is $|\Theta_{2,3}^B > (-)\rangle$; the three other states correspond to situations where the two particles always exit through the same channel (Hong–Ou–Mandel effect). If the two detectors D_a and D_b each register a particle, the measurements project the remaining particles 1 and 4 into the completely entangled state $|\Theta_{1,4}^B > (-)\rangle$.

internally entangled, but not with the 3, 4 pair, which also has only internal entanglement. By performing an appropriate measurement on one particle from each pair, one projects the two remaining particles into a completely entangled state, even if these two particles have never interacted.

We have already mentioned the possibility to entangle a larger number of particles [209] with similar schemes. Other protocols have been implemented to create quantum entanglement, and to put six ions in a NOON state [251] and up to eight into a W state (coherent superposition of states where a single excitation is localized on any of the ions, with the same probability amplitude) [252]. In quantum optics, the techniques of parametric down-conversion in non-linear crystals have been used in experiments providing entanglement in 2, 4, or 6 photons [253, 254]. In §18, we have discussed how entanglement swapping has been used to create quantum correlations between remote ions with the help of photons (in Figure 6.1, particles 1 and 4 are then ions, particles 2 and 3 are photons), leading to Bell inequality violations that are closer and closer to a loophole-free experiment.

6.3.3 Decoherence

We now discuss more precisely a phenomenon that we have already introduced, decoherence, which acts to reduce the lifetime of coherent superpositions. The process is extremely efficient if the number of correlated particles is large.

We have introduced decoherence in §2.1 as the initial part of the phenomenon associated with the Von Neumann infinite regress: coherent superpositions tend to constantly propagate towards the environment, involving more and more complex correlations with it. During decoherence, entanglement does not really disappear, but goes further and further into the environment; in practice, it becomes rapidly completely impossible to detect.

6.3.3.a Mechanism

To see more precisely how this happens, let us for instance consider a state:

$$|\Psi\rangle = \left[\alpha |\varphi_a\rangle + \beta |\varphi_b\rangle \right] \otimes |\mathbf{k}_0\rangle \quad (6.35)$$

which is the product of two states: one describes an atom in a coherent superposition of two orthogonal states $|\varphi_a\rangle$ and $|\varphi_b\rangle$, localized in two different regions of space; the other describes another particle, a photon for instance, initially in state $|\mathbf{k}_0\rangle$ (we assume that all these states are normalized).

Initially, the atom is described by a state with quantum properties that depend on the relative phase of α and β , which are therefore coherent; decoherence is a process where these coherent properties will disappear. This happens when the

photon interacts with the atom and is scattered into a quantum state that depends on the location where scattering took place: if the scattering atom is in state $|\varphi_a\rangle$, the photon reaches state $|\mathbf{k}_+\rangle$; if it is in state $|\varphi_b\rangle$, it reaches state $|\mathbf{k}_-\rangle$ ¹⁸ (states $|\mathbf{k}_\pm\rangle$ are normalized). After scattering, the atom has not changed its position, and returns to the same internal ground state (assuming that momentum transfer is negligible), so that the state vector describing the whole system after interaction becomes:

$$|\Psi'\rangle = \alpha |\varphi_a\rangle \otimes |\mathbf{k}_+\rangle + \beta |\varphi_b\rangle \otimes |\mathbf{k}_-\rangle \quad (6.36)$$

Assume now that we are interested only in the atom, not the photon; the reason might be, for instance, that detecting the scattered photon is impossible or very difficult (e.g. it may be a far-infrared photon). It is then useful to calculate the partial trace (§11.2.3) over this photon, in order to obtain the density operator that describes the atom only. A calculation similar to that leading to (6.16) shows that this partial trace can be written, in the basis of the two states $|\varphi_a\rangle$ and $|\varphi_b\rangle$:

$$\rho = \begin{pmatrix} |\alpha|^2 & \alpha\beta^* \langle \mathbf{k}_- | \mathbf{k}_+ \rangle \\ \alpha^*\beta \langle \mathbf{k}_+ | \mathbf{k}_- \rangle & |\beta|^2 \end{pmatrix} \quad (6.37)$$

If the scalar product $\langle \mathbf{k}_- | \mathbf{k}_+ \rangle$ was equal to 1, the density matrix of the atoms would not be affected at all by the scattering of the single photon. But this would be assuming that the photon is scattered exactly into the same state, independently of the spatial location in $|\varphi_a\rangle$ or $|\varphi_b\rangle$ of the scatterer; in other words, that it carries no information about the location of the atom! This is very unlikely if the distance between the two locations is much larger than the photon wavelength. Actually, it is much more realistic to assume that this scalar product is close to zero, which means that the off-diagonal element of (6.37) is also almost zero. We then conclude that the scattering of even a single photon destroys the coherence between atomic states, as soon as they are located in different places.

The loss of coherence becomes even worse when more and more photons (assumed to be all in the same initial state $|\mathbf{k}_0\rangle$) are scattered, since one should then replace (6.36) by state:

$$|\Psi''\rangle = \alpha |\varphi_a\rangle \otimes |\mathbf{k}_+\rangle |\mathbf{k}'_+\rangle |\mathbf{k}''_+\rangle \dots + \beta |\varphi_b\rangle \otimes |\mathbf{k}_-\rangle |\mathbf{k}'_-\rangle |\mathbf{k}''_-\rangle \dots \quad (6.38)$$

with obvious notation (the states with n primes correspond to the $(n-1)$ th scattered photon); the same calculation as above then provides the following value for the

¹⁸ We could also have assumed that the photon is focussed so that it can interact only with atoms in one of the two states, but is not scattered by the other state; the conclusions of our discussion would not have been changed.

partial trace ρ :

$$\begin{pmatrix} |\alpha|^2 & \alpha\beta^* \langle \mathbf{k}_- | \mathbf{k}_+ \rangle \langle \mathbf{k}'_- | \mathbf{k}'_+ \rangle \dots \\ \alpha^* \beta \langle \mathbf{k}_+ | \mathbf{k}_- \rangle \langle \mathbf{k}'_+ | \mathbf{k}'_- \rangle \dots & |\beta|^2 \end{pmatrix} \quad (6.39)$$

Since we now have, in the off-diagonal elements, the product of many scalar products $\langle \mathbf{k}_- | \mathbf{k}_+ \rangle$, it is clear that these elements have even smaller modulus than when a single photon is scattered. Actually, as soon as the two states $|\mathbf{k}_+\rangle$ and $|\mathbf{k}_-\rangle$ are not strictly identical, the off-diagonal elements tend exponentially to zero when the number of scattering events increases.

This is a completely general property: all objects have a strong tendency to leave a trace in the environment because they develop correlations with any elementary particle which passes by, so that they lose their coherence in the process. To illustrate this property, one can for instance in (6.35) replace the coherent state of a single atom by a GHZ type state (coherent state of a large number of atoms) and write the initial state as:

$$|\Psi\rangle = \left[\alpha |\varphi_a, \varphi_a, \dots, \varphi_a\rangle + \beta |\varphi_b, \varphi_b, \dots, \varphi_b\rangle \right] \otimes |\mathbf{k}_0\rangle \quad (6.40)$$

The larger the number of scattering atoms initially in states $|\varphi_a\rangle$ and $|\varphi_b\rangle$, the sooner many photons will be scattered, and the faster the off-diagonal elements will tend to zero. Nevertheless, the scattering of a single photon into two orthogonal states is already sufficient to completely destroy the coherence. Consequently, the large size of a physical system in a coherent superposition makes the superposition extremely fragile. The coherence actually propagates into a coherence involving the environment and more and more complex correlations with it (the scattered photon, in turn, may correlate with other particles); soon the coherence becomes practically impossible to detect. The phenomenon is unavoidable, unless of course the scattering properties of both states symbolized by $|\varphi_a\rangle$ and $|\varphi_b\rangle$ are exactly the same, which excludes any significant spatial separation between the states. This illustrates how fragile coherent superpositions of macroscopic objects are, as soon as they involve states that can be seen as distinct.

Nevertheless, it has been pointed out that the coupling of a quantum system with an environment can also, in some cases, be used to “engineer” the master equation that gives the time evolution of the density operator of the system, and to drive it into given quantum states; the method is called “quantum reservoir engineering” [255–257]. Therefore, dissipative coupling to a reservoir does not necessarily lead to states without quantum coherence; in fact, controlled decoherence may be used to prepare interesting quantum states.

6.3.3.b Revisiting the Schrödinger cat

We are now in a position where we can come back in more detail to some questions that we already discussed above in §2.2, and which are related to decoherence and/or the Schrödinger cat. Obviously, a cat that is alive or dead will not scatter all photons in the same way, otherwise we could not even see the difference! Clearly, decoherence will take place almost immediately. Does this simple remark provide an obvious solution to the paradox? More generally, can we regard decoherence (as some authors have done) as an “explanation” of the postulate of state vector reduction: when the superposition of the initial system becomes incoherent, are we not in the presence of a statistical mixture that resembles the description of a classical object with well-defined (but unknown) properties?

The answer to this question was already given in §2: this explanation is unsatisfactory because the purpose of the postulate of state vector reduction is not to explain decoherence, already contained in the Schrödinger equation, but the uniqueness of the result of the measurement, which is beyond this equation. In fact, the effect of the state vector reduction may sometimes be just the contrary: it puts back the measured sub-system into a pure state, the perfect opposite of a statistical mixture, so that the real question is then to understand how the (re)emergence of a pure state should be possible [49], not of a statistical mixture. Indeed, in common life, as well as in laboratories, one never observes superposition of results; Nature seems to operate in such a way that a single result always emerges from a single experiment. How can something that was indeterminate become determinate, and where does this disturbance come from? This will never be explained by the Schrödinger equation, since all that it can do is to endlessly extend its ramifications into the environment, without ever selecting one of them.

Another way to answer the question is to emphasize its logical structure. The central point of the paradox is the necessity for some kind of limit of the validity of the linear Schrödinger equation, since a linear equation can never predict the emergence of a single result in an experiment. The difficulty is where and how to create this border. Logically, it is clear that this problem will never be solved by invoking any process that is already contained inside the linear Schrödinger equation¹⁹, such as decoherence or any other similar process that remains linear. No one doubts that a typical measurement process will involve decoherence at some preliminary stage, but the real questions are what happens after, and how does a single result emerge. Decoherence is only the preliminary stage of the measurement process itself, when the off-diagonal elements of the density matrix vanish; in a second step, all diagonal elements but one should vanish (emergence of uniqueness), and this is the process that is very difficult to explain.

¹⁹ Staying in the middle of a country, one never reaches its borders ...

Pressed to this point of the discussion, some physicists reply that one can always assume that, at some later stage, the superposition resolves into one of its branches only. This is of course true, but this would amount to first throwing a problem out through the door, and then letting it come back through the window! (see discussions in Chapter 2, for instance on the status of the state vector and the necessity to resolve the Wigner friend paradox). We have discussed in §1.2.3.a the fundamental differences between a classical probability and how probabilities appear in quantum mechanics. A more consistent attitude would then be to consider that the natural complement of decoherence is the Everett interpretation of quantum mechanics (§10.11). Indeed, this provides a consistent interpretation of quantum mechanics, where the emergence of a single result does not have to be explained, since it is assumed never to take place (the Schrödinger equation then has no limit of validity). But, of course, if one takes this point of view, one has to deal with all the intrinsic difficulties of the Everett interpretation, which we will discuss later (§10.11). A general discussion of the relations between decoherence and the measurement problem, and its role in the various interpretations of quantum mechanics, is given in [258].

Concerning terminology, we have already mentioned in §2.2 that, during the last few years, it has become rather frequent to use the words “Schrödinger cat” (SC) to characterize coherent states such as (5.9), even sometimes for small values of N (sometimes for a single ion!). This is a redefinition of the words, since the essential property of the original cat was precisely to have a macroscopic number of degrees of freedom, as opposed to the few radioactive atoms or ions. Let us then assume that someone succeeded in preparing an all-or-nothing state with a very large value of N – would that be a much better realization of the Schrödinger cat as meant by its inventor? To some extent, yes, since the cat can be seen as a symbol of a system of many particles that go to an orthogonal single-particle quantum state, when one goes from one component of the state vector to the other. Indeed, it is likely that many of the atoms of a cat take part in different chemical bonds if the cat is alive or dead, which puts them in a different quantum state. But it seems rather hard to invent a reason why every atom, every degree of freedom, should necessarily be in two orthogonal states in the original story, while this is the essential property of “all-or-nothing states”. In a sense, they do more than realizing a standard Schrödinger cat; the two concepts are related but remain somewhat different, even for large values of N .

From an experimental point of view, decoherence is an interesting physical phenomenon that is certainly worth studying in itself, as recent experiments have illustrated [259]. A result of these studies and of the related calculations, among others, is to specify the basis in the space of states that is relevant to the decoherence process, as a function of the coupling Hamiltonian, as well as the characteristic time

constants that are associated with it. One can reasonably expect that more experiments on decoherence will follow this initial breakthrough and provide a more detailed understanding of many aspects of the phenomenon.

6.3.4 Purification, distillation

Distributing entangled states among partners may play an important role in several processes, such as quantum cryptography or teleportation, which we will study in §7.2. But, to realize this distribution, it is not sufficient to have a source that emits pairs of particles in a strongly entangled state. The reason is that the particles must propagate towards Alice and Bob who, in many applications, may be in remote places; this propagation may then change the state of the pair, transforming it either into a different pure state or a statistical mixture, and reduce the entanglement. The problem that then arises is to find a method to restore the initial entanglement, and to obtain particles that return to a state of complete entanglement. For this purpose, the only possible actions are local operations: each partner, in his/her laboratory, can modify the property of his/her particles by applying appropriate evolutions (for instance, they can apply a magnetic field to locally rotate one spin, or a phase plate to change the polarization of a photon), or perform local measurements with results that can be communicated to the other partner by classical channels.

It has been remarked that this operation is indeed compatible with the rules of quantum mechanics [260, 261], provided one accepts to reduce the number of useful pairs available to Alice and Bob. The loss of quality of entanglement is, so to say, converted into a reduction of the efficiency of the source, but then one is sure that the remaining pairs are indeed in the requested quantum state. This process is possible only if the initial loss of quality of the entanglement is not too strong. It conserves the entropy of entanglement, inasmuch as the local Hamiltonian operations performed by Alice and Bob do not change the Von Neumann entropies of their respective sub-systems. In fact, the efficiency of the process (number of pairs after the entanglement concentration operation divided by the total number of received pairs) is merely given by this entropy. The method can actually be used to produce any state of the two particles, provided the two partners can exchange information through a classical channel and act accordingly. It has a deep relation with the protocols for quantum correction of errors, an essential component of quantum computation (§7.4) [262]

Experimentally, this method has been successfully tested [263], nevertheless with a rate for success that is smaller than the maximal rate in theory; the involved protocols are difficult to implement, so that making some compromise is necessary.

6.4 Quantum dynamics of a sub-system

We now study the evolution of a test system S interacting with another (possibly much larger) quantum system B . Under general conditions, we will see that one can obtain a “master equation” giving the time evolution of the partial density operator ρ_S of S . This is in particular the case when the coupling between S and B is sufficiently weak, and for instance when B is a large system playing the role of a thermal reservoir. The master equation gives more quantitative information on the process of decoherence, generally called “transverse relaxation” in this context. It also contains other types of evolution such as “longitudinal relaxation” or “effective Hamiltonian” due to the coupling with B . In order to introduce a master equation in a general and convenient way, it is useful to first introduce the notion of Kraus operators.

6.4.1 Kraus operators

The space of states of S is an N_S -dimensional space $\mathcal{E}_S$ spanned by an orthonormal basis $\{|u_n\rangle\}$, with $n = 1, 2, \dots, N_S$. Initially, S is in state:

$$|\varphi_0\rangle = \sum_{n=1}^{N_S} c_n |u_n\rangle \quad (6.41)$$

while B is in a state $|\Phi_0\rangle \in \mathcal{E}_B$, where $\mathcal{E}_B$ is the space of states of B spanned by the N_B kets of an orthonormal basis $\{|\Theta_q\rangle\}$. We assume that the initial state $|\Psi\rangle$ of the whole system at $t = 0$ is a product:

$$|\Psi\rangle = |\varphi_0\rangle \otimes |\Phi_0\rangle \quad (6.42)$$

Between time $t = 0$ and $t = \tau$, the two systems interact and become entangled; we call $|\Psi'\rangle$ their state at time τ after interaction. We wish to obtain state $|\Psi'\rangle$ and the corresponding reduced density operator ρ'_S of system S .

6.4.1.a A first calculation

We begin with a simple calculation in order to define N_B operators M_q acting in the space of states of S , called Kraus operators. In the next section, we refine this definition and show how the number of Kraus operators can be limited in a way that depends only on the dimension N_S of $\mathcal{E}_S$, instead of N_B .

Assume that the initial state of S is one of the $|u_n\rangle$:

$$|\Psi\rangle = |u_n\rangle \otimes |\Phi_0\rangle \quad (6.43)$$

The corresponding final state $|\Psi'\rangle$ can then be expanded over the basis of the tensor products $|u_{n'}\rangle \otimes |\Theta_q\rangle$ as:

$$|\Psi'\rangle = \sum_{n'=1}^{N_S} \sum_{q=0}^{N_B-1} x_{n',q}^{(n)} |u_{n'}\rangle \otimes |\Theta_q\rangle \quad (6.44)$$

(for reasons that will become clear in §6.4.3, it is convenient to let the values of index q range from 0 to $N_B - 1$). We then have:

$$|\Psi'\rangle = \sum_{q=0}^{N_B-1} |\bar{u}_q^n\rangle \otimes |\Theta_q\rangle \quad (6.45)$$

with:

$$|\bar{u}_q^n\rangle = \sum_{n'=1}^{N_S} x_{n',q}^{(n)} |u_{n'}\rangle \quad (6.46)$$

The states $|\bar{u}_q^n\rangle$ of S appearing in this expansion are neither necessarily orthogonal nor normalized. We can introduce the N_B linear operators M_q acting in $\mathcal{E}_S$ defined by:

$$M_q |u_n\rangle = |\bar{u}_q^n\rangle \quad (6.47)$$

for any value of n . We then have:

$$|\Psi'\rangle = \sum_{q=0}^{N_B-1} M_q |u_n\rangle \otimes |\Theta_q\rangle \quad (6.48)$$

Assume now that the initial state of S is any state (6.41) in $\mathcal{E}_S$. By linearity, the state after interaction can be written as:

$$\begin{aligned} |\Psi'\rangle &= \sum_{n=1}^{N_S} c_n \sum_{q=0}^{N_B-1} M_q |u_n\rangle \otimes |\Theta_q\rangle \\ &= \sum_{q=0}^{N_B-1} M_q |\varphi_0\rangle \otimes |\Theta_q\rangle \end{aligned} \quad (6.49)$$

Since the unitary interaction between S and B does not change the norm of the state vector, this state is still normalized to 1. But by construction the $|\Theta_q\rangle$ are

orthonormal; we therefore have:

$$\sum_{q=0}^{N_B-1} \langle \varphi_0 | M_q^\dagger M_q | \varphi_0 \rangle = 1 \quad (6.50)$$

for any $|\varphi_0\rangle$. This can only be true²⁰ if:

$$\sum_{q=0}^{N_B-1} M_q^\dagger M_q = 1 \quad (6.51)$$

6.4.1.b An upper limit of the number of Kraus operators

The summations over q in the preceding expressions contain N_B terms; this situation may become intractable when N_B is very large (for instance if B is a macroscopic thermal reservoir). We now show that these summations can actually be limited to a maximum of $(N_S)^2$ terms, that is the square of the dimension of the space of state of S .

We begin with a simple lemma. Suppose that a state of the whole system $S + B$ can be written as:

$$|\Psi\rangle = \sum_s \lambda_s |\varphi_s\rangle \otimes |\Phi_s\rangle \quad (6.52)$$

where the summation contains an arbitrary number of terms, and where all kets $|\Phi_s\rangle$ belong to the same subspace $\mathcal{E}'_B$ of $\mathcal{E}_B$ with dimension $N'_B \leq N_B$. The constants λ_s and the kets $|\varphi_s\rangle \in \mathcal{E}_S$ can have any value. Then $|\Psi\rangle$ can always be written as a sum of N'_B terms as:

$$|\Psi\rangle = \sum_{q=1}^{N'_B} |\theta_q\rangle \otimes |\Xi_q\rangle \quad (6.53)$$

where the $|\Xi_q\rangle$ form an orthonormal basis in subspace $\mathcal{E}'_B$; the kets $|\theta_q\rangle \in \mathcal{E}_S$ are neither necessarily normalized nor mutually orthogonal in general. Obtaining (6.53) from (6.52) is straightforward since it is sufficient to replace all kets $|\Phi_s\rangle$ by their expansion on any orthonormal basis $\{|\Xi_q\rangle\}$, and to group all kets of $\mathcal{E}_S$ appearing in a product with a given $|\Xi_q\rangle$ into their sum $|\theta_q\rangle$.

²⁰ Operator $\sum_{q=0}^{N_B-1} M_q^\dagger M_q$ is Hermitian, and can therefore be diagonalized. Relation (6.50) then shows that all its eigenvalues are necessarily equal to 1, which means that the operator himself is equal to the identity operator acting in $\mathcal{E}_S$.

Consider now an initial state of the whole system where S is initially in one of the states $|u_n\rangle$:

$$|\Psi\rangle = |u_n\rangle \otimes |\Phi_0\rangle \quad (6.54)$$

The associated state $|\Psi'\rangle_n$ after interaction can be written in the form of a Schmidt decomposition with N_S terms (we have seen in §6.2.1 that such a decomposition cannot have more terms than the smallest dimension of the two spaces of states, and we assume $N_B \geq N_S$):

$$|\Psi'\rangle_n = \sum_{i=1}^{N_S} \sqrt{q_i} |\varphi_i^n\rangle \otimes |\Xi_i^n\rangle = \sum_{i=1}^{N_S} |\theta_i^n\rangle \otimes |\Xi_i^n\rangle \quad (6.55)$$

where the $|\Xi_i^n\rangle$ form an orthonormal ensemble spanning a subspace $\mathcal{E}_B^n$ of $\mathcal{E}_B$ with dimension N_S , and where:

$$|\theta_i^n\rangle = \sqrt{q_i} |\varphi_i^n\rangle \quad (6.56)$$

(some of the $|\theta_i^n\rangle$ vanish if some of the q_i are zero).

Now assume that the initial state of S is a linear superposition of two components:

$$|\Psi\rangle = \left[c_n |u_n\rangle + c_p |u_p\rangle \right] \otimes |\Phi_0\rangle \quad (6.57)$$

By linearity, the final state $|\Psi'\rangle_{np}$ can be written:

$$|\Psi'\rangle_{np} = c_n |\Psi'\rangle_n + c_p |\Psi'\rangle_p \quad (6.58)$$

where $|\Psi'\rangle_n$ is given by (6.55) and $|\Psi'\rangle_p$ by a similar relation with p replacing n . This ket has the form (6.52) corresponding to a subspace $\mathcal{E}_B^n \oplus \mathcal{E}_B^p$ of $\mathcal{E}_B$, obtained by direct sum of $\mathcal{E}_B^n$ and $\mathcal{E}_B^p$. The dimension N_{np} of this direct sum is equal to, or smaller than, the sum $2N_S$ of their dimensions (this maximum value $2N_S$ can be reached only if $N_B \geq 2N_S$ and if all kets $|\Xi_i^n\rangle$ and $|\Xi_i^p\rangle$ are independent). The final state can therefore be written as:

$$|\Psi'\rangle_{np} = \sum_{j=1}^{N_{np}} |\theta_j\rangle \otimes |\Xi_j^{np}\rangle \quad (6.59)$$

where the $|\Xi_j^{np}\rangle$ form an orthonormal basis in $\mathcal{E}_B^n \oplus \mathcal{E}_B^p$.

The reasoning can be continued by recurrence and applied to an initial state (6.42) containing the general expansion (6.41) of $|\varphi_0\rangle$. The final state of the whole

system will then be:

$$|\Psi'\rangle = \sum_{k=1}^{N'} |\theta_k\rangle \otimes |\Xi_k\rangle \quad (6.60)$$

where N' is the dimension of the subspace $\mathcal{E}_B^{n=1} \oplus \mathcal{E}_B^{n=2} \oplus \dots \oplus \mathcal{E}_B^{n=N_S}$ of $\mathcal{E}_B$ (with $N' \leq N_S \times N_S$), and where the $|\Xi_k\rangle$ provide an orthonormal basis in this subspace.

If we apply this relation to an initial state (6.54), using basis $\{|\Xi_k\rangle\}$ we obtain:

$$|\Psi\rangle = |u_n\rangle \otimes |\Phi_0\rangle \Rightarrow |\Psi'\rangle_n = \sum_{k=1}^{N'} |\bar{u}_k^n\rangle \otimes |\Xi_k\rangle \quad (6.61)$$

Changing the dummy index k into q , and still defining the Kraus operators M_q by (6.47), we then obtain:

$$|\Psi'\rangle_n = \sum_{q=0}^{N'-1} M_q |u_n\rangle \otimes |\Xi_q\rangle \quad (6.62)$$

By linearity, for any initial state $|\varphi_0\rangle$ of S , we then have:

$$|\Psi'\rangle = \sum_{q=0}^{N'-1} M_q |\varphi_0\rangle \otimes |\Xi_q\rangle \quad (6.63)$$

We have therefore recovered exactly the same expansions as above, but with a different range of summation, which may be much smaller if the dimension of the space of states $\mathcal{E}_B$ is very large. The number of terms in the sum is $N' \leq (N_S)^2$, so that $(N_S)^2$ gives an upper limit for the number of relevant Kraus operators associated with a given initial state $|\Phi_0\rangle$ (of course, one can also decide to always write sums with $(N_S)^2$ terms exactly, but then some Kraus operators M_q may vanish). The unitarity relation (6.51) still holds, for the same reasons as above. Note also that the Kraus operators are not uniquely defined: there is a large flexibility for the choice of the basis $|\Xi_q\rangle$ in the subspace of dimension N' , which may result in different M_q operators.

6.4.2 Density operator, Kraus sum

The density operator of system S after the interaction is given by a trace over the states of B :

$$\rho'_S = \text{Tr}_B \{ |\Psi'\rangle \langle \Psi'| \} \quad (6.64)$$

where $|\Psi'\rangle$ is given by (6.63). Since the $|\Xi_q\rangle$ are orthonormal, we have:

$$\rho'_S = \sum_{q=0}^{N'-1} M_q |\varphi_0\rangle \langle \varphi_0| M_q^\dagger \quad (6.65)$$

Therefore, the density operator of S after interaction can be expressed as a function of its value $\rho_0 = |\varphi_0\rangle \langle \varphi_0|$ before entanglement as:

$$\rho'_S = \sum_{q=0}^{N'-1} M_q \rho_0 M_q^\dagger \quad (6.66)$$

where the Kraus operators obey the normalization condition (6.51) and where $N' \leq (N_S)^2$. The right-hand side of (6.66) is called a “Kraus sum”; it provides the general expression of the partial density operator of S at the end of the process.

The Kraus operators therefore characterize, not only the way the state of the whole system has evolved from an initial product state (6.42), but also the evolution of the density operator ρ_S of the test system S . If, initially, the state of system S is defined by a density operator that is a statistical mixture and not a pure state, a simple linear superposition shows that relation (6.66) is still valid with the same operators M_q . If, initially, system B is also defined by a statistical mixture, a linear superposition can still be used. Nevertheless, since the definition of the Kraus operators depend on the initial state $|\Phi_0\rangle$, adequate averages should be introduced with the square roots of the probabilities associated with the initial states of B . Physically, it is not surprising that the Kraus operators M_q , which give the evolution of the partial density operator ρ_S , should depend on the state of the physical system B to which S is coupled; for instance, if B is a thermal bath, the evolution of ρ_S depends on the temperature of this bath.

6.4.3 Master equation, Lindblad form

We now assume that the interaction between the two systems is weak. The evolution operator $U(0, \tau)$ is then not very different from 1, and $|\Psi'\rangle$ not very different from $|\Psi\rangle$. If we choose a basis $\{|\Theta_q\rangle\}$ having $|\Phi_0\rangle$ as its first vector, we have:

$$|\Psi'\rangle \simeq |\Psi\rangle = |\varphi_0\rangle \otimes |\Theta_0\rangle \quad (6.67)$$

and the contribution of value $q = 0$ is then dominant in (6.49), with $M_0 \simeq 1$. We can write M_0 as:

$$M_0 = 1 + J + iK \quad (6.68)$$

with the following definitions of J and K :

$$J = \frac{M_0 + M_0^\dagger}{2} - 1 \quad K = i \frac{M_0^\dagger - M_0}{2} \quad (6.69)$$

They are both Hermitian operators, with small values if the interaction is weak. To first order in this interaction, we have:

$$M_0^\dagger M_0 = 1 + 2J + \dots \quad (6.70)$$

and the contribution of the term $q = 0$ in the right-hand side of (6.66) is:

$$\rho_0 + [J, \rho_0]_+ + i[K, \rho_0] + \dots \quad (6.71)$$

where $[C, D]_+$ is the anti-commutator $CD + DC$ and $[C, D]$ the commutator $CD - DC$.

If we insert (6.70) into (6.51), we obtain:

$$1 + 2J + \sum_{q=1}^{N'-1} M_q^\dagger M_q = 1 \quad (6.72)$$

or:

$$J = -\frac{1}{2} \sum_{q=1}^{N'-1} M_q^\dagger M_q \quad (6.73)$$

To first order, we can then write (6.66) in the form:

$$\rho'_S - \rho_0 = i[K, \rho_0] + \sum_{q=1}^{N'-1} \left(M_q \rho_0 M_q^\dagger - \frac{1}{2} M_q^\dagger M_q \rho_0 - \frac{1}{2} \rho_0 M_q^\dagger M_q \right) \quad (6.74)$$

where $\rho_0 = |\varphi_0\rangle \langle \varphi_0|$ is the initial density operator of S .

Now we assume that, during time dt , system S interacts in succession with ndt identical systems B , all identical and initially in the same state $|\Phi_0\rangle$ (n has the dimension of an inverse time). We then obtain the variation $d\rho_S$ of ρ_S during this time in the form:

$$\frac{d\rho_S}{dt} = \frac{-i}{\hbar} [H_{eff}, \rho_S] + \sum_{q=1}^{N'-1} \left(L_q \rho_S L_q^\dagger - \frac{1}{2} L_q^\dagger L_q \rho_S - \frac{1}{2} \rho_S L_q^\dagger L_q \right) \quad (6.75)$$

where H_{eff} is the “effective Hamiltonian” acting on S :

$$H_{eff} = -\hbar n K \quad (6.76)$$

and L_q the “Lindblad operator”:

$$L_q = \sqrt{n} M_q \quad (6.77)$$

Differential equation (6.75) is called a “master equation” for the evolution of the partial density operator ρ_S . The right-hand side term has a general form called a “Lindblad form”. The number of Lindblad operators can exceed neither the square of the dimension of the space of states of the test system, nor the dimension of the space of states of the coupled system B .

This type of master equation occurs in many physical situations. Its derivation can be generalized in several ways:

- (i) If S is initially described by a density operator ρ_0 corresponding to a statistical mixture, instead of a pure state, this operator can be expanded as:

$$\rho_0 = \sum_m p_m |\varphi_0^m\rangle\langle\varphi_0^m| \quad (6.78)$$

The above reasoning applies to each component $|\varphi_0^m\rangle\langle\varphi_0^m|$ of this operator and leads to (6.74) with ρ_0 replaced by $|\varphi_0^m\rangle\langle\varphi_0^m|$. Linearity then implies that equations (6.74) and (6.75) are still valid.

- (ii) If B is also described by a density operator corresponding to a statistical mixture and if the initial density operator of the whole system is the product of this operator by ρ_0 , the situation is different since the M_q operators depend on the initial state $|\Phi_0\rangle$ of B . Nevertheless, Lindblad forms can still be obtained in this case (see for instance §4.3 of [264]).
- (iii) We have assumed that system S interacts in succession and for a short time with many other systems B . A more frequent situation is that S constantly (but weakly) interacts with a single system B , for instance a heat bath. In such a case, one can show that similar pilot equations can be obtained, provided the correlation times associated with B are sufficiently short. Physically, it is plausible that a big system with short correlation times should be equivalent to a series of independently prepared systems. We refer for instance to Chapter IV of [265] for more details.
- (iv) The proper evolution in time of systems S and B can be included in the equations.

7

Applications of quantum entanglement

Quantum entanglement does not only provide a field of fundamental studies, but can also be harnessed as a tool for applications. In this chapter, after introducing two general theorems that are useful in the context of this discussion (§7.1), we propose a few examples: quantum cryptography (§7.2), teleportation (§7.3), and quantum computing (§7.4). For the moment, only quantum cryptography has given rise to real applications, and has been used in practical (and even commercial) applications; its purpose is the sharing of cryptographic keys between several partners by using a protocol where privacy is guaranteed by fundamental laws of physics. As for quantum computation, it is based on the general manipulation of quantum information, and is probably a more futuristic field of research in terms of applications, but it is certainly a domain of intense activity throughout the world.

Strictly speaking, none of these subjects in itself brings a really new view on the interpretation of quantum mechanics. Nevertheless, in addition to their strong intrinsic interests, they provide very direct and particularly interesting applications of its basic principles. This is the reason we study them in this chapter. We will only summarize the main ideas; the interested reader is invited to read the proposed references.

7.1 Two theorems

The two theorems that follow are similar; one deals with the creation and duplication of quantum states, the other with their determination.

7.1.1 *No-cloning*

The duplication of a quantum state, often called “quantum cloning”, is an operation where one starts from one particle in an arbitrary quantum state and reaches a situation where two particles are in the same state $|\varphi\rangle$. However, it turns out that,

within the laws of quantum mechanics, this operation is fundamentally impossible, a property often called the “no-cloning theorem” [266, 267]. The proof involves the linearity and the unitarity of the evolution of the state vector (the norm and the scalar product between states is conserved during the evolution).

We assume that the complete system contains:

- the “source”, initially in an arbitrary state $|\varphi\rangle$
- the “target”, a physical system having the same space of states than the source (or an isomorphic space of states). Initially, the target is in a normalized state $|\xi_0\rangle$ and should be transferred into the same state $|\varphi\rangle$ as the source.
- an environment, with any space of state, and which is initially in the normalized state $|\Phi_0\rangle$.

We wish to study what kind of evolution can transform the initial state:

$$|\Psi_i\rangle = |\varphi\rangle \otimes |\xi_0\rangle \otimes |\Phi_0\rangle \quad (7.1)$$

into a final state $|\Psi_f\rangle$ given by:

$$|\Psi_f\rangle = |\varphi\rangle \otimes |\varphi\rangle \otimes |\Phi_f(\varphi)\rangle \quad (7.2)$$

We now consider two different values $|\varphi_1\rangle$ and $|\varphi_2\rangle$ of $|\varphi\rangle$, associated to initial states $|\Psi_i^{(1)}\rangle$ and $|\Psi_i^{(2)}\rangle$ for the whole system; according to (7.1), the scalar product of these two states is:

$$\langle \Psi_i^{(1)} | \Psi_i^{(2)} \rangle = \langle \varphi_1 | \varphi_2 \rangle \quad (7.3)$$

After evolution, the product scalar should become, according to (7.2):

$$\langle \Psi_f^{(1)} | \Psi_f^{(2)} \rangle = [\langle \varphi_1 | \varphi_2 \rangle]^2 \langle \Phi_f(\varphi_1) | \Phi_f(\varphi_2) \rangle \quad (7.4)$$

This can be equal to (7.3) only in two cases:

- either $\langle \varphi_1 | \varphi_2 \rangle = 0$,
- or:

$$\langle \varphi_1 | \varphi_2 \rangle \times \langle \Phi_f(\varphi_1) | \Phi_f(\varphi_2) \rangle = 1 \quad (7.5)$$

If this second possibility occurs, since all vectors are normalized, each of the scalar products in this expression has a modulus that is less than, or equal to, 1 (Schwarz inequality); the equality necessarily implies that the vectors themselves are equal (within an irrelevant phase factor). Relation (7.5) therefore implies at the same time that $|\varphi_1\rangle = |\varphi_2\rangle$ and $|\Phi_f(\varphi_1)\rangle = |\Phi_f(\varphi_2)\rangle$, while we have assumed above that states $|\varphi_1\rangle$ and $|\varphi_2\rangle$ are different. We then reach a contradiction, and the only remaining

possibility is the first: the scalar product $\langle \varphi_1 | \varphi_2 \rangle$ should vanish. Unitarity therefore requires that, if state $|\varphi_1\rangle$ can be cloned, the only other states $|\varphi_2\rangle$ that can be cloned are orthogonal; it is impossible to clone an arbitrary linear combinations of source states.

Suppose that we release one of the assumptions, namely that state $|\varphi\rangle$ is invariant in the whole process; can we then make cloning possible? We now assume that, instead of putting two systems into the initial state $|\varphi\rangle$ of the source, the process puts them both into another state $|\bar{\varphi}\rangle$ that is a function of $|\varphi\rangle$ given by:

$$|\bar{\varphi}\rangle = \bar{U} |\varphi\rangle \quad (7.6)$$

where $\bar{U}$ is a unitary operator. Equation (7.2) is then replaced by:

$$|\Psi_f\rangle = |\bar{\varphi}\rangle \otimes |\bar{\varphi}\rangle \otimes |\Phi_f(\varphi)\rangle \quad (7.7)$$

After all, this would also be a useful sort of cloning, since from the knowledge of $|\bar{\varphi}\rangle$ one could infer that of $|\varphi\rangle$. But this process also is forbidden by the rules of quantum mechanics. To see why, it is sufficient to apply the same reasoning as above to the unitary operator obtained by multiplying the evolution operator by the product $\bar{U}^{-1}(\text{source}) \times \bar{U}^{-1}(\text{target})$; one then obtains the same equations and reaches the same contradictions. Releasing the invariance of the state of the source therefore does not help.

If it is impossible to clone states exactly, is it possible at least to do it in some approximated way? We now show that this is impossible as well. Without writing the strict equality (7.2) for the final state, let us now assume a weaker condition:

$$|\Psi_f\rangle = |\bar{\varphi}\rangle \otimes |\tilde{\varphi}\rangle \otimes |\Phi_f(\varphi)\rangle \quad (7.8)$$

where $|\bar{\varphi}\rangle$ and $|\tilde{\varphi}\rangle$ are good approximations of the initial target state $|\varphi\rangle$. The conservation rule of the scalar product then gives:

$$\langle \varphi_1 | \varphi_2 \rangle = \langle \bar{\varphi}_1 | \bar{\varphi}_2 \rangle \langle \tilde{\varphi}_1 | \tilde{\varphi}_2 \rangle \langle \Phi_f(\varphi_1) | \Phi_f(\varphi_2) \rangle \quad (7.9)$$

Since $\langle \bar{\varphi}_1 | \bar{\varphi}_2 \rangle \simeq \langle \varphi_1 | \varphi_2 \rangle$, we necessarily have

- either $\langle \varphi_1 | \varphi_2 \rangle = 0$,
- or:

$$\langle \tilde{\varphi}_1 | \tilde{\varphi}_2 \rangle \times \langle \Phi_f(\varphi_1) | \Phi_f(\varphi_2) \rangle \simeq 1 \quad (7.10)$$

which implies that $\langle \tilde{\varphi}_1 | \tilde{\varphi}_2 \rangle \simeq 1$ and, by symmetry, that $\langle \bar{\varphi}_1 | \bar{\varphi}_2 \rangle \simeq 1$. As a consequence, the states that can be approximately cloned are necessarily either almost orthogonal, or almost identical; again, arbitrary linear combinations cannot be cloned.

In Appendix F, we discuss why, if state cloning was possible in an EPR experiment, Bob could apply it to the particle he receives after Alice has made her measurement along direction a . Since his particle is also polarized along this direction, he could, if he had multiple copies of it, determine the polarization with accuracy and know which direction a Alice has chosen. This scheme could be used to transmit information about the orientations a without any minimal delay proportional to the distance, in contradiction with relativity. In this case, the consistency of physics is guaranteed by the no-cloning theorem; see [268] for a historical discussion of this theorem, and [269] for a study of multiple cloning.

7.1.2 No single shot state determination

A similar theorem is the following: given a single quantum system in a state $|\varphi\rangle$, it is impossible to determine $|\varphi\rangle$ exactly with any sequence of measurements. The physical reason is that, whatever first measurement is performed on the system, if some information is indeed obtained, at the same time the state is modified (state vector reduction). Further measurements will then have access only to this modified state, which make the accurate determination of the initial state $|\varphi\rangle$ impossible. This theorem is necessary to ensure consistency with the no-cloning theorem: if it was possible to determine the quantum state with arbitrary accuracy, one could then construct a filter (Stern–Gerlach system for spins) to put an arbitrary number of particles into the same state $|\varphi\rangle$, in violation with the no-cloning theorem.

The no-determination theorem is valid only when a single realization of the quantum system is given: if many copies of the system in state $|\varphi\rangle$ are available, then it becomes possible to determine this state with an accuracy that is an increasing function of the number of copies. Different methods have been proposed, in particular “quantum tomography”, which relies on the successive measurements of incompatible observables with several realizations of the same quantum state, and a mathematical reconstruction of the most probable initial quantum state [270–272]; it has been used in many experiments. See also the discussion of Chapter 15 of [273] and the method of so-called “weak measurements” (§8.3.1) [274, 275], which has been used to determine the wave function of single photons with the appropriate use of small birefringent phase shifter, a Fourier filter, and polarization measurements [276].

A related interesting theoretical result is the following: if several copies of the initial system are available and described by the same density operator ρ (pure state or statistical mixture), it turns out that a single experimental setup is sufficient to determine the quantum state [277]. A condition is that a larger auxiliary quantum system is allowed to interact in a controlled way with the initial system; at the end of the process, one measures a factorized observable acting on the two systems.

The interaction with the auxiliary system (sometimes called “ancilla”) transforms the information related to non-commuting observables in the initial system into commutative information for the total quantum system.

7.2 Quantum cryptography

Quantum cryptography is not, as one could believe from these words, a method of cryptography to code or decode messages by quantum methods. The purpose of the method is only the transmission of a secret key between two partners through quantum measurements, the subsequent use of this key remaining classical. This explains why the more accurate name “quantum key distribution” is often used for it.

7.2.1 Sharing cryptographic keys

The method makes use of the specific properties of quantum systems to ensure the transmission of a cryptographic key between two partners, without any risk of a third person having access to the key by monitoring the transmission line. It has several variants, some making use of series of single-particle events [278] as in the BB84 protocol, some involving several entangled particles, typically in an EPR quantum state [279, 280]. For a general introduction, see [281]; for a review with more details, see [282].

The basic idea is to build a scheme that leads to a perfectly secure way to transmit a cryptographic key – such a key is a random sequence of 0 and 1, which is used to code, and then decode, a secret message. In a first step, the two remote correspondents Alice and Bob share this key, and then use it to code¹ and decode all messages they exchange. If the key is perfectly random and if each of its elements is used only once (which means that the key should be at least as long as the messages themselves), it becomes totally impossible to anyone to decode any message without knowing the key, even if the message is sent publicly. But the risk is that, during the initial process of key communication, a spy (traditionally named Eve, after “eavesdropping”) is able to have access to the communication line used by Alice and Bob in order to learn the key; from this moment, she will be able to decode all messages sent with this key. The exchange of keys is therefore a crucial step in cryptography. The usual strategy is to make the best possible use of the traditional methods of confidentiality: storing the key in a safe, secure transportation, etc., but

¹ The simplest method is to write the message with a binary coding of its characters, and then to perform a binary sum of each bit of the message with the corresponding bit of the key (the bit with the same rank); the result, which looks totally random and is therefore completely unreadable, is sent by Alice to Bob. He can then perform the same binary addition again and recover the initial message.

it is always difficult to assess the safety of such methods, where human factors play such an important role.

By contrast, the quantum key distribution relies on fundamental laws of physics, which are impossible to break: as clever and inventive spies may be, they will not be able to violate the laws of quantum mechanics! The basic idea is that Alice and Bob will create their common key by making quantum measurements, for instance on particles in an EPR correlated state; in this way they can generate series of random numbers that can be subsequently used as a secret communication key. What happens then if Eve tries to intercept the photons during the process of key creation? For example, she could couple some elaborate optical device to the optical fiber where the photons propagate between Alice and Bob, and then make measurements. If Eve just absorbs the photons she measures, this immediately changes the correlation properties observed by Alice and Bob, and the spying process can be detected. Of course, this does not necessarily stop the spying, but at least Alice and Bob know which data have been perturbed and can use only the others in the future as a perfectly safe cryptographic key.

7.2.2 Examples of protocols for key exchange

Quantum cryptography is now a big field of research; there is a large variety of schemes and protocols for quantum cryptography, some based on the use of EPR correlated particles [279], others not [280]. Some of the schemes have been implemented in practice, and provide the distribution of keys over distances exceeding 100 kilometers.

7.2.2.a BB84 protocol

We assume that Alice sends to Bob single photons, one by one, either in state $|H\rangle$ with horizontal polarization to signal a bit 0 of the key, or in state $|V\rangle$ with vertical polarization to signal a bit 1 (Figure 7.1). Bob, when he measures the polarization of the photons he receives, can reconstruct the key, but there is not guarantee that Eve has not been able to interfere and obtain the value of the bit. This is because, when Eve does the same polarization measurement as Bob in her laboratory along the transmission line, even if she has to absorb the photon, she can also re-send another photon with the polarization corresponding to the result of her measurement. Bob, when measuring the polarization of the photons he then receives from Eve, obtains exactly the same result as if Eve had not made any measurement; neither he nor Alice can know that they have been spied.

The BB84 (for Bennett and Brassard, 1984 [278]) protocol provides a way to suppress this risk. The protocol involves two different bases for the polarizations, that of the two states $|H\rangle$ and $|V\rangle$ (horizontal and vertical polarizations) and a

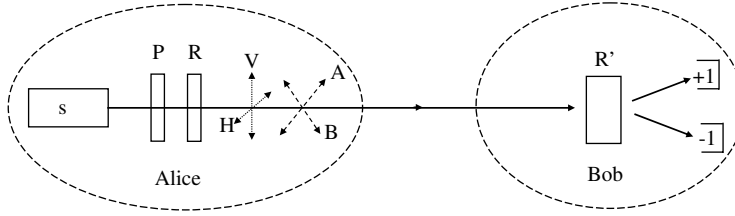


Figure 7.1. Scheme of the BB-84 protocol for secure exchange of a cryptographic key. Alice sends photons one by one to Bob and, for each of them, randomly chooses a polarization among four possible values: H (horizontal), V (vertical), A (first diagonal at 45°), and B (second diagonal at -45°). For this purpose, she uses two optical devices P and R that produce these four possible polarizations. By convention, H and A correspond to a bit 0, V and B to a bit 1. Bob, on his side, uses optical system R' to choose randomly along which pair of polarizations, (H,V) or (A,B), he will make each measurement, which provides result ± 1 . The extraction of the shared secret key from these results is discussed in the text and summarized Figure 7.2.

second base with two other states:

$$\begin{aligned} |A\rangle &= \frac{1}{\sqrt{2}} [|H\rangle + |V\rangle] \\ |B\rangle &= \frac{1}{\sqrt{2}} [|H\rangle - |V\rangle] \end{aligned} \quad (7.11)$$

These two states correspond to the two linear polarizations at $\pm 45^\circ$ with respect to polarizations $|H\rangle$ and $|V\rangle$. By convention, $|H\rangle$ and $|A\rangle$ are associated with a bit of the key equal to 0, while $|V\rangle$ and $|B\rangle$ are associated with a bit equal to 1 (this is just a possibility; any other convention associating 0 with one state of each basis, and 1 with the two others, would work equally well).

A crucial point is that, to send each bit of the key, Alice must choose in a completely random way to use either basis $\{|H\rangle, |V\rangle\}$ or basis $\{|A\rangle, |B\rangle\}$. On his side, Bob also chooses his direction of measurement completely randomly, so that there is a 50% chance that his basis is different from Alice's. Relations (7.11) then show that he can obtain his two results with the same probability $1/2$, independently of the bit sent by Alice, so that he receives no information. But if, by chance, he chooses the same basis, the information is contained in his measurement. The protocol therefore has two steps: first, Alice sends particles to Bob and, second, and only later, they communicate with each other publicly (through a classical channel, which does not have to be secret) to exchange information about what basis they used. Of course, they do not communicate any information on their results, otherwise privacy would be lost! This allows Bob to retain only the bits corresponding to cases where the two bases are the same (perfectly correlated results), and to

reject all the others. The secret key is then made of the retained results only, with a number of bits that is approximately half the number of particles sent by Alice. This method may look complicated, but we will see that it ensures that any attempt by Eve to get access to the data becomes detectable by Alice and Bob.

What can Eve do to have access to the key sent by Alice? While the particles are travelling from Alice to Bob, she may attempt to capture each of them and to perform a measurement of their polarizations; to remain unnoticed, each time she measures a particle, she re-sends towards Bob another particle with the polarization corresponding to the result she has obtained. But, before Alice and Bob communicate about what basis they choose, no information is available on this basis, so that Eve does not know if she must analyze the particle with basis $\{|H\rangle, |V\rangle\}$ or basis $\{|A\rangle, |B\rangle\}$. We have seen in §7.1 that she cannot determine the quantum state of a single particle, and moreover that the no-cloning theorem prevents her from duplicating the received particle into several in the same state. Actually, as Bob, Eve has one chance out of two to choose the wrong basis. If for instance Alice used the first basis and Eve the second, Eve will receive a state $|H\rangle$ or $|V\rangle$ but re-send a particle in one of the states $|A\rangle$ or $|B\rangle$, which changes the flux of particles received by Bob and creates anomalies. Since there is a chance in two that Alice will choose the wrong basis, and also a chance in two that this will change the result observed by Bob, the perturbations created by Alice introduce an anomaly in 25% of the cases.

The strategy used by Alice and Bob is therefore to sacrifice some of the “good” bits that they have obtained when their bases coincide, to exchange them publicly, and to check if they are always identical as expected (Figure 7.2). It does not matter if Eve learns the values of these bits, which will not be used anymore. Alice and Bob compare their results and, if they differ by roughly 25%, know that someone has been spying on their communication. Whatever Eve can do inevitably modifies the correlation properties of the photons measured at the two ends of the line, so that Alice and Bob can become aware of the problem – provided they take the trouble to carefully compare their data and correlation rates.

Of course, our discussion is simplified, and only gives the general principles. In practice, when Alice and Bob choose the same basis, the coincidence between their results is not perfect, even in the absence of Eve. The reason is that various perturbations may occur during the propagation of the particles and during the measurements. If the corresponding error rate does not exceed a given value, one can use the techniques associated with classical error correction codes (as those used when reading CD or computer hard disks, for instance parity controls); the size of the key is then reduced even more, but random errors are very efficiently reduced. For a discussion of various methods of “privacy amplification” and “quantum secret growing”, see [282].

Alice's polarizations	 V	 A	 V	 B	 B	 H	 B	 V
Bits sent	1	0	1	0	1	1	0	1
Bob's measurement polarisation basis								
Bob's results	1	0	0	0	1	1	0	1
Bits of common key	1	0		0		1		1

Figure 7.2. A typical run for a secret key exchange experiment. Alice sends her photons one by one by choosing randomly among four polarizations H, V, A, and B; then Bob measures the received polarization, either in the (H,V) basis or in the (A,B) basis. After the measurement is performed, they publicly communicate the choice of basis, either (H,V) or (A,B), for each pair, but not the results. The common key is then obtained from the results for which the two bases turned out to be the same; all results where the basis are different are ignored. In the case shown, only five bits are retained, but by repeating the experiment many times one can obtain a key of arbitrary length.

The general scheme works with any sort of quantum particles but, for practical reasons, they are often photons propagating inside optical fibres. A difficulty is that the fibres may change the polarizations of the photons and destroy the correlations used to build the key. To overcome this problem, Bennett [283] has proposed another technique where time resolution is used, and where both Alice and Bob use a Mach–Zehnder interferometer with adjustable phase shift in one arm. Each photon emitted in Alice's laboratory can reach Bob's detectors through several different paths, two of which interfere in the time pulse he measures. The quantum interference effect that then takes place is very similar to the interference between different polarizations in the previous scheme, and a similar protocol of quantum measurements (involving non-commuting operators) can be used to transmit secure keys. This scheme has been realized in an experiment in 1993 [284] and is still in use nowadays.

7.2.2.b EPR protocol

In the preceding protocol, the quantum particles carrying the information are sent by one of the two partners and received by the other. It is also possible to design protocols that are more symmetric, where both receive particles belonging to pairs emitted by a common source. A particularly interesting case occurs when these pairs are in an entangled state such as that used in the Bell theorem; this makes a connection between the completeness of quantum mechanics (or its non-locality),

as discussed by EPR, and the quantum distribution of keys [279]. To emphasize the similarity with the discussion of the Bell inequalities in §4.1, we now come back to a situation with two spin $1/2$ particles, initially in the singlet state written in (4.1); but the transposition in terms of vertical or tilted polarizations of photons is straightforward.

The protocol to be used in this case remains rather similar to the BB84 protocol; it also gives a central role to the random choice of basis, but this time both Alice and Bob make a completely random choice of the components of the spins they measure. One can assume that they limit themselves to two directions of measurement, either along axis Oz , or along a perpendicular axis Ox . As above, two situations may occur: either they make different choices and their results have no correlation; or they make the same choice and, for an ideal experiment, the correlations are perfect. In this case also, Alice and Bob communicate only after the reception and measurement of all particles, and only about the choice of directions they have made – never about the results obtained. Each partner only retains the cases where it happens that the two choices have coincided. With this protocol, often called “EPR protocol”, the cryptographic key that Alice and Bob extract from their measurements is completely random: none of them can choose the series of bits of the key (while, in the BB84 protocol, Alice may do it), but this difference has no effect on the subsequent transmission of messages.

The effects of Eve’s intrusions are similar to those in the BB84 protocol: since she does not know which base is used by Alice and Bob, any action from her implies a significant probability that it will change the correlations observed by the two partners (in the cases when they have chosen the same basis). More precisely, when Eve performs a measurement on a spin and finds a result, the other spin is projected onto an eigenstate of the spin component that she has chosen, with the opposite eigenvalue. If Eve sends on the line another spin having the polarization that she has measured, and if by chance Alice and Bob choose a direction of measurement that is parallel to Eve’s, they will observe a perfect anti-correlation of their results, exactly as if Eve had done nothing; her action is not visible. But if she chooses a perpendicular direction, there is one chance in two that they will observe the same result, which would be impossible in the absence of Eve’s perturbation; an error rate of 25% then appears. As in the BB84 protocol, the result is that, if Alice and Bob decide to sacrifice a fraction of their results to check that no perturbation has occurred, they can easily verify that their communication has remained confidential.

Moreover, they can use a broader range for the directions of measurement, in particular those leading to a quantum violation of the BCHSH inequalities [107]; this provides another test to check that no action has been performed on the spins between the source and the measurements. In fact, the direction chosen by Eve and the result she observes play a role that is very similar to an additional variable λ ,

which determines the polarization properties of the particles received by Alice and Bob. The average over λ then leads very naturally to the BCHSH inequalities (in fact, one is precisely in the case studied in Appendix C). Checking that significant violations of the inequalities are still obtained is therefore a good criterion for the absence of Eve's intermediate measurements; this illustrates an interesting relation between the Bell theorem and quantum cryptography.

Conceptually, the EPR protocol nevertheless remains rather different from BB84 since, at any moment when Eve may attempt to intrude into the system, the information that Alice and Bob will use to build the key still does not exist: all possible results of all spin components are still potentially present in the singlet entangled state, since no measurement has been performed – in other words, the useful information does not exist yet. If Eve attempts to acquire some knowledge, it is (so to say) she who takes the responsibility of projecting the state vector and to attribute definite polarizations to the spins; it is rather intuitive that this operation should be detectable by Alice and Bob.

One can go even further in this direction and conceive schemes where, for the events that are useful for creating the key, no particle at all propagates along the line between Alice and Bob where Eve could intervene. Such a scheme has already been discussed at the end of §2.4 – see also [60]. This particular protocol involves one single particle only, and constructs the key from events where the particle has remained in Alice's laboratory; it never propagated between Alice's and Bob's locations, because of a destructive quantum interference effect. This is an extreme case where, for all events that are retained for constructing the key, no particle has propagated between the two partners, which prevents Eve from obtaining any useful information.

Other protocols have been proposed. It is possible for instance to use protocols making use of six quantum states, or individual quantum systems having a space of states with dimension larger than 2; again, for more details, we refer the reader to the review article [282].

7.3 Teleporting a quantum state

The notion of quantum teleportation [285] is also related to quantum non-locality; the idea is to take advantage of the correlations between two entangled particles, which initially are for instance in an “all or nothing” state (5.9) (for $N = 2$), in order to reproduce at a distance the arbitrary spin state of a third particle. The scenario is the following (Figure 7.3): initially, two entangled particles propagate towards two remote regions of space; one of them reaches the laboratory of the first actor, Alice, while the second reaches that of the second actor, Bob; a third particle in an arbitrary quantum state $|\varphi\rangle$ is then provided to Alice in her laboratory; the final

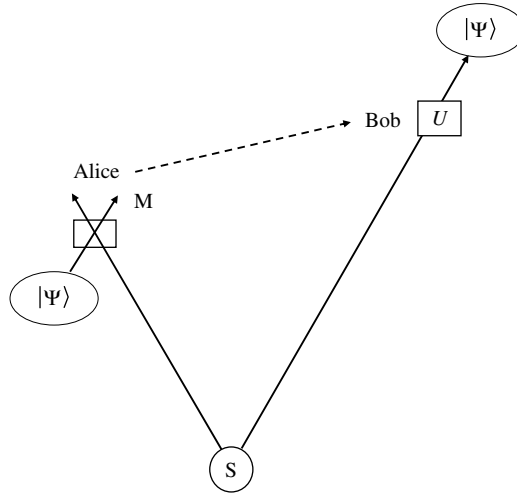


Figure 7.3. A source S emits two spin $1/2$ entangled particles, which fly to Alice and Bob. In addition, Alice receives a particle in an arbitrary state $|\Psi\rangle$, which she does not know. She performs a measurement M involving both this particle and one of the particles of the pair emitted by S . She then sends a message to Bob by a classical communication channel (shown by a broken line in the figure) containing the result of her experiment (two bits of classical information). Bob uses this information to perform a unitary transformation U on his particle which, then, is transferred into exactly the same state $|\Psi\rangle$ as the initial remote particle. This process is often called “quantum teleportation”.

purpose of all this scenario is to put Bob’s particle into exactly the same state $|\varphi\rangle$, whatever it is (without, of course, transporting the particle itself). One then says that state $|\varphi\rangle$ has been teleported.

What procedure can be followed to realize teleportation? One could naively think that the best strategy for Alice is to start by performing a measurement on the particle in state $|\varphi\rangle$ to be teleported. This is not true: one can show that it is more efficient to perform a “combined measurement” that involves at the same time this particle as well as her particle from the entangled pair. In fact, for the teleportation process to work, an essential feature of this measurement is that no distinction between the two particles used by Alice should be made. With photons one may for instance, as in [286], direct the particles onto opposite sides of the same optical beam splitter, and measure on each side how many photons are either reflected or transmitted; this device does not allow one to decide from which initial direction the detected photons came, so that the condition is fulfilled. Then, Alice communicates the result of the measurement to Bob; this is done by some classical method such as telephone, e-mail, etc., in other words by a method that is not instantaneous but submitted to the limitations due to the finite velocity of light. Finally, Bob modifies

the state of his particle by applying a local unitary transformation that depends on the classical information he has received. This operation puts his particle into exactly the same state $|\varphi\rangle$ as the initial state of the third particle, which realizes the “teleportation” of the state. The whole scenario is “mixed” because it involves a combination of transmission of quantum information (through the entangled state) and classical information (the phone call from Alice to Bob). Here we discuss only the general ideas, without explicitly writing the calculations that show how the measurement process modifies the state vector, and which unitary transformation Bob should apply, even if these calculations are not very complicated. A more detailed description can be found in the original reference [285], or for instance in §9.8 of [287], or again in §6.5.3 of [288].

Teleportation may look either magical, or trivial, depending how one looks at it. The possibility of reproducing at a distance a quantum state from classical information is not in itself a big surprise. Suppose for instance that Alice decided to choose what the teleported state should be, and filtered the spin (she sends particles through a Stern–Gerlach system until she gets a $+1$ result²); she could then ask Bob by telephone to align his Stern–Gerlach filter in the same direction, and to repeat the experiment until he also observes a $+1$ result. This might be called a trivial scenario of teleportation, based only on the transmission of classical information. But real quantum teleportation does much more than this! First, the state that is transported is not necessarily chosen by Alice, but can be completely arbitrary. Second, a message with only binary classical information is used (the result of the combined experiment made by Alice³), which certainly does not provide sufficient information to reconstruct a quantum state. In fact, a quantum state depends on continuous parameters, while results of experiments correspond to discrete information only. Somehow, in the teleportation process, the finite binary information has turned into continuous information! The latter, in classical information theory, would correspond to an infinite number of bits.

From Bob’s point of view, the amount of received information has two components: classical information sent by Alice, with a content that is completely “uncontrolled”, since it is not decided by her, but just describes the random result of an experiment; quantum information is contained in the teleported state itself (what we will call a “qubit” in the next section). If the teleportation is repeated many times with the same state prepared by Alice, by successive measurements on the teleported particles Bob will be able to determine its quantum state with

² For filtering a spin state, one obviously needs to use a non-destructive method for detection after the Stern–Gerlach magnet. One could for instance imagine a laser detection scheme, designed in such a way that the atom goes through an excited state, and then emits a photon by returning to the same internal ground state (closed optical pumping cycle – this is possible for well-chosen atomic transition and laser polarization).

³ Alice sends to Bob two classical bits corresponding to the result of a measurement performed on two particles.

arbitrary accuracy, including the direction that was chosen by Alice; he therefore receives a real message from her (for a discussion of the optimum strategy that Bob should apply, see [289]).

If one wishes to describe teleportation in a sensational way, one could explain that, even before Bob receives any classical information, he has already received “almost all the information” on the quantum state, in fact all the controllable information (the content of the classical message is random); this “information” has come to him instantaneously, exactly at the time when Alice performed her combined experiment, without any delay proportional to the distance covered. The rest of information, which is the “difference” between continuous “information” and discrete information, comes only later and is, of course, subject to the minimum delay associated with relativity. But this is based on an intuitive notion of “difference between quantum/controllable and classical/non-controllable information” that we have not precisely defined; needless to say, this should not be taken as a real violation of the basic principles of relativity!

Finally, has really “something” been transported in the teleportation scheme, or just information? Not everyone agrees on the answer to this question. In any case, what is perfectly clear is that the essence of the teleportation process is completely different from any scenario of classical communication between human beings. The relation between quantum teleportation and Bell-type non-locality experiments is discussed in [290]; see also [291] as well as [292] for a review of recent results, as well as [293] for a teleportation experiment of an N-photon GHZ state with the help of only one pair of entangled particles. Recently, teleportation of photon states over a distance of 16 kilometers in free space has been reported [294].

7.4 Quantum computation and information

What we said above of cryptography and teleportation is even more true concerning quantum computing. This field of research gives rise to a large activity, with new results constantly appearing in an abundant literature; giving an up-to-date report is out of the question here. We will thus limit ourselves to a brief introduction of some of the major ideas, and refer the interested reader who wishes to learn more to [295–299].

7.4.1 General idea

The general idea of quantum computing [300] is to base numerical calculations, not on classical “bits”, which can be only in two discrete states (corresponding to 0 and 1 in usual binary notation), but on quantum bits, or “qubits”. These qubits

are quantum systems, each having access to a two-dimensional space of states; this means that qubits can, not only be in one of the two states $|0\rangle$ and $|1\rangle$, but also in any linear superposition of them. For a single qubit, it is already clear that a continuum of states is a much “larger” ensemble than two discrete states only. For a large collection of bits, the difference is even more pronounced: for classical bits, the dimension of the state space increases linearly with the number of bits (for instance, the state of three classical bits defines a vector with three components, each equal to 0 or 1); for qubits, the dimension increases exponentially (this is a property of the tensor product of spaces; for instance, for three qubits the dimension of space is $2^3 = 8$). If one assumes that a significant number of qubits is available, one gets access to a space state with an enormous “size”, where many kinds of interference effects can take place.

Now, if one could somehow make all branches of the state vector “work in parallel” to perform independent calculations, it is clear that one could perform much faster calculations, at least in theory. For instance, suppose that one wishes to calculate the solutions of a system of equations that depend on a parameter; one could imagine an algorithm where the system of qubits goes to a superposition of states associated with all possible solutions of the equations, simultaneously for all values of the parameter. But then the difficulty is to have access to these components: one cannot directly measure the state vector as a classical variable. Nevertheless, appropriate quantum measurement processes can be designed to make some use of this “quantum parallelism”, which in fact opens up many possibilities. For instance, the notion of unique computational complexity of a given mathematical problem, which puts limits on the efficiency of classical computers, no longer applies in the same way.

A summary of the history and prehistory of quantum computing can be found for instance in [301]. Feynman, in a conference given in 1981 at MIT, remarked that it seems to be in general impossible to simulate the evolution of any quantum system on a classical computer with reasonable efficiency. This led him to propose a basic model for a quantum computer that would perform this task. In 1985, David Deutsch [300] described a “universal quantum computer”, or “quantum Turing machine”, that can simulate any other quantum computer, as a classical Turing machine can simulate any classical computer. The whole field, which is presently a domain of very active research, expanded rapidly in the decade 1990–2000 with the introduction of quantum gates and algorithms.

7.4.2 Quantum gates and algorithms

A “quantum computer” is generally considered as a combination of basic elements called “logical quantum gates”, which are connected with each other in a way that

is appropriate for the calculation to be performed. The simplest quantum gates act on a single qubit:

- the “X gate”, which transforms $|0\rangle$ into $|1\rangle$, and $|1\rangle$ into $|0\rangle$; its action is represented by the σ_x Pauli matrix.
- the “Hadamard gate H ”, which acts according to:

$$\begin{aligned} H|0\rangle &= \frac{1}{\sqrt{2}} [|0\rangle + |1\rangle] \\ H|1\rangle &= \frac{1}{\sqrt{2}} [|0\rangle - |1\rangle] \end{aligned} \tag{7.12}$$

(if the two basis states $|0\rangle$ and $|1\rangle$ correspond to photon states that are polarized horizontally and vertically, the effect of H is to turn these linear polarizations by 45°). Quantum gates acting on two qubit states also exist, such as the “cNOT gate” (for “controlled not”), which acts on both a control qubit and a target qubit. For an introduction to various gates and a rapid discussion of how they can be used to implement quantum algorithms, see for instance §6.5 of [288] or Chapter 8 of [302].

Among quantum algorithms, one often cites first the Shor algorithm [303]. Shor noticed that the factorization of large integer numbers into prime factors may become enormously faster than with classical methods; an introduction to the relation between quantum mechanics and the factorization of integer numbers is given in [304], and with more details in Chapter 3 of [299]. The Grover algorithm [305] is a quantum algorithm that allows searching an unsorted data base in a much more efficient way than classical computation (the gain is quadratic as a function of the number of objects in the base); the algorithm is probabilistic since it gives the correct answer with high probability, but the probability of failure can be decreased at will by repeating the algorithm. The Deutsch–Josza algorithm [306] provides another case where the gain offered by the quantum computation is exponential. Similar enhancements of the speed of computation are expected in the simulation of many-particle quantum systems [307]. Depending on the problem, the theoretical gain in speed provided by quantum algorithms is polynomial or exponential, but sometimes there is no gain at all. More recently, a new algorithm has been proposed to obtain useful information concerning the solutions of very large systems of linear equations, by providing approximate values of mathematical quantities depending on the solution [308].

Fundamentally, there are many differences between classical bits and quantum qubits. We have already mentioned that classical bits have two reference states that are fixed once and for all, while qubits can use any orthogonal basis in their space of states, but this is not the only difference by far. Classical bits can be copied at will and ad infinitum, while the no-cloning theorem mentioned in the preceding

section (see also Appendix F) applies to qubits. On the other hand, classical bits can be transmitted only into the forward direction of light cones, while the use of entanglement and teleportation may remove this limitation for qubits. Another difference is that the information is less directly coded in quantum qubits than in classical bits: in order to transmit and receive usable information from qubits, one has to specify what kind of measurements should be made with them (this is related to the flexibility concerning space state basis mentioned above). Since the quantum measurement process involves randomness, often the algorithm will provide results with randomness also, so that repeating it may be necessary. In the end, as all human beings, Alice and Bob can communicate only at a classical level, for instance by adjusting the macroscopic settings of their measurement apparatuses and observing the red and green light flashes associated with the results of measurements. Paraphrasing Bohr (see the end of §1.2.3), we could say that “there is no such concept as quantum information; information shared by humans is inherently classical, but may be transmitted through quantum qubits”; nevertheless, the whole field is now sometimes called “quantum information theory”.

One of the first proposal of a scheme of a quantum computer involved cold trapped atomic ions, in particular the realization of a cNOT gate, see [246]. Since then, many laboratories have realized experiments to demonstrate the feasibility of quantum computations with ions [247]. In 2001, a group at IBM [309] used the techniques of nuclear magnetic resonance to implement the Shor algorithm, and to factorize the number $N = 15$ into 3×5 . For the moment, we are still very far from practical applications!

7.4.3 Quantum error correction codes

Decoherence is the big enemy of quantum computation, since it constantly tends to destroy the useful coherent superpositions; this immediately reduces the full quantum information to its classical, boolean, component (made of diagonal matrix elements only). It is now perfectly clear that a “crude” quantum computer based on the naive use of non-redundant qubits will never work, at least with more than a very small number of them. It has been remarked that this kind of quantum computer would simply be a sort, or resurgence, of the old analog computers (errors in quantum information form a continuum), in an especially fragile version!

But it has also been pointed out that an appropriate use of quantum redundancy may allow one to design efficient error correcting schemes [310, 311]. Decoherence can be corrected by using a system containing more qubits, and by projecting its state into some subspaces in which the correct information about the significant qubit survives without error [312]; the theoretical schemes involve collective measurements of several qubits, which give access to some combined information on

all them, but none on a single qubit. It turns out that it is theoretically possible to “purify” quantum states by combining several systems in perturbed entangled states and by applying to them local operations, in order to extract a smaller number of systems in non-perturbed states [313]. Chapter 5 of [299] discusses in more detail how this “miracle” can happen and how error correcting systems can be built from the basic quantum logic gates.

Quantum error correction codes apply to various situations, including quantum computation as well as communication or cryptography [314]. Similarly, the notion of “quantum repeaters” [315] has been introduced to correct for the effect of imperfections and noise in quantum communication. Another very different approach to quantum computation has been proposed, based on a semi-classical concept where qubits are still used, but communicate only through classical macroscopic signals, which are used to determine the type of measurement performed on the next qubit [316]. This kind of computer should be much less sensitive to decoherence.

Another possibility is to accurately control the coupling of a quantum system to its environment. While, for many forms of dissipation, decoherence tends to destroy the quantum effects that are used for quantum computation, dissipation may also sometimes have the opposite effect. With a good control of the coupling of the system to the outside reservoir, the environment may actually drive the system to a steady state in which the outcome of the quantum computation is encoded [317].

Generally speaking, whether or not it will be possible one day to beat or control decoherence in a sufficiently large system for practical quantum computing still remains to be seen. Moreover, although the factorization into prime numbers is an important question (in particular for cryptography), as well as the many-body quantum problem, it would be nice to apply the principles of quantum computation to a broader scope of problems! The question as to whether or not quantum computation will become a practical tool one day remains open to debate [297, 318]. In any case, this is an exciting new field of research.

8

Quantum measurement

The process of measurement plays an important role in quantum theory. Measurements can be direct, if the physical system S interacts directly with the measurement apparatus M (as we have assumed until now), or indirect. In the latter case, the physical system first interacts with an ancillary system B , which may have a space of states that is very different from that of S , for instance much larger; after this interaction has finished, M is used to perform a measurement on B , without any direct interaction with A . Because S is then “protected” from any direct interaction with the measurement apparatus, the state of S is not necessarily strongly modified, and may even be only weakly affected. In both cases, the process implies entanglement between several physical systems. In this chapter, we study how this entanglement is created and used for measurements as well as the notion of weak and continuous measurements. These questions play an important role in several of the interpretations of quantum mechanics that we discuss in Chapter 10.

8.1 Direct measurements

The Von Neumann model of quantum measurement [4] provides a general frame for describing the process in terms of correlations appearing (or disappearing) in the state vector associated with the whole system $S + M$. In this model, the two systems S and M are initially described by a product state $|\Psi_0\rangle$ and interact during the time of measurement, so that they become entangled; they then reach a final state $|\Psi'\rangle$ and do not interact anymore.

8.1.1 Ideal measurement, Von Neumann model

We assume that the physical quantity measured on S is described by an Hermitian operator A acting in its space of states $\mathcal{E}_S$. This operator has normalized eigenvectors $|a_n\rangle$ associated with eigenvalues a_n (for the sake of simplicity, we assume that they

are non-degenerate):

$$A|a_n\rangle = a_n|a_n\rangle \quad (8.1)$$

Initially, the state $|\varphi_0\rangle$ of S is any linear combination of the $|a_n\rangle$:

$$|\varphi_0\rangle = \sum_{n=1}^{N_S} c_n |a_n\rangle \quad (8.2)$$

with arbitrary complex coefficients c_n ; the only condition is that the sum of their squared modulus is equal to 1 (normalization condition). As for the measurement apparatus M , we assume that before measurement it is always in the same quantum state $|\Phi_0\rangle$. The initial state of the whole system is then:

$$|\Psi_0\rangle = |\varphi_0\rangle \otimes |\Phi_0\rangle \quad (8.3)$$

8.1.1.a Basic measurement process

Let us first study the special case where, initially, S is in one of the eigenstates associated with the measurement: $|\varphi_0\rangle = |a_n\rangle$. The initial state of the whole system is then:

$$|\Psi_0\rangle = |a_n\rangle \otimes |\Phi_0\rangle \quad (8.4)$$

After the measurement, system S is left in the same state $|a_n\rangle$. Nevertheless, the measurement apparatus reaches a state $|\Phi'_n\rangle$ that is different from $|\Phi_0\rangle$ and depends on n ; it has to be so since, to make the result accessible experimentally, the position of its “pointer” used for reading the result on the apparatus (macroscopic pointer moving in front of a scale, state of light emitting diodes, writing of the result into a computer memory, etc.) must depend on n . The various states $|\Phi_n\rangle$ are actually mutually orthogonal, since the pointer necessarily involves a very large number of atoms reaching a different state to make the reading by a macroscopic observer possible¹. The effect of measurement is therefore summarized by:

$$|\Psi_0\rangle = |a_n\rangle \otimes |\Phi_0\rangle \implies |\Psi'\rangle = |a_n\rangle \otimes |\Phi'_n\rangle \quad (8.5)$$

where $|\Phi'_n\rangle$ is a normalized state. No correlation or entanglement has then appeared between the measurement apparatus and the measured system; this corresponds to the simple situation where the result of the measurement is certain.

¹ This is a necessary but not sufficient condition for M to constitute a good measurement apparatus. The states $|\Phi'_n\rangle$ should also provide a stable recording of the measurement result, which should not quickly be erased under the effect of the proper evolution of M or its coupling to the environment – we come back to this point in §8.1.2.b.

In the general case, the initial state of system S is a superposition (8.2) of eigenstates associated with the measurement. The state replacing (8.4) is the linear combination with the same coefficients:

$$|\Psi_0\rangle = \sum_n c_n |a_n\rangle \otimes |\Phi_0\rangle \quad (8.6)$$

The linearity of the Schrödinger equation then implies that:

$$|\Psi_0\rangle \implies |\Psi'\rangle = \sum_n c_n |a_n\rangle \otimes |\Phi'_n\rangle \quad (8.7)$$

which is a state where the measurement apparatus is entangled with the measured system S . After the measurement, one can then no longer attribute a state vector to system S (pure state), but only a density operator obtained by partial trace. Since the $|\Phi'_n\rangle$ are mutually orthogonal and normalized, this density operator is given by:

$$\rho'_S = Tr_M \{ |\Psi'\rangle \langle \Psi'| \} = \sum_n |c_n|^2 |a_n\rangle \langle a_n| \quad (8.8)$$

This result looks very natural: it expresses that the measured system has a probability $|c_n|^2$ to be in each of the states $|a_n\rangle$ associated with the results a_n , which corresponds exactly to Born's probability rule. It is indeed a very useful formula that summarizes some features of the measurement postulate of quantum mechanics. Nevertheless, as we have already emphasized (for instance in §§2.1 and 2.2), it does not contain one essential component of this postulate: the uniqueness of the result of measurement (emergence of macroscopic uniqueness). Indeed, all possible results are still contained in the partial trace, considered as equally possible even after the measurement. The system is described mathematically by a sum containing all possible results of the experiment simultaneously, while in reality one single result is observed in a given realization. This was expected: equation (8.8) is only a consequence of the Schrödinger equation, which is unable to stop the endless progression of the Von Neumann chain.

As we have seen in §1.2.2.a, to solve this problem Von Neumann introduced a specific postulate: the state reduction postulate, which forces the uniqueness of the result of the measurement. According to this postulate, a single result a_r is indeed observed after measurement and, in the summations of (8.7) and (8.8), all components except the component $n = r$ should be suppressed². Under these conditions, the state vector after measurement becomes a product, entanglement has disappeared, and S is in a pure state again.

² Non-standard interpretations of quantum mechanics (additional variables, modal interpretation, non-linear dynamics, etc.) solve the problem in different ways (Chapter 10).

8.1.1.b Effects of the interaction, pointer observable

Relations (8.5) or (8.7) imply that some interaction has occurred between S and M . We can for instance assume that their interaction Hamiltonian H_{int} has the form:

$$H_{int} = g A P_M \quad (8.9)$$

where A is the operator introduced above (acting on S only), P_M an operator acting on M only, and g a coupling constant. We also assume that, in the space of states of M , operator P_M has a conjugate operator X_M :

$$[X_M, P_M] = i\hbar \quad (8.10)$$

This means that P_M acts as the infinitesimal translation operator with respect to X_M , in other words that the action of its exponential on any eigenvector $|x_M\rangle$ of X_M :

$$X_M |x_M\rangle = x_M |x_M\rangle \quad (8.11)$$

provides a translation of the eigenvalue x_M :

$$e^{-i\Delta x P_M/\hbar} |x_M\rangle = |x_M + \Delta x\rangle \quad (8.12)$$

where Δx is any real number.

We now assume that $|\Phi_0\rangle$ (state of the apparatus before measurement) is an eigenstate of X_M with eigenvalue x_0 , and ignore any other source of evolution³ of the combined system other than the mutual interaction between S and M . The evolution operator between time $t = 0$ before measurement and time $t = \tau$ after the end of interaction is:

$$U(0, t) = e^{-ig\tau A P_M/\hbar} \quad (8.13)$$

Its application to (8.4) then provides:

$$U(t) |a_n\rangle \otimes |\Phi_0(x_0)\rangle = |a_n\rangle \otimes |\Phi(x_0 + g\tau a_n)\rangle \quad (8.14)$$

where the variables written between brackets in the state of the measurement apparatus⁴ refer to the eigenvalues of X_M . In terms of the kets $|\Phi'_n\rangle$ introduced in (8.5), we therefore have:

$$|\Phi'_n\rangle = |\Phi(x_0 + g\tau a_n)\rangle \quad (8.15)$$

³ This is a valid approximation if the interaction time τ is very short and g very large. One could include the proper evolutions of systems S and M by using the interaction representation associated with the sum of their Hamiltonians. Nevertheless, since here we are mostly interested in the effects of their mutual interaction, for the sake of simplicity we ignore the effects of the separate proper evolutions.

⁴ Needless to say, a measurement apparatus is macroscopic and has many more degrees of freedom than the only position of its pointer. For the sake of simplicity, these degrees of freedom are not made explicit in our notation.

These relations show that, for the measurement apparatus M , the eigenvalue of X has been translated by an amount $g\tau a_n$ that depends on the eigenvalue a_n of S . For the measurement apparatus, the observable X_M therefore plays the role of the position of a “pointer” which indicates the result of measurement after interaction between both systems.

If, initially, the system is in an eigenstate of A , it will remain in the same eigenstate after measurement. One then says that the measurement is a “quantum non-demolition” measurement⁵ (QND). Of course, in general, the system S is not initially in an eigenstate of A , and the interaction with the measurement apparatus changes its pure state into a complicated statistical mixture obtained by inserting (8.15) into (8.7); the usual “Von Neumann regress” (§2.1) takes place.

8.1.2 Effects of the environment

If, without the state reduction postulate, formula (8.8) does not suffice to introduce a uniqueness of the observed measurement result, one could at least hope that it determines without any ambiguity the basis into which the projection towards a single state associated with the measurement takes place. Actually, this is not necessarily the case. The notion of pointer states, which involves the environment of the measurement apparatus, is useful to remove this ambiguity.

8.1.2.a Ambiguous entanglement

To see why formula (8.8) may be ambiguous, let us consider a simple example: the measurement of the Oz component of a spin $1/2$, associated with eigenvectors $|+\rangle$ and $|-\rangle$. Initially, the spin is in the following state:

$$|\varphi_i\rangle = \frac{1}{\sqrt{2}} [|+\rangle + |-\rangle] \quad (8.16)$$

After interaction with the measurement apparatus, the state vector $|\Psi'\rangle$ of the whole system is:

$$|\Psi'\rangle = \frac{1}{\sqrt{2}} \left[|+\rangle \otimes |\Phi'_+\rangle + |-\rangle \otimes |\Phi'_-\rangle \right] \quad (8.17)$$

where the $|\Phi'_\pm\rangle$ are the two possible states that the measurement apparatus can reach.

⁵ QND measurements can for instance be performed in quantum optics [319]. In Chapter 9, we discuss other examples of QND measurements.

Looking at this expression of the vector, one could have the impression that it directly indicates that the basis of measurement is indeed the basis of the eigenvectors $|\pm\rangle$ associated with the Oz component of the spin. But, if one introduces the basis of spin states:

$$|\xi_{\pm}\rangle = \frac{1}{\sqrt{2}} \left[|+\rangle \pm e^{i\xi} |-\rangle \right] \quad (8.18)$$

(corresponding to the eigenstates of the spin component along axis $O\xi$ in plane xOy making angle ξ with Ox), $|\Psi'\rangle$ can be written as well:

$$|\Psi'\rangle = \frac{1}{\sqrt{2}} \left\{ |\xi_+\rangle \otimes |\Xi_+\rangle + |\xi_-\rangle \otimes |\Xi_-\rangle \right\} \quad (8.19)$$

with:

$$|\Xi_{\pm}\rangle = \frac{1}{\sqrt{2}} \left[|\Phi'_+\rangle \pm e^{-i\xi} |\Phi'_-\rangle \right] \quad (8.20)$$

Equation (8.19) has exactly the same form as (8.17); it is sufficient to replace the two vectors $|\pm\rangle$ by the two other vectors $|\xi_{\pm}\rangle$ as well as vectors $|\Phi'_{\pm}\rangle$ of the measurement apparatus by $|\Xi_{\pm}\rangle$. Looking at expression (8.19) of the entangled state vector after the measurement, one could now get the impression that it is the spin component along direction $O\xi$ that has been measured, in contradiction with what was obtained from expression (8.17) of the same ket. This example illustrates that, for an initial state of S such as (8.16), the nature of only the final entanglement between the system and the measured apparatus is not sufficient to determine which observable has been measured (of course, even less what is the result obtained in a given realization of the experiment).

To avoid this ambiguity, one can generalize our discussion and consider the case where $|\varphi_i\rangle$ can be any linear superposition⁶ of states $|\pm\rangle$. We now discuss another method, which takes into account the unavoidable interaction between the measurement apparatus and its environment.

8.1.2.b Pointer states

A first obvious condition is that the states $|\Phi'_n\rangle$ of the measurement apparatus should store the information related to the result of the measurement in a robust way, and not destroy it immediately under the effect of the proper evolution of M . This

⁶ If one replaces (8.16) by a more general superposition $|\varphi_i\rangle = \alpha |+\rangle + \beta |-\rangle$, one obtains $|\Psi'\rangle = \alpha |+\rangle \otimes |\Phi'_+\rangle + \beta |-\rangle \otimes |\Phi'_-\rangle$. For any α and β , one cannot find basis vectors $|\xi_{\pm}\rangle |\Theta_{\pm}\rangle$ other than that used in (8.17), in which this state can also be written in general $|\Psi'\rangle = \alpha |\xi_+\rangle \otimes |\Theta'_+\rangle + \beta |\xi_-\rangle \otimes |\Theta'_-\rangle$, so that the ambiguity is lifted.

condition is fulfilled if X_M is a constant of motion of M , in other words if X_M commutes with the proper Hamiltonian H_M of M .

Moreover, a measurement apparatus cannot remain completely isolated from its environment, even at a microscopic level. This would require that none of its macroscopic number of atoms, electrons, etc. interacts with any particle of the environment to become correlated with it in some way (for instance a photon that is scattered into a different state, depending on the position of the measurement pointer, as in the scheme discussed in §6.3.3.a). One can even remark that, because of its very purpose, the apparatus must be able to interact and correlate with the experimenter, when he acquires the knowledge of the result of measurement. This impossibility has stimulated Zeh [320] and Zurek [321, 322] to push further the analysis, and to include the environment of the measurement apparatus in the process of quantum measurement.

Models have been developed which, at the price of some simplification, lead to tractable calculations. This gives a good idea of the general nature of the physical phenomena occurring during the entanglement with the environment. In the Coleman–Hepp model [323, 324], the system under study is coupled to a semi-infinite linear array of N spin $1/2$ particles at fixed positions, representing the apparatus. The calculation then shows that, if “local” observables are built from a finite number of spin operators, their measurement cannot make the distinction between a coherent superposition of the various positions of the pointer and an incoherent statistical mixture. Zurek considers an environment E made of a large number of either harmonic oscillators [325], or spins $1/2$ that are all coupled to the measurement apparatus M [322]; the apparatus is also simplified and treated as a two-level system (equivalent to a spin $1/2$). This coupling introduces a Hamiltonian $H_{ME}(i)$ between the i th spin of the environment and the spin of the measurement apparatus; $H_{ME}(i)$ is proportional to the product of the components of the two spins along Oz with a coupling constant g_i , which varies randomly as a function of i . One can then show that a preferred basis exists in the space of states of the measurement apparatus M (as simplified in this model), the “pointer states basis” in which the entanglement with E tends to destroy the coherences. What determines this basis, preferred to all the others, is the form of the coupling Hamiltonian between the measurement apparatus and its environment. For a more realistic measurement apparatus than a single spin, for instance an apparatus that includes a pointer moving in front of a scale, the preferred basis corresponds to states where the position of the pointer is well defined (as opposed, for instance, to the eigenbasis of its momentum, or to any other state that is significantly extended in space). Zurek calls this phenomenon “einselection” (for “environment induced selection”).

This is an important idea: the very physical constitution of a measurement apparatus determines the way it is coupled to the environment. The Hamiltonian describing

this coupling determines the basis of the pointer states. Actually, if the proper dynamics of the apparatus is neglected, pointer states are merely the eigenvectors of the operator which commutes with the interaction Hamiltonian between the apparatus and the environment [321]; if this dynamics is included, the situation is more complicated. Several necessary conditions therefore determine if a device can be considered as satisfactory to give access to a physical quantity of S ; first, the coupling between S and M must be appropriate to transfer the desired information from S to M ; second, this information must be recorded in a way that makes it stable in time with respect to both the proper evolution of M and the coupling between M and E ; finally, it should also be robust against small perturbations.

Needless to say, from a fundamental point of view, once more the same remark applies: uniqueness does not emerge from this theory. It just determines the basis of states of the apparatus that are sufficiently robust with respect to the environment to store the result of the measurement. An independent postulate is therefore still needed. Zurek [326] (see also §VI-D of [322] and §III-F of [258]) has nevertheless proposed to use the notion of “*envariance*” (environment-assisted invariance), a symmetry exhibited by correlated quantum systems and related to causality, to describe the nature of statistical ignorance and derive the Born rule and macroscopic classicality, without any notion of measurement or state vector collapse.

8.1.3 *Hund paradox*

A similar analysis applies to a problem in molecular physics, the so-called “Hund paradox” [327], namely the origin of the stability of chiral states of molecules. Chiral molecules in their ground state have access to two configurations, one left-handed configuration associated with quantum state $|L\rangle$, and one right-handed configuration associated with quantum state $|R\rangle$. A mirror image symmetry transforms each of these states into the other, but none of them is invariant: each has chirality, which means that a solution of molecules in one of these states exhibits “rotary power” (it rotates the plane of polarization of an optical beam that crosses the solution).

These two states are always coupled by tunnel effect (as for the two symmetrical configurations of the Ammonia molecule), so that the real ground state of the molecule is the symmetrical combination:

$$|G\rangle = \frac{1}{\sqrt{2}} [|L\rangle + |R\rangle] \quad (8.21)$$

This state is invariant under mirror symmetry, as expected: the non-degenerate ground state of any quantum system has the same symmetry as its Hamiltonian (in this case the Hamiltonian of the molecule, which is invariant in this symmetry). It has no chirality and therefore no rotary power.

Nevertheless, by measuring the rotation of the plane of polarization of light when it crosses the solution, one observes that molecules are “enantiomers”: they are either in state $|L\rangle$ or $|R\rangle$, which indeed rotate polarizations in opposite directions, but they are not in the real ground state $|G\rangle$ with no rotary power. One can even sort molecules of one configuration, $|L\rangle$ or $|R\rangle$ only, to enhance the rotation of the solution. The intriguing question is then: why aren’t the molecules in a state with the same symmetry as the Hamiltonian of the molecule, such as state $|G\rangle$? More generally, why can’t the molecules be found in any superposition:

$$|D_{\pm}\rangle = \frac{1}{\sqrt{2}} \left[|L\rangle \pm e^{i\xi} |R\rangle \right] \quad (8.22)$$

since this possibility is a direct consequence of the superposition principle?

The answer to the question involves the coupling to the environment, as above for pointer states. The molecule constantly collides with others or scatters photons so that it correlates with them in a way that is different for states $|L\rangle$ and $|R\rangle$ because interactions are local; those two states therefore provide the preferred basis with respect to the environment, while $|G\rangle$ is a combination of such states and therefore very fragile against decoherence [328]. This explains why state $|G\rangle$ is not observed in the optical properties of the solution.

Actually, if the coupling with the other molecules is sufficiently fast and the tunnel effect relatively slow, the interaction with the environment completely blocks the tunnel effect. Assume a molecule is initially in state $|L\rangle$; if the molecule was completely isolated, it should oscillate between states $|L\rangle$ and $|R\rangle$ under the effect of its internal Hamiltonian, at a frequency fixed by the rate of the tunnel effect. Nevertheless, the effect of the environment blocks this oscillation, so that state $|R\rangle$ is never reached by the molecule (this is sometimes called the “quantum Zeno effect”).

8.2 Indirect measurements

To introduce the notions of indirect measurements and of POVM, we begin with a simple case: a two-level system coupled to another quantum system B with an arbitrarily large space of states.

8.2.1 A simple model: two-level system

We assume that the space of states $\mathcal{E}_S$ of S is two dimensional and call $|+\rangle$ and $|-\rangle$ two states forming an orthonormal basis in $\mathcal{E}_S$. Between times $t = 0$ and $t = \tau$, system S is coupled to another system B called “ancillary system”, with space of states $\mathcal{E}_B$ spanned by an orthonormal basis $|\Phi_m\rangle$ (with $q = 0, 1, \dots, N_B$; the

dimension N_B may be much larger than 2). At time τ , a measurement is performed on the ancillary system, without any direct interaction with S ; we ignore the proper evolution of S and B during this time interval, and take into account only their interaction (otherwise we should use the interaction representation with respect to the sum of the proper Hamiltonians of both systems, which would complicate the calculations).

8.2.1.a Interaction and entanglement

The interaction Hamiltonian is:

$$H_{int.} = g \sigma_z(S) \Xi(B) \quad (8.23)$$

where g is a coupling constant, $\sigma_z(S)$ the operator acting in the space of states of S (defined by the third Pauli matrix in the $\{|\pm\rangle\}$ basis), and $\Xi(B)$ an arbitrary operator acting on B only. Initially, the ensemble of $S + B$ is in the state:

$$|\Psi_0\rangle = |\varphi_0\rangle \otimes |\Phi_0\rangle \quad (8.24)$$

where $|\varphi_0\rangle$ is any state in $\mathcal{E}_S$:

$$|\varphi_0\rangle = \alpha |+\rangle + \beta |-\rangle \quad (8.25)$$

Both $|\varphi_0\rangle$ and $|\Phi_0\rangle$ are normalized. Since the proper evolutions of S and B are ignored, the evolution operator between $t = 0$ and $t = \tau$ is:

$$U(0, \tau) = e^{-ig\tau \sigma_z(S) \Xi(B)/\hbar} \quad (8.26)$$

If S is initially in one of the states $|+\rangle$ or $|-\rangle$, relation (8.5) becomes:

$$|\Psi_0\rangle = |\pm\rangle \otimes |\Phi_0\rangle \implies |\Psi'\rangle = |\pm\rangle \otimes |\Phi_B^\pm\rangle \quad (8.27)$$

with:

$$|\Phi_B^\pm\rangle = e^{\mp ig\tau \Xi(B)/\hbar} |\Phi_0\rangle \quad (8.28)$$

The state of S is then unchanged, and will not be affected by whatever measurement is performed on B , since the two systems are not entangled. Note that $|\Phi_B^\pm\rangle$ are normalized vectors (they are obtained by the action of unitary operators onto $|\Phi_0\rangle$) but that they are in general not mutually orthogonal: for instance, if the coupling constant g tends to zero, they both tend to the same vector $|\Phi_0\rangle$ in the space of states of B .

In the more general case where S is initially in state (8.25), the ensemble $S + B$ after interaction is described by the entangled state:

$$|\Psi'\rangle = U(0, \tau) |\varphi_0\rangle |\Phi_0\rangle = \alpha |+\rangle |\Phi_B^+\rangle + \beta |-\rangle |\Phi_B^-\rangle \quad (8.29)$$

so that the effect of the measurement performed on B may change the state of S , as we now discuss.

8.2.1.b Measurement on the ancillary system

At some time $t \geq \tau$, a measurement is performed on system B ; the possible results b_m , corresponding to the orthonormal kets $|\Phi_m\rangle$, are assumed to be non-degenerate for simplicity. Expanding over the corresponding basis allows to write $|\Psi'\rangle$ as:

$$|\Psi'\rangle = \sum_m \left[\alpha \langle \Phi_m | \Phi_B^+ \rangle |+\rangle + \beta \langle \Phi_m | \Phi_B^- \rangle |-\rangle \right] \otimes |\Phi_m\rangle \quad (8.30)$$

If the result of measurement is b_r , the projection postulate reduces this ket to one of its components in the summation over m , so that it becomes the product:

$$|\Psi_{mes.}^r\rangle = \left[\alpha \langle \Phi_r | \Phi_B^+ \rangle |+\rangle + \beta \langle \Phi_r | \Phi_B^- \rangle |-\rangle \right] \otimes |\Phi_r\rangle \quad (8.31)$$

System S is then in a pure state again, state $|\varphi_{mes.}^r\rangle$ given by:

$$|\varphi_{mes.}^r\rangle = \alpha \langle \Phi_r | \Phi_B^+ \rangle |+\rangle + \beta \langle \Phi_r | \Phi_B^- \rangle |-\rangle \quad (8.32)$$

or:

$$|\varphi_{mes.}^r\rangle = \Pi_r |\varphi_0\rangle \quad (8.33)$$

where Π_r is the operator acting in $\mathcal{E}_S$ defined by:

$$\Pi_r = \langle \Phi_r | \Phi_B^+ \rangle |+\rangle \langle +| + \langle \Phi_r | \Phi_B^- \rangle |-\rangle \langle -| \quad (8.34)$$

In the general case, the state of S is indeed changed by the action of the indirect measurement.

In general, Π_r is not a projection operator, since its square is not the same operator:

$$[\Pi_r]^2 = \langle \Phi_r | \Phi_B^+ \rangle^2 |+\rangle \langle +| + \langle \Phi_r | \Phi_B^- \rangle^2 |-\rangle \langle -| \quad (8.35)$$

(except in the special case where $|\Phi_r\rangle$ is equal to one of the $|\Phi_B^\pm\rangle$ and orthogonal to the other). For this reason, the operation is often called a “non-projective measurement”, as opposed to projective measurements, such as those discussed in §8.1.

A similar calculation gives:

$$\Pi_r^\dagger \Pi_r = \Pi_r \Pi_r^\dagger = \langle \Phi_B^+ | \Phi_r \rangle \langle \Phi_r | \Phi_B^+ \rangle |+\rangle \langle +| + \langle \Phi_B^- | \Phi_r \rangle \langle \Phi_r | \Phi_B^- \rangle |-\rangle \langle -| \quad (8.36)$$

8.2.1.c Probabilities and sum rule

The probability $\mathcal{P}_r$ of obtaining result b_r is nothing but the square of the norm of ket (8.31), which by (8.33) and (8.34) can be simply expressed as an average value over the initial state $|\varphi_0\rangle$ of S :

$$\mathcal{P}_r = |\alpha \langle \Phi_r | \Phi_B^+ \rangle|^2 + |\beta \langle \Phi_r | \Phi_B^- \rangle|^2 = \langle \varphi_0 | \Pi_r^\dagger \Pi_r | \varphi_0 \rangle \quad (8.37)$$

A summation over index r in (8.36) then introduces a closure relation over basis $\{|\Phi_r\rangle\}$, and one obtains:

$$\sum_r \Pi_r^\dagger \Pi_r = \langle \Phi_B^+ | \Phi_B^+ \rangle |+\rangle \langle +| + \langle \Phi_B^- | \Phi_B^- \rangle |-\rangle \langle -| = |+\rangle \langle +| + |-\rangle \langle -| \quad (8.38)$$

or:

$$\sum_r \Pi_r^\dagger \Pi_r = 1 \quad (8.39)$$

which means that the sum of probabilities gives 1, as expected.

As a consequence, the series of operators $\Xi_r = \Pi_r^\dagger \Pi_r$ provides a decomposition of the identity operator in the space of states of S . In general, the Ξ_r are not mutually orthogonal (their binary products are not zero), as one can easily see from (8.34). With mutually orthogonal projectors in a two-dimensional space associated with a spin 1/2, the identity could only be obtained by summing two projectors over two opposite spin directions; it is impossible to build more than two orthogonal projectors in $\mathcal{E}_S$. With the non-orthogonal Ξ_r , a decomposition of identity is obtained with N_B operators, the dimension of the space of state of B , which can be much larger than 2. For instance, one can obtain the identity in the space of states of a spin 1/2 by summing projectors over any series of K spin directions defined by vectors in space having a zero sum, and dividing them by $K/2$. Since, moreover, the Ξ_r are not necessarily projectors, there is a large flexibility for forming series of operators summing to 1 as in (8.39).

8.2.2 Generalization: POVM

We now assume that the dimension of the space of states $\mathcal{E}_S$ has any value N_S , instead of necessarily 2. We also assume that the two systems evolve under the effect of their interaction only, with a unitary evolution given for instance by (8.13). We can then use the results already obtained in §6.4.1 and, using definition (6.47) of operators M_q as well as (6.49), write the state vector $|\Psi'\rangle$ after

interaction as:

$$|\Psi'\rangle = \sum_{q=0}^{N_B-1} M_q |\varphi_0\rangle \otimes |\Theta_q\rangle \quad (8.40)$$

where the $|\Theta_q\rangle$ form an orthonormal basis in the space of states of B .

We can choose for this basis the eigenstates of the measurement performed on B . The probability $\mathcal{P}_r$ to obtain result b_r associated with a particular $|\Phi_r\rangle$ is given by the norm of the corresponding term $q = r$ in the expansion of the state written in (8.40). This probability is thus:

$$\mathcal{P}_r = \langle \varphi_0 | M_r^\dagger M_r | \varphi_0 \rangle \quad (8.41)$$

Consider the series of operators:

$$K_r = M_r^\dagger M_r \quad (8.42)$$

Relation (6.51) indicates that their sum is equal to the identity operator:

$$\sum_{r=0}^{N_B-1} K_r = 1 \quad (8.43)$$

and the average of each K_r in any state of S is always a positive number (a probability). Such an ensemble of operators are called a POVM (an acronym for “Positive Operator Valued Measure”). As in the case where S is a spin 1/2, the number of operators in a POVM is not equal to the dimension of space $\mathcal{E}_S$, but to the dimension of space for system B . Their number may therefore be much larger, as we have seen above when the ancillary system is a spin 1/2. It may also be much smaller if, for instance, the ancillary system B is a spin 1/2: in this case, the number of POVM operators acting in the space of states of S cannot exceed 2, whatever the dimension of this space is. A well-known example of a two-operator POVM acting in a two-dimensional space of states $\mathcal{E}_S$ spanned by the two vectors $|\pm\rangle$ is given by the two operators:

$$M_0 = |-\rangle \langle -| \quad M_1 = |-\rangle \langle +| \quad (8.44)$$

satisfying:

$$M_0^\dagger M_0 + M_1^\dagger M_1 = |-\rangle \langle -| + |+\rangle \langle +| = 1 \quad (8.45)$$

If M_0 is a projector, M_1 is not such an operator (its square vanishes); the cross products of these operators are $M_0 M_1 = M_1$ and $M_1 M_0 = 0$.

8.3 Weak and continuous measurements

Situations where the entanglement produced during the interaction stage between S and M remains weak are appropriate for performing weak and continuous measurements.

8.3.1 Measurements of weak values

The notion of “weak values” was introduced by Aharonov, Albert, and Vaidman [274] (see also Chapter 16 of [273], and [275]). As in §8.1.1.b, let us consider the evolution of the coupled system $S + M$, but in the opposite case where the indication readable on the pointer is inaccurate and provides little information on the state of S . This case is obtained if, in the expression (8.9) of the coupling Hamiltonian, the coupling constant g is small and can be treated to first order only. The initial state of the system is:

$$|\Psi_0\rangle = |\varphi_0\rangle |\Phi_0(x_0)\rangle \quad (8.46)$$

where $|\varphi_0\rangle$ has the general form (6.41), expressed in terms of the eigenvectors $|a_n\rangle$ of A . Instead of assuming that $|\Phi_0\rangle$ is an eigenstate of the pointer observable X_M , we now assume that it is a state $|\Phi_0(x_0)\rangle$ with Gaussian components on the eigenvectors $|x_M\rangle$ of this operator; we choose a broad Gaussian function centered around $x_M = x_0$:

$$\langle x_M | \Phi_0(x_0) \rangle = (2\pi\sigma^2)^{-1/4} e^{-(x_M - x_0)^2/4\sigma^2} \quad (8.47)$$

where σ is the mean-square fluctuation of the position of the pointer in this initial state. After interaction between time $t = 0$ and time $t = \tau$, the state vector of the whole system is obtained by action on $|\Psi_0\rangle$ of the evolution operator (8.13), and given by⁷:

$$\begin{aligned} |\Psi'\rangle &= (2\pi\sigma^2)^{-1/4} \sum_n c_n \int dx_M e^{-(x_M - x_0)^2/4\sigma^2} e^{-ig\tau a_n P_M/\hbar} |a_n\rangle \otimes |x_M\rangle \\ &= (2\pi\sigma^2)^{-1/4} \sum_n c_n \int dx_M e^{-(x_M - x_0)^2/4\sigma^2} |a_n\rangle \otimes |x_M + g\tau a_n\rangle \end{aligned} \quad (8.48)$$

With a change of the integration variable x_M into $x = x_M + g\tau a_n$, we obtain:

$$|\Psi'\rangle = (2\pi\sigma^2)^{-1/4} \sum_n c_n \int dx e^{-(x - x_0 - g\tau a_n)^2/4\sigma^2} |a_n\rangle \otimes |x\rangle \quad (8.49)$$

⁷ As in §8.2.1, we consider that the only evolution of the whole system is due to the interaction Hamiltonian between S and M , ignoring the proper evolution of the two isolated systems.

or, to first order in $g\tau$ (we assume $g\tau a_n \ll \sigma$ for any a_n):

$$|\Psi'\rangle = (2\pi\sigma^2)^{-1/4} \sum_n c_n \int dx \left[1 + g\tau a_n \frac{d}{dx_0} + \dots \right] e^{-(x-x_0)^2/4\sigma^2} |a_n\rangle \otimes |x\rangle \quad (8.50)$$

The summation over n of the term in 1 in the bracket reconstructs the ket $|\varphi_0\rangle$ written in (8.2), while the second term in the bracket introduces $A|\varphi_0\rangle$. We then obtain:

$$|\Psi'\rangle = (2\pi\sigma^2)^{-1/4} \left[1 + g\tau A \frac{d}{dx_0} + \dots \right] |\varphi_0\rangle \otimes |\Phi_0(x_0)\rangle \quad (8.51)$$

Now, assume that a measurement is performed on S , projecting this system into state $|\varphi_f\rangle$ corresponding to a non-degenerate eigenvalue of some operator acting in $\mathcal{E}_S$. After this measurement, the state becomes:

$$\begin{aligned} |\Psi''\rangle &\sim \left[\langle\varphi_f|\varphi_0\rangle + g\tau \langle\varphi_f|A|\varphi_0\rangle \frac{d}{dx_0} \right] |\varphi_f\rangle |\Phi_0(x_0)\rangle \\ &= \langle\varphi_f|\varphi_0\rangle \left[1 + \frac{g\tau \langle\varphi_f|A|\varphi_0\rangle}{\langle\varphi_f|\varphi_0\rangle} \frac{d}{dx_0} \right] |\varphi_f\rangle |\Phi_0(x_0)\rangle \end{aligned} \quad (8.52)$$

If the ratio between the matrix elements $\langle\varphi_f|A|\varphi_0\rangle$ and $\langle\varphi_f|\varphi_0\rangle$ is real, we finally obtain the simple relation:

$$|\Psi''\rangle \sim |\varphi_f\rangle |\Phi_0(x_0 + \Delta x_0)\rangle \quad (8.53)$$

with:

$$\Delta x_0 = g\tau \frac{\langle\varphi_f|A|\varphi_0\rangle}{\langle\varphi_f|\varphi_0\rangle} \quad (8.54)$$

Δx_0 is the average shift of the position of the pointer for this particular series of events; it can be large if $\langle\varphi_f|\varphi_0\rangle$ has a very small modulus. If the ratio between the matrix elements is complex:

$$\frac{\langle\varphi_f|A|\varphi_0\rangle}{\langle\varphi_f|\varphi_0\rangle} = R + iJ \quad (8.55)$$

one can show that the additional term in J changes the phase of the state of M after the measurement, which to first order corresponds to the substitution:

$$e^{-(x-x_0-\Delta x_0)^2/4\sigma^2} \implies e^{i\xi(x)} e^{-(x-x_0-\Delta x_0)^2/4\sigma^2} \quad (8.56)$$

where $\xi(x)$ is a phase variable⁸. Measurements of the position of the pointer variable are not affected by this change (but a measurement of the conjugate variable P_M

⁸ To first order, the phase shift $\xi(x)$ is given by:

$$\xi(x) = g\tau J e^{(x-x_0)^2/\sigma^2} \frac{d}{dx_0} \left[e^{-(x-x_0)^2/\sigma^2} \right] = g\tau J \frac{x-x_0}{2\sigma^2} \quad (8.57)$$

would give access to it). In what follows, we limit ourselves to the case where the ratio between the matrix elements is real.

The experimental procedure is therefore the following: first, one prepares S in a state $|\varphi_0\rangle$ and M in a state $|\Phi_0(x_0)\rangle$ associated with a wide distribution of the pointer variable x_M ; the two systems are left to interact for a short time τ ; one then measures S with an observable that has $|\varphi_f\rangle$ as a non-degenerate eigenstate, and one selects only the events where S is found in this state (this is a post-selection process); finally, one observes the average shift Δx_0 of the pointer variable of M . The interesting result is that, with this combined scheme of pre- and post-selection performed on S , the observed shift can be much larger than any of the shifts $g\tau a_n$ associated with the eigenvalues of A ; the amplification factor is the real part of the ratio written in (8.55). Δx_0 is called the “weak value” associated with this measurement [273, 274] – even if its value may be surprisingly large. Note that our first-order calculation implies that the weak value should always remain much smaller than the width σ associated with the initial state of the measurement apparatus. To detect this amplified shift with a reasonable accuracy, a large number of successive measurements is therefore necessary in order to reduce the uncertainty by averaging the individual fluctuations of the observations.

A good amplification factor is obtained if $\langle\varphi_f|\varphi_0\rangle$ is small but $\langle\varphi_f|A|\varphi_0\rangle$ large. Situations where $|\varphi_0\rangle$ and $|\varphi_f\rangle$ are almost orthogonal are therefore favorable, but this implies that the post-selection process will reject most events, which means that the experiment has to be repeated a very large number of times to provide a significant result. For a spin 1/2 for instance, one can assume that:

$$|\varphi_0\rangle = |+\rangle \quad \text{and} \quad |\varphi_f\rangle = \varepsilon |+\rangle + \alpha |-\rangle \quad (8.58)$$

with ε real and very small, and α almost equal to 1. If the measurement apparatus M measures the spin Pauli operator $A = \sigma_x$, then:

$$A|\varphi_0\rangle = A|+\rangle = \frac{1}{\sqrt{2}}|-\rangle \quad (8.59)$$

The scalar product of this ket with $|\varphi_f\rangle$ is not infinitesimal when ε tends to zero, but almost equal to $1/\sqrt{2}$. As a consequence, the amplification factor (8.55) is $\sim 1/(\varepsilon\sqrt{2})$, which can be arbitrarily large if ε is very small.

Another interesting situation occurs when:

$$|\varphi_0\rangle = |+\rangle \quad \text{and} \quad |\varphi_f\rangle = \frac{1}{\sqrt{2}}[|+\rangle + |-\rangle] \quad (8.60)$$

The spins then start in an eigenstate of σ_z and reach an eigenstate of σ_x , both with eigenvalues 1. We assume that:

$$A = \frac{1}{\sqrt{2}} [\sigma_x + \sigma_z] \quad (8.61)$$

which corresponds to a spin measurement along a direction in the xOz plane at 45° of both axes Ox and Oz . An easy calculation then shows that the amplification factor is $\sqrt{2}$. In other words, the spins interact with M as if their component at 45° was not ± 1 , as for any component of $\vec{\sigma}$, but the geometrical sum $1/\sqrt{2} + 1/\sqrt{2} = \sqrt{2}$ obtained by adding the initial value of the spin with the final value (as if they were classical perpendicular vectors). A curious property is that this value exceeds the largest eigenvalue $+1$ of any component of $\vec{\sigma}$. The authors of [275] propose to interpret this situation in a non-standard way, by considering that the pre-selection and post-selection processes, when combined, select spins having two orthogonal components perfectly well defined (which is of course totally impossible in the usual formulation of quantum mechanics, where two perpendicular spin components cannot have well-defined values at the same time).

Similar ideas apply to the time evolution of a quantum system S coupled to an external system B through an Hamiltonian gQ_SQ_B , where Q_S acts in the space of states of S and Q_B in the space of states of B . Pre-selecting the state of B in a quantum superposition of eigenstates of Q_B and post-selecting its state in another superposition may lead to a quantum evolution of S that is equivalent to the superposition of different Hamiltonians, and of evolutions at different periods of time [329].

The general idea of measuring weak values is not restricted to quantum physics, but applies to any wave theory. In classical optics for instance, when detecting the weak field scattered by an object, it is well known that one can enhance the contrast by inserting almost orthogonal polarizers on the input and output beams. A variety of experiments have been made to measure weak values, either in a regime of classical optics, or in a purely quantum regime involving entanglement between particles [330–338]. The measurement of weak values has also been used [339] to test the Leggett–Garg inequalities (§4.2.3) or to obtain a “direct measurement” of a quantum wave function [276].

8.3.2 Continuous measurements

We now consider a situation similar to weak measurement but, instead of assuming that S is post-selected in a given state, we calculate the evolution of its state when the position of the pointer of M is measured, still ignoring the proper evolution of the isolated systems S and M . We start from the expression (8.49) of the quantum

state of the combined system $S + M$, which provides:

$$|\Psi'\rangle = \left(2\pi\sigma^2\right)^{-1/4} \int dx e^{-(x-x_0-g\tau A)^2/4\sigma^2} |\varphi_0\rangle \otimes |x\rangle \quad (8.62)$$

and we assume that a measurement of the position of the pointer of M has given a result in the interval:

$$\left[x_r - \frac{dx_r}{2}, x_r + \frac{dx_r}{2} \right] \quad (8.63)$$

where $dx_r \ll \sigma$. After the measurement, the state $|\Psi''\rangle$ is obtained by projecting $|\Psi'\rangle$ over the eigenstates of X_M corresponding to this interval:

$$|\Psi''\rangle = \left(2\pi\sigma^2\right)^{-1/4} e^{-(x_r-x_0-g\tau A)^2/4\sigma^2} |\varphi_0\rangle \otimes \int_{x_r-dx_r/2}^{x_r+dx_r/2} dx |x\rangle \quad (8.64)$$

Since this ket is a tensor product, we can attribute a ket $|\varphi''\rangle$ to system S :

$$|\varphi''\rangle \sim e^{-(x_r-x_0-g\tau A)^2/4\sigma^2} |\varphi_0\rangle \quad (8.65)$$

In what follows, we assume that the dimensionless parameter:

$$\varepsilon = \frac{g\tau a}{\sigma} \quad (8.66)$$

(where a is the largest of the modulus of all eigenvalues of A) is small, and perform a second-order calculation with respect to this parameter; the reason why a first-order calculation is not sufficient will become apparent at the end of §8.3.2.c.

8.3.2.a Probability of the result

The probability $\mathcal{P}(x_r)dx_r$ of obtaining a result in the interval (8.63) is given by the square of the norm of $|\Psi''\rangle$, which gives:

$$\mathcal{P}(x_r)dx_r = \left(2\pi\sigma^2\right)^{-1/2} \langle\varphi_0| e^{-(x_r-x_0-g\tau A)^2/2\sigma^2} |\varphi_0\rangle dx_r \quad (8.67)$$

By expanding the exponential to second order in ε , we obtain:

$$\begin{aligned} \mathcal{P}(x_r) = & \left(2\pi\sigma^2\right)^{-1/2} e^{-(x_r-x_0)^2/2\sigma^2} \\ & \times \left\{ 1 + \frac{g\tau}{\sigma} \frac{x_r-x_0}{\sigma} \langle A \rangle_0 + \left(\frac{g\tau}{\sigma}\right)^2 \langle A^2 \rangle_0 \left[-\frac{1}{2} + \frac{(x_r-x_0)^2}{2\sigma^2} \right] + \dots \right\} \end{aligned} \quad (8.68)$$

where the averages of A and its square in state $|\varphi_0\rangle$ are defined by:

$$\langle A \rangle_0 = \langle \varphi_0 | A | \varphi_0 \rangle \quad \text{and} \quad \langle A^2 \rangle_0 = \langle \varphi_0 | A^2 | \varphi_0 \rangle \quad (8.69)$$

To second order in ε , we then obtain:

$$\begin{aligned} \mathcal{P}(x_r) = & \left(2\pi\sigma^2\right)^{-1/2} \\ & \times \left\{ e^{-(x_r - x_0 - \Delta x_0)^2 / 2\sigma^2} + \left(\frac{g\tau}{\sigma}\right)^2 e^{-(x_r - x_0)^2 / 2\sigma^2} F(x_r - x_0) + \dots \right\} \end{aligned} \quad (8.70)$$

where:

$$\Delta x_0 = g\tau \langle A \rangle_0 \quad (8.71)$$

and where F is an even function⁹ of $(x_r - x_0)$. To first order in $\varepsilon = g\tau a/\sigma$, one can set $F = 0$ and the distribution is a shifted Gaussian distribution, centered around:

$$\bar{x}_r = x_0 + \Delta x_0 \quad (8.72)$$

If one calculates the second-order correction to the average of $(x_r - x_0 - \Delta x_0)$, one obtains no second-order correction: the term in $(x_r - x_0)$ vanishes because of the parity of F , and the term in Δx_0 combined with the second term in the right-hand side of (8.70) gives a third-order correction.

We conclude that the average value of the result of the measurement is not exactly the most probable initial position of the pointer x_0 , but a value shifted by an amount Δx_0 given by (8.71); the shift depends on the initial state $|\varphi_0\rangle$ of S . We then introduce the dimensionless random variable ξ_r defined by:

$$\xi_r = \frac{x_r - x_0 - \Delta x_0}{\sigma} \quad (8.73)$$

which characterizes the random result of the measurement, has zero average value, and a unit mean-square deviation.

8.3.2.b Evolution of the state

With this notation, the state (8.65) of S after measurement becomes:

$$|\varphi''\rangle \sim e^{-[\xi_r - \frac{g\tau}{\sigma}(A - \langle A \rangle_0)]^2 / 4} |\varphi_0\rangle \quad (8.74)$$

⁹ The expression of F is:

$$F = -\frac{1}{2} \langle A^2 \rangle_0 + \frac{(x_r - x_0)^2}{2\sigma^2} [\langle A^2 \rangle_0 - (\langle A \rangle_0)^2]$$

or, by expanding the exponential to second order:

$$|\varphi''\rangle \sim e^{-(\xi_r)^2/4} \times \left\{ 1 + \frac{g\tau}{2\sigma} \xi_r [A - \langle A \rangle_0] + \left(\frac{g\tau}{2\sigma} \right)^2 \left[\frac{1}{2} (\xi_r)^2 - 1 \right] [A - \langle A \rangle_0]^2 + \dots \right\} |\varphi_0\rangle \quad (8.75)$$

The prefactor $e^{-(\xi_r)^2/4}$ is irrelevant, and we only need to calculate the square of the norm of the ket written in the second line of this expression, that is the product of this ket by the associated bra. This product introduces average values in the initial state $|\varphi_0\rangle$, which we note $\langle \rangle_0$ as above. The first-order term contains $(\langle A \rangle_0 - \langle A \rangle_0)$ and vanishes; the zero and second-order terms then give:

$$1 + \left(\frac{g\tau}{2\sigma} \right)^2 \langle [A - \langle A \rangle_0]^2 \rangle_0 \left[(\xi_r)^2 + 2 \left[\frac{1}{2} (\xi_r)^2 - 1 \right] \right] + \dots \quad (8.76)$$

To obtain a normalized state, we must therefore multiply the ket by:

$$\left\{ 1 + \left(\frac{g\tau}{2\sigma} \right)^2 \langle [A - \langle A \rangle_0]^2 \rangle_0 \left[2 (\xi_r)^2 - 2 \right] + \dots \right\}^{-1/2} \quad (8.77)$$

or:

$$\left\{ 1 - \left(\frac{g\tau}{2\sigma} \right)^2 \langle [A - \langle A \rangle_0]^2 \rangle_0 \left[(\xi_r)^2 - 1 \right] + \dots \right\} \quad (8.78)$$

If we multiply the curly bracket in the right-hand side (8.75) by this bracket, we obtain the normalized ket $|\widehat{\varphi}''\rangle$ after measurement:

$$|\widehat{\varphi}''\rangle = \left\{ 1 + \frac{g\tau}{2\sigma} \xi_r [A - \langle A \rangle_0] - \frac{1}{2} \left(\frac{g\tau}{2\sigma} \right)^2 (\xi_r)^2 [A - \langle A \rangle_0]^2 + \left(\frac{g\tau}{2\sigma} \right)^2 \left[(\xi_r)^2 - 1 \right] \left[[A - \langle A \rangle_0]^2 - \langle [A - \langle A \rangle_0]^2 \rangle_0 \right] + \dots \right\} |\varphi_0\rangle \quad (8.79)$$

8.3.2.c Wiener process: stochastic differential equation

We now assume that the system undergoes a continuous series of measurements separated by a time interval δt , with measurement apparatuses that are all identical and all in the same initial state $|\Phi_0(x_0)\rangle$. Since we wish to obtain a continuous process, we assume that two parameters go simultaneously to zero: the time interval δt and the parameter $\varepsilon = g\tau a/\sigma$ characterizing the perturbation introduced by each measurement. At this stage, in order to introduce the appropriate continuous limit, we recall a few properties of Brownian motion.



Brownian motion Consider a particle moving on an axis Ox by random jumps occurring constantly with a time interval δt between them. Each jump changes the position by an amount $\pm \delta l$, the two opposite values having the same probability $1/2$. We are interested in the continuous limit where both δl and δt tend to zero. We consider a fixed time interval dt , which we divide into N smaller intervals $\delta t = dt/N$, corresponding to times $t_0, t_1, \dots, t_r, \dots, t_N$ at which jumps $\delta x_r = \pm \delta l$ occur. Each jump is characterized by the dimensionless variable $\xi_r = \delta x_r / \delta l = \pm 1$.

We first assume that the ratio $\delta l / \delta t$ keeps a constant value:

$$\frac{\delta l}{\delta t} = c \quad (8.80)$$

Since, during the time interval dt , the particle makes jumps in both directions with equal probabilities, the average $\overline{dx}$ of the variation of its position x vanishes:

$$\overline{dx} = 0 \quad (8.81)$$

Moreover, the average of the square of this variation is given by:

$$\overline{dx^2} = \overline{[\xi_0 + \xi_1 + \dots + \xi_N]^2} (\delta l)^2 = N (\delta l)^2 \quad (8.82)$$

This is because, in the square of the sum, all cross terms containing the product of two different ξ_r have zero average value, so that only N square terms remain. The mean-square deviation of the distance covered by the particle during dt is therefore:

$$\sqrt{\overline{dx^2}} = \sqrt{N} \delta l = \sqrt{\frac{dt}{\delta t}} \delta l = c \sqrt{dt \delta t} \quad (8.83)$$

which tends to zero in the limit $\delta t \rightarrow 0$. We then see that, in the continuous limit where $\delta l / \delta t$ is constant, the particle does not move anymore; the infinitesimal distance covered by each jump is too small.

We now assume that the ratio $(\delta l)^2 / \delta t$ keeps a constant value in the continuous limit:

$$\frac{(\delta l)^2}{\delta t} = D \quad (8.84)$$

The average value of $\overline{dx}$ is still equal to zero, as in (8.81), but now we have:

$$\overline{dx^2} = N (\delta l)^2 = \frac{dt}{\delta t} (\delta l)^2 = D dt \quad (8.85)$$

During a time interval dt , the particle now covers an average distance $\sqrt{\overline{dx^2}}$ proportional to $\sqrt{dt}$. This regime is called a Brownian motion regime.

Note that, for the sake of simplicity, we have assumed that the jumps take only two opposite values $\pm\delta l$ so that $\xi_r = \pm 1$. However, the same results hold if the jumps take a continuous range of values, in other words if ξ_r is any stochastic variable with zero average value and unit mean-square deviation. For a more detailed study of this random motion and its various applications in physics, see for instance [340].

The square of the displacement dx^2 is a random variable with significant fluctuations. But if, instead of considering the square of the sum of the ξ_r as in (8.82), we introduce the sum of the squares of the ξ_r :

$$ds = \left[(\xi_0)^2 + (\xi_1)^2 + \dots + (\xi_N)^2 \right] (\delta l)^2 \quad (8.86)$$

we obtain another variable with an average value:

$$\overline{ds} = N (\delta l)^2 = \frac{dt}{\delta t} (\delta l)^2 = D dt \quad (8.87)$$

If $\xi_r = \pm 1$, this variable does not fluctuate, and has no random character at all. If ξ_r takes a continuum of values, the fluctuations of ds are given by $\sqrt{N} (\delta l)^2 = D\sqrt{\delta t dt}$, which tends to zero in the limit $\delta t \rightarrow 0$. Although initially defined as a random function, ds actually becomes a deterministic function varying linearly in time in the continuous limit. Relation (8.87) will be useful in what follows.

If one considers an ensemble of particles defining a statistical distribution $\rho(x, t)$ at time t and each undergoing a Brownian motion, one can show that the evolution of this distribution obeys the “diffusion equation”:

$$\frac{d}{dt} \rho(x, t) = D \frac{d^2}{dx^2} \rho(x, t) \quad (8.88)$$

which has applications in many domains of physics (heat conduction, transport theory, etc.).

Stochastic evolution of the state vector We now apply the previous considerations to the evolution of the state vector. For the sake of simplicity, we assume that system S evolves only under the effect of its coupling with the series of measurement apparatuses. We also assume that the change of state written in (8.79) occurs at small time intervals δt , and that the constant $g\tau/2\sigma$ characterizing the size of the jump of the state vector is related to δt by a relation:

$$\left(\frac{g\tau}{2\sigma} \right)^2 = D \delta t \quad (8.89)$$

This relation is similar to (8.84), with $g\tau/2\sigma$ playing the role of δl in Brownian motion. At the end of the calculation, we will take the limit $\delta t \rightarrow 0$.

By analogy with the position undergoing steps $\xi_r \delta l$, we define the variations of a random function W by:

$$\delta W_r = \xi_r \frac{g\tau}{2\sigma} = \xi_r \sqrt{D \delta t} \quad (8.90)$$

If a time interval dt is split into N smaller intervals δt , the variation dW of W during dt is given by:

$$dW = W(t + dt) - W(t) = \sum_{r=1}^N \delta W_r \quad (8.91)$$

In the limit where δt and the coupling constant g tend to zero, at constant D , W is not a regular function. It is actually a highly singular function, with a derivative that is always infinite. W is an example of what is called a “Wiener process”. We briefly come back on the properties of Wiener processes in §11.

The sum of the squares of the variations δW_r has the same properties as the function ds introduced in (8.86), which is not stochastic but merely equal to $D dt$:

$$\sum_{r=1}^N (\delta W_r)^2 = D dt \quad (8.92)$$

We can now use (8.79) to express the variation $|\delta\varphi\rangle_r = |\widehat{\varphi}'\rangle - |\varphi\rangle_0$ of the state vector $|\varphi\rangle$ of S during one single infinitesimal measurement process, in terms of the Wiener process W :

$$\begin{aligned} |\delta\varphi\rangle_r = & \left\{ \delta W_r [A - \langle A \rangle_0] - \frac{1}{2} (\delta W_r)^2 [A - \langle A \rangle_0]^2 \right. \\ & \left. + \left[(\delta W_r)^2 - \left(\frac{g\tau}{2\sigma} \right)^2 \right] \left[[A - \langle A \rangle_0]^2 - \langle [A - \langle A \rangle_0]^2 \rangle_0 \right] + \dots \right\} |\varphi\rangle \end{aligned} \quad (8.93)$$

Consider a sufficiently short time interval dt during which the evolution of $|\varphi\rangle$ is very small. We sum the variations $|\delta\varphi\rangle_r$ during the N smaller intervals δt to obtain $|d\varphi\rangle$. The sum of the terms linear in δW_r introduces a term in dW , as in (8.91). The quadratic term in δW_r of the first line of (8.93) gives a term in $D dt$, according to (8.92). As for the quadratic term in the second line, it gives zero when relations (8.92) and (8.89) are inserted.

Finally, we take the continuous limit where both δt and $g\tau/2\sigma$ tend to zero at constant D in (8.89) and obtain:

$$|d\varphi\rangle = \left\{ dW [A - \langle A \rangle] - \frac{1}{2} D dt [A - \langle A \rangle]^2 \right\} |\varphi\rangle \quad (8.94)$$

with¹⁰:

$$\langle A \rangle = \langle \varphi | A | \varphi \rangle \quad (8.95)$$

Note that this stochastic differential equation is significantly different from any usual differential equation, the Schrödinger equation for instance: the term in dW is stochastic and singular, with an amplitude proportional to $\sqrt{dt}$ instead of dt as usual. We now understand why a second-order calculation with respect to ε was necessary: the term in $D dt$ arises from the square of the variations $(\delta W_r)^2$, which is proportional to ε^2 , and would have been missed by a first-order calculation only¹¹. This term is called an “Ito term”, and the corresponding sum an “Ito integral” [341]. Although the stochastic term in dW is much larger than the Ito term at a given time, it is also stochastic and may have opposite effects on the evolution of the state vector at successive times. By contrast, the Ito term is non-stochastic and has cumulative effects, so that they are not negligible in the long run when compared to the much larger stochastic terms. Note also that equation (8.94) is non-linear, since $\langle A \rangle$ depends on the state $|\varphi\rangle$, which may lead to evolutions that are very different from a usual Schrödinger evolution. As expected, we check that $|d\varphi\rangle$ vanishes if $|\varphi\rangle$ is any eigenstate of A with eigenvalue a_n : in this case all successive measurements provide the same result $x_r = x_0 + g\tau a_n$, and the process is no longer stochastic (actually no evolution takes place). This type of non-linear stochastic differential equation has been studied by Gisin [342, 343] and discussed in the context of quantum measurements [344]. For introductory text on stochastic evolutions of the state vector and continuous measurements, see for instance [345, 346].

Properties of the Wiener process Since ξ_r is a stochastic variable with zero average, the definition (8.90) implies that the statistical average of the variations of W vanishes:

$$\overline{\delta W_r} = 0 \quad \text{or} \quad \overline{dW} = 0 \quad (8.96)$$

(the upper bar denotes an ensemble average over many realizations of the process). Since the mean-square deviation of ξ is 1, we also have:

$$\overline{[\delta W_r]^2} = D\delta t \quad (8.97)$$

Moreover, since the successive results of measurements result from successive quantum processes, which are all independent and fundamentally random, the

¹⁰ Using (8.92), we can check that $(\langle \varphi | + \langle d\varphi |) (|\varphi\rangle + |d\varphi\rangle) = \langle \varphi | \varphi \rangle + 0(dt^2)$, in other words that (8.94) conserves the norm of the state vector.

¹¹ A second-order calculation is indeed sufficient since third-order terms give a contribution in $dt^{3/2}$, playing no role in the limit $dt \rightarrow 0$.

stochastic variables $x_r - x_0$ are independent for different values of r . The same is therefore true of the ξ_r and of the δW_r . Combining this property with (8.97), we obtain:

$$\overline{\delta W_r \delta W_{r'}} = \delta_{r,r'} D \delta t \quad (8.98)$$

where $\delta_{r,r'}$ is a Kronecker delta symbol.

We can also introduce the time derivative $W'(t)$ of the Wiener process. This derivative may be defined as a random process obtained by dividing the variation of W by the time interval:

$$W'_r = \frac{\delta W_r}{\delta t} = \xi_r \sqrt{\frac{D}{\delta t}} \quad (8.99)$$

In the continuous limit, we then have:

$$\overline{W'(t) W'(t')} = D \delta(t - t') \quad (8.100)$$

where $\delta(t - t')$ denotes the Dirac delta function of $(t - t')$. This equality can be checked by remarking that, if $t' \neq t$, the average of the product of two W'_r vanishes for different values of the index r (they contain independent random variables ξ_r). Moreover, integrating (8.100) over t' introduces an integral that is the continuous limit of the discrete sum:

$$\sum_{r'} \delta t \left[\frac{1}{\delta t^2} \overline{\delta W_r \delta W_{r'}} \right] = \sum_{r'} \delta t \frac{1}{\delta t^2} \delta_{r,r'} D \delta t = D \quad (8.101)$$

Therefore, $\overline{W'(t) W'(t')}$ is entirely concentrated in an infinitesimal interval around $t' = t$ and has an integral equal to D , which defines a delta function multiplied by D . Equation (8.100) expresses that the time variations of the Wiener process have no memory. They are what is often called a “white noise” in physics, with completely independent values at different times, even if the times are almost equal.

Now, instead of writing the evolution of the state vector in terms of an unknown Wiener process as in (8.94), it may seem more natural to express it as a function of the results of measurements, which are directly observable. In a given run of the experiment, what is actually obtained is a whole string of results $(x_r - x_0)$ – it is convenient here to take x_0 as an origin for expressing the results. By analogy with (8.73), we define the stochastic variable ζ_r by:

$$\zeta_r = \frac{x_r - x_0}{\sigma} = \xi_r + \frac{\Delta x_0}{\sigma} \quad (8.102)$$

and, to obtain a proper continuous limit as in (8.90), we define the stochastic function R by:

$$\delta R_r = \zeta_r \frac{g\tau}{2\sigma} = \delta W_r + \frac{g\tau}{2\sigma} \frac{\Delta x_0}{\sigma} \quad (8.103)$$

R is called the “measurement record” for each realization of the series of measurements. Using expression (8.71) as well as (8.89), we then obtain:

$$\delta W_r = \delta R_r - 2 \left(\frac{g\tau}{2\sigma} \right)^2 \langle A \rangle = \delta R_r - 2 \langle A \rangle D \delta t \quad (8.104)$$

and, inserting this result into (8.94) and going to the continuous limit:

$$|d\varphi\rangle = \left\{ dR [A - \langle A \rangle] - \frac{1}{2} D dt \left[[A + \langle A \rangle]^2 - 4 \langle A \rangle^2 \right] \right\} |\varphi\rangle \quad (8.105)$$

This equation provides the evolution of the state vector as a function of the measurement record R .

Stochastic evolution of the density operator We finally study the evolution of the density operator $\rho = |\varphi\rangle \langle \varphi|$. Its infinitesimal variation is given by:

$$\delta (|\varphi\rangle \langle \varphi|)_r = [|\varphi\rangle + |d\varphi\rangle_r][\langle \varphi| + \langle d\varphi|_r] - |\varphi\rangle \langle \varphi| \quad (8.106)$$

where $|d\varphi\rangle_r$ is given by (8.94) and $\langle d\varphi|_r$ is the corresponding bra. Usually, the term in $|d\varphi\rangle_r \langle d\varphi|_r$ can be ignored to first order, since it is in dt^2 . Here the situation is different since, in terms of the Wiener process, $|d\varphi\rangle$ contains a term in dW_r , with a square proportional to δt as shown by (8.92). This term then introduces a contribution:

$$D dt [A - \langle A \rangle] \rho [A - \langle A \rangle] \quad (8.107)$$

When this term is added to the linear terms contained in $|d\varphi\rangle_r$ and the associated bra, the evolution of ρ is obtained in the form:

$$d\rho = \left\{ dW ([A, \rho]_+ - 2 \langle A \rangle \rho) - \frac{1}{2} D dt [A, [A, \rho]] \right\} \quad (8.108)$$

containing an anticommutator $[A, \rho]_+$ and a double commutator $[A, [A, \rho]]$. This equation is sometimes called a Belavkin equation [347]. As above for the state vector, one can replace dW by $dR - 2 \langle A \rangle D dt$ to express the variation of the density operator as a function of the measurement record, instead of the Wiener

process. We can easily check that if ρ is a statistical mixture of projectors over eigenstates of A , since it commutes with A , no evolution takes place under the effect of the successive measurements, for the same physical reasons as above.

The mathematical tools we have discussed in this chapter (Wiener processes, stochastic differential equations, etc.) turn out to be useful in some interpretations of quantum mechanics we discuss in Chapter 10, in particular those involving a stochastic Schrödinger dynamics.

Experiments: quantum reduction seen in real time

From a fundamental point of view we must acknowledge that, since about 1935, our conceptual understanding of quantum mechanics has not progressed so much. Really new ideas are few and far between – except of course the major line initiated by the contribution of Bell [6]. This is in big contrast with the rest of physics, where new theoretical and experimental discoveries in many fields have flourished, very often with the help of the tools of quantum mechanics. The fantastic evolution of the experimental techniques has completely changed the situation. At the beginning of quantum mechanics, the observation of the tracks of single particles in Wilson chambers [348] played the essential role in the introduction of the postulate of state vector reduction, but otherwise it was impossible to observe continuously a single electron, atom, or ion; the experiments that theorists were proposing in discussions on the principles of quantum measurement were therefore “thought experiments” (“Gedanken Experiment”), as for instance in the famous Solvay meetings [1, 21]. But nowadays, after almost of century of experimental progress, experiments that were then unthinkable have become a reality.

A huge number of contemporary experiments involves the laws of quantum mechanics in general; several books would not be sufficient to describe all of them. Nevertheless, in the majority of experiments, what is really observed is the sum over a very large number of particles of one individual microscopic observable (sum of atomic dipoles for instance), which is accurately described by the average value of this observable. The Schrödinger equation can then be used to calculate this average, which is subsequently treated as if it was a classical variable. No particular use of the Born rule or of the projection postulate is then necessary – a typical illustration is given by many NMR (nuclear magnetic resonance) experiments in chemistry and physics. Similarly, in many coincidence experiments, the relevant physical observable is the product of quantum operators associated to individual counting rates, and the Schrödinger equation can be used again to obtain the average of this product. Of course, this does not necessarily mean that the projection postulate is

completely irrelevant for these experiments! In quantum optics and atomic physics for instance, optical detection with photomultipliers and diodes is often used; the projection postulate then determines the size and nature of the random “shot noise” observed in the experiments. This noise limits the accuracy by adding a fluctuating component to the signal, while the latter itself has a smooth deterministic behavior as a function of the experimental parameters; since very often what is studied in detail is the signal rather than the noise, the projection postulate then plays only an auxiliary role.

Here we will focus our interest only on the small fraction of experiments where the effect of the projection postulate is particularly visible, for instance those where a single quantum particle is observed, and where “quantum jumps are visible in real time”. The observations then exhibit more quantum features than what is predicted by the continuous Schrödinger evolution, and in this sense go beyond this equation. Our purpose is not to give a complete review, but just to provide a few particularly illustrative examples.

9.1 Single ion in a trap

It is possible to observe a single Barium ion contained in a radio-frequency trap by monitoring its fluorescence under laser irradiation [349, 350]; the lifetime of the excited resonance level is about 10^{-9} second, meaning that even if only a thousandth of the fluorescence light is collected (as typical in these experiments) one can still detect a flux of a million of photons per second, directly visible with the naked eye. Under these conditions, it is possible, as the authors of [349] write, to directly “watch the reduction of the wave function by the measurement process on the oscilloscope screen”.

The relevant energy levels of a Barium ion used in the experiment performed by Dehmelt and coworkers is shown in Figure 9.1: the ground state is g , two excited states are e_1 and e_2 , and m_1 and m_2 are two metastable states. The transition $g - e_1$ from the ground state g to a first excited state e_1 is strongly illuminated with a laser, while the fluorescence at the corresponding wavelength is also constantly monitored. Another laser excites the transition from e_1 to a first metastable level m_1 , so that the ion cannot be trapped in this metastable level. The role of this second laser is to create a $g - e_1 - m$ closed circuit from which the ion cannot escape, and which is also used for laser cooling of the ion. If no other excitation of the ion was used, it would constantly fluoresce. But another, much weaker, light source excites the $g - e_e$ transition to a second excited level e_2 . Sometimes, when the ion reaches level e_2 , it does not fall back into the ground state g , but into the metastable level m_2 , where it can no longer be excited optically; it then stops fluorescing, and one then says that it has been “shelved” in the metastable state. Nevertheless, since

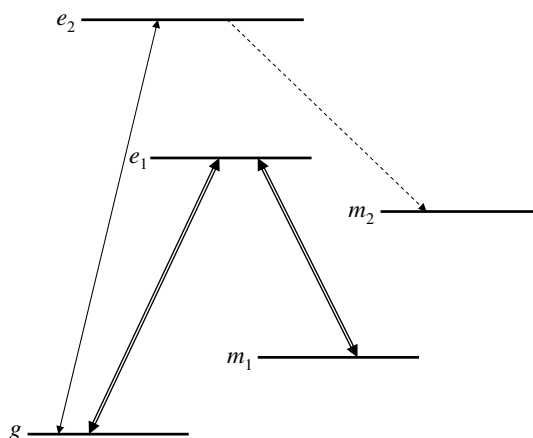


Figure 9.1. Energy levels of the Barium ion involved in the experiment. The ground state of the ion is g , e_1 and e_2 are two excited states, and m_1 and m_2 two metastable states. Two lasers of high intensity excite the transitions $g - e_1$ and $m_1 - e_1$ for laser cooling of the ion motion in the trap; the fluorescence emitted by spontaneous emission from level e_1 to g is constantly monitored. The transition $g - e_2$ is excited only weakly with a lamp; when the ion reaches e_2 , it sometimes falls (dotted line) into the second metastable level m_2 . It is then “shelved” and no longer fluoresces, until it falls spontaneously into the ground state and starts fluorescing again.

this level has a finite lifetime, the ion eventually falls back spontaneously into the ground state, and starts fluorescing again.

If one applies the Schrödinger equation to this situation, one has to include in the quantum system the ion itself as well as the electromagnetic field, with several modes populated, and treat the effect of interactions (absorption, stimulated emission, spontaneous emission). The solution of the equation that is obtained in this way is, after some time, a superposition of two components: one where the ion is not shelved and strong spontaneous radiation is emitted, and another component where the ion is shelved and emits no radiation at the frequency of the $e_1 - g$ transition. The ion fluoresces and does not fluoresce at the same time, as the Schrödinger cat is alive and dead; the state vector provides an average fluorescence intensity that is intermediate between the two situations. In other words, within the Schrödinger equation, everything remains continuous, and “jumps” never occur.

But, since in the experiment the fluorescence intensity is continuously measured, the two components of the state vector also contain macroscopically different states of the measurement apparatus, and therefore the projection postulate applies: a fundamentally random process takes place, and spontaneously the system chooses one of the components only. Consequently, the fluorescence has either its maximal

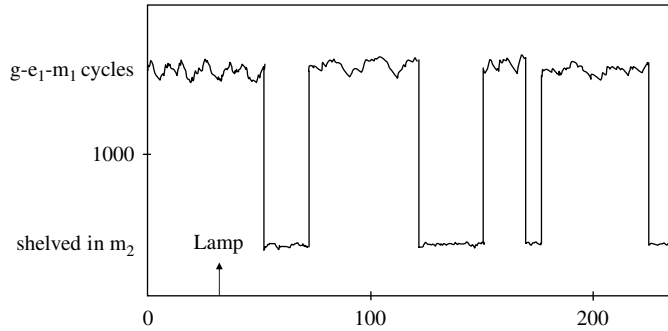


Figure 9.2. Fluorescence signal of a single ion at the wavelength of the transition $g - e_1$ as a function of time t in the experiment of [349]; the horizontal scale is seconds, and the vertical scale gives the number of counts per second delivered by a photomultiplier. Until $t = 30$ s., the weak lamp remains off, so that the ion cycles between levels g , e_1 , and m_1 and constantly fluoresces (upper noisy line). The weak lamp is then switched on and may transfer the ion to metastable level m_2 , where it stops emitting light. When this happens, the ion remains shelved for some time in the metastable level, and a dark period occurs (lower line) until the ion spontaneously returns to the ground state. The whole cycle then starts again from the beginning.

value (that corresponding to the ion circulating between the three levels g , e_1 , and m_1), or zero; it does not vary continuously between these values.

The experimental results fully confirm this prediction: as shown in Figure 9.2, the observed fluorescence undergoes “jumps” between a fixed value and zero. The times at which these jumps occur are not controlled experimentally, since they are a consequence of the fundamentally random character of quantum measurement. Consequently, if many ions were fluorescing at the same time, each of them would switch between the fluorescing and the dark regime at different times, and on average one would observe only the average value of the fluorescence – the value predicted by the continuous Schrödinger equation. But, with a single ion the quantization of the signal is directly observed, and the system can undergo a very large number of jumps between the two values.

Experiments performed at the same time in Toschek’s group [350] on the same ion have provided similar results in a case where only three levels are involved, also exhibiting very clear “quantum jumps”. Similar experiments with other ions were performed in other laboratories, in particular with the Mercury ion [351, 352]. In this case, three atomic levels are relevant, the ground state, a resonance excited level $^2P_{1/2}$ with short lifetime (2 ns.), and a metastable level $^2D_{5/2}$ with lifetime 0.1 s. When the transition between the ground state and resonance level is constantly excited, the ion fluoresces with no interruption. But when the transition to the metastable states is also excited, sometimes the ion is “shelved” in this

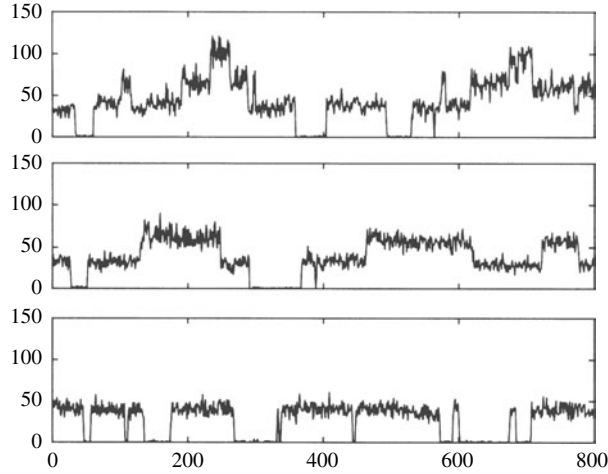


Figure 9.3. Lowest curve: fluorescence signal of a single Hg^+ ion at the resonance wavelength corresponding to the transition between the $^2\text{S}_{1/2}$ ground state of the ion and the $^2\text{P}_{1/2}$ excited state. The units of the horizontal axis are ms, and the vertical axis gives number of counts per ms. When the transition to a metastable $^2\text{D}_{5/2}$ level is simultaneously excited, quantum jumps are clearly visible, as in Figure 9.2. The middle curve is obtained when two ions sit in the trap; three cases may then occur, depending on the number (0, 1, or 2) of ions “shelved” in the metastable state, and resulting in three possible fluorescence intensities. The upper curve shows the fluorescence when three ions are trapped, and provide four possible levels of fluorescence at the resonance transition (figure kindly provided by D. Wineland and W. Itano).

state and stops fluorescing until it falls back to the ground state by spontaneous emission. Results are shown in Figure 9.3. As in Figure 9.2 the fluorescence of the ion exhibits discontinuities between two regimes, often called “quantum jumps” as a reminiscence of the historical Bohr theory of atoms (§1.1.1).

Long before these results were obtained, Schrödinger had analyzed a thought experiment where the light emitted by an atom was used in an interference experiment [353]; he had pointed out that quantum jumps, if they occur when an atom emits photons, cannot be instantaneous, but necessarily have a finite duration related to the radiative width of the atomic levels. Here, the situation is similar, since the change from the dark regime (no fluorescence) to the bright regime (fluorescence) is triggered by the spontaneous emission from level e_2 to m_2 , while the transition from bright to dark is triggered by the transition from m_2 to the ground state g – in other words two “jumps” play a role, instead of one, but Schrödinger’s remark applies as well.

Then, how is this compatible with the observations of Figure 9.2? The explanation is that the signal of Figure 9.2 is averaged over time, so that it does not show

the discontinuities that would appear under a more precise observation. At a much smaller time scale, the signal provided by the photomultiplier monitoring the fluorescence is a series of sudden “clicks” occurring at random times, corresponding to the detection of individual photons; Figure 9.2 actually shows nothing but the averaged frequency of these clicks. What is really observed is transitions between periods where the clicks occur frequently and others where they are rare; the transition time between these periods cannot be measured more precisely than the time between two consecutive clicks. As a consequence, the time at which the “jumps” occur can only be measured within some uncertainty, related to the emission rate and therefore to the radiative width of the levels, as implied by Schrödinger’s argument – for a more precise discussion see [354]. A more detailed theoretical study of the fluorescence intermittence phenomenon can be found in [355, 356].

9.2 Single electron in a trap

A single electron orbiting in a cyclotron trap can also give rise to observations reaching the quantum limit, as demonstrated by Gabrielse *et al.* [357]. In their experiment, a cold Penning trap was used to store one electron. Such a trap has a strong homogenous axial magnetic field and a quadrupole electric field created by electrodes. In the magnetic field, the quantized levels of a charged particle are the equidistant orbital “Landau levels” with energy $n\hbar\nu_c$, where ν_c is the cyclotron frequency and n an integer (the energy results from a combined effect of the coupling of the magnetic field with the orbital and spin variables). What is measured is the quantum number n characterizing this energy.

Two different motions of the electron in the trap are actually relevant in this experiment. The main motion is the usual cyclotron motion, perpendicular to the magnetic field, which is the central focus of the experiment. But there is also a longitudinal motion, parallel to the field, which has a frequency that depends weakly on the energy of the cyclotron motion (experimentally, this coupling is introduced by a small distortion of the homogeneous magnetic field). It is therefore possible to have access to the cyclotron energy by a very precise measurement of the frequency of longitudinal oscillations of the electron, which gives access to n . In the experiment, a single electron orbits in the magnetic field, while its longitudinal frequency is constantly monitored with great accuracy, which amounts to a constant quantum measurement of n . Figure 9.4 shows an example of the variation in time of this frequency: it clearly exhibits steps occurring when the system is projected into an energy eigenstate under the effect of quantum measurement. The general comments of §9.1 apply to this experiment as well, which provides another illustration of a “quantum non-demolition” measurement monitored in real time. More recently, a new version of the cyclotron experiment has provided an

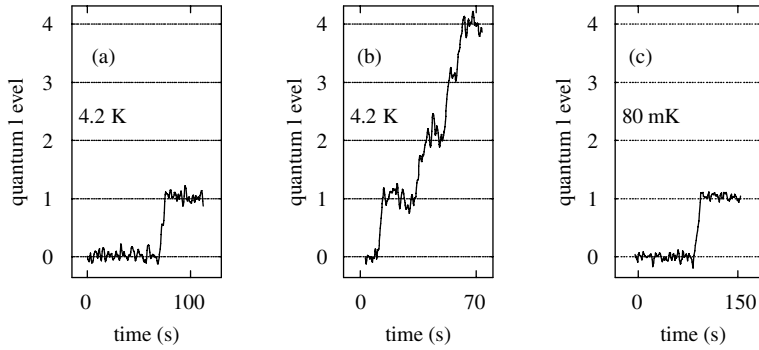


Figure 9.4. Real time detection of the quantum orbital number of a single electron in a Penning trap. In figures (a) and (b), the excitation between the Landau levels is obtained under the effect of blackbody radiation in the experiment, at a few degrees Kelvin. In figure (c), the temperature is much lower so that this process does not occur anymore; the excitation of the electron towards excited Landau levels is then obtained by applying a microwave field (figure kindly provided by G. Gabrielse).

extremely accurate measurement of the magnetic moment of the electron and of the fine structure constant [358].

9.3 Measuring the number of photons in a cavity

In the experiments described in the two preceding sections, the observed quantum system was a material particle (particle with rest mass), either an ion or an electron. One could think that photons are less appropriate to be observed individually, since they are so easily absorbed in various processes, including their detection with photomultipliers or diodes. But recent experiments have obtained comparable results by observing photons in a cavity, illustrating the properties of quantum measurement in a way that is even more spectacular. Until relatively recently, the only way to detect a photon was to absorb it in a detector (photomultiplier for instance), so that further measurements on the same particle were impossible. But various “quantum non demolition” methods now make it possible to measure the presence and the number of photons without destroying them [319, 359] – see also §6.2 of [264]. The combination of this possibility with the methods of cavity electrodynamics has resulted in experiments where the number of photons stored in a cavity is constantly monitored in real time [360].

Rydberg atoms are atoms in high energy levels, Rydberg levels, very close to the ionization threshold. Among them, the “circular Rydberg levels” (those where the rotation quantum number l takes its maximal value) interact with photons in a particularly simple and well-controlled way. These atoms have a very large dipole

moment, which means that they are strongly coupled to the electromagnetic field; they can therefore be used as probes of its properties, even if it contains very few photons. Moreover, they can be ionized and detected with high efficiency and selective access to the various Rydberg levels. Current experimental techniques allow the production of beams of such atoms, which can be sent through a resonant electromagnetic cavity to test the number of photons it contains.

One could make use of the absorption of photons by the atoms, by sending atoms one by one through the cavity, controlling their energy level before the crossing of the field and measuring the energy level at the output – but obviously this process would absorb photons and not lead to a quantum non-demolition measurement. A much better technique is to choose a case where the frequency of the photons in the cavity differs significantly from the resonance frequency of the atoms (Bohr frequency associated with the transition between the two relevant Rydberg levels), so that the absorption probability of photons remains negligible. Due to non-resonant interactions between atoms and photons, the atoms create index (dispersive) effects for the photons and, conversely, the photons shift the atomic energy levels; one can show that the frequency of the transition between the two atomic levels is increased by an amount that is proportional to the number of photons in the cavity.

The idea is then to send into the cavity atoms that are in a coherent superposition of two Rydberg states, let them interact with the photons inside the cavity, and at the output measure the change of phase of the coherent superposition induced by this interaction. Since the phase change is proportional to the number of photons in the cavity, one obtains in this way a measurement of their number. This purely dispersive technique is non-destructive: it changes neither the energy of the electromagnetic field (that is the number of photons) nor the energy of the atom at the output of the cavity. It provides quantum non-demolition measurements (§8.1.1.b). The experiment is shown schematically in Figure 9.5.

One major experimental difficulty is that a sufficient number of atoms must cross the cavity in order to give a reasonably accurate measurement, before the electromagnetic field has decayed by absorption in the walls of the cavity. As a consequence, one has to use a cavity with an extremely high quality factor Q . This has been achieved with the help of a high quality superconducting cavity so that, in the experiment, hundreds of atoms can cross the cavity before the number of photons changes. By constantly monitoring the state of the outgoing atoms, one has access to a really continuous and non-destructive measurement of the number of photons. Initially, the cavity is fed with a coherent state that does not have a well-defined number of photons; nevertheless, as soon as a few atoms have crossed the cavity and their phase has been measured at the output of the cavity, the system is projected into a state with fixed number of photons. Then, the photons disappear

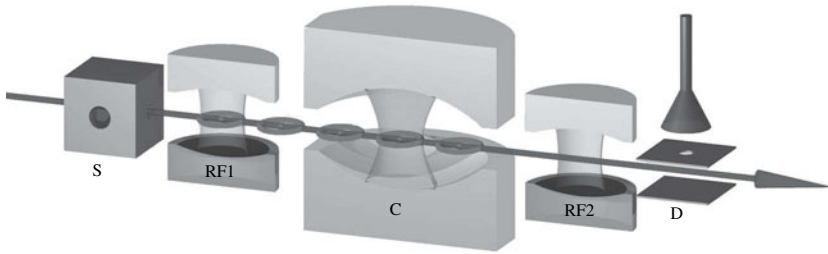


Figure 9.5. A beam of atoms in a given highly excited state (circular Rydberg level) is created by a source *S*. Each atom crosses, first a region where a device *RF1* excites coherently a transition between the initial and another close Rydberg level, then crosses a cavity *C* with resonant frequency shifted with respect to the atomic transition, and a region where another device *RF2* (locked in phase with *RF1*) excites again the same transition. Devices *RF1* and *RF2* perform what is called Ramsey spectroscopy on the atom: the first interaction creates a coherent dipole associated with the two relevant Rydberg levels and the second detects the phase of this dipole, including its evolution between the two regions and the effect of cavity *C*. Since the dipole accumulates a dephasing when it crosses *C*, and since this dephasing depends on the number of photons contained in *C*, the final population of the Rydberg level depends on this number of photons. One then has direct access to it by measuring this population in detector *D*. Since the cavity is not resonant at the atomic resonant frequency, the number of photons in the cavity is not changed by the interaction; the method therefore provides quantum non-destructive (QND) measurement (this figure has been kindly provided by J.M. Raimond).

one by one as they are absorbed in the walls of the cavity, which occurs at a slow rate since the superconducting cavity has a very high quality factor. This produces a series of quantized changes in the observed signal until no photon remains in the cavity. Two examples of recordings made in this way are shown in Figure 9.6.

In this case, the discussion concerning the time it takes to obtain this projection (see end of §9.1) can be given with particular clarity, in terms of the number of atoms that are required before one can ascertain with good probability that a jump has occurred [360]. What one observes is clearly not the continuous evolution predicted by the Schrödinger equation, which has the same solution for all identical realizations of the experiment; in each experiment, one sees marked steps, which fall at times that are different for each realization. The Schrödinger equation provides only the average of the observations over many experiments; the steps seen in single experiment cannot be explained within this equation only, but require an additional ingredient. This is another particularly spectacular case where one can see “the reduction of the state vector occurring directly under your eyes”.

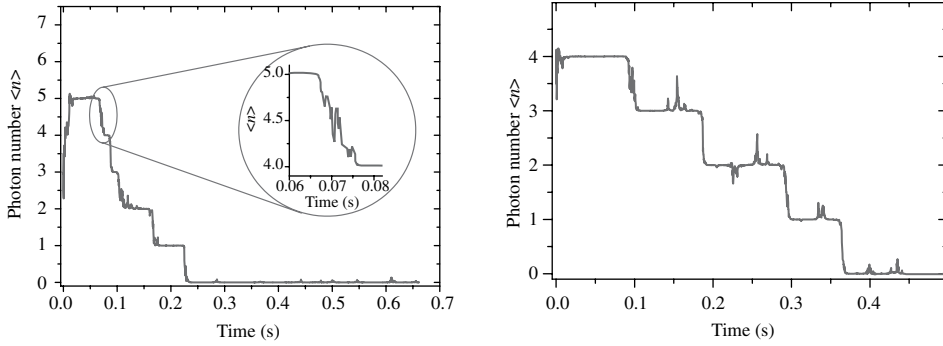


Figure 9.6. Continuous monitoring of the number of photons contained in an electromagnetic cavity, obtained by measuring the phase shift accumulated by Rydberg atoms crossing the cavity. Initially, the cavity is loaded with a coherent field of small intensity which, under the effect of quantum measurement, is projected to a state with a well-defined number of photons (five in the left part of the figure, four in the right part). Then the photons are progressively absorbed by the wall of the cavity and disappear, one by one. This process is constantly observed in real time (figure kindly provided by J.M. Raimond).

9.4 Spontaneous phase of Bose–Einstein condensates

The notion of “identical particles” plays a much more important role in quantum mechanics than in classical physics. It leads to the introduction of a special “symmetrization” postulate, which determines the possible forms of the state vector of a system of identical particles, with a distinction between two classes of particles: bosons and fermions. The latter obey the Pauli exclusion principle: two fermions can never occupy the same individual quantum state. By contrast, an arbitrary number of bosons can occupy the same quantum state, as for instance photons do in a monomode laser beam. Another example is given by Bose–Einstein condensates; in a very dilute gas at very low temperatures, such condensates can be represented by states where all N bosons are in the same individual state $|\varphi\rangle$:

$$|\Psi\rangle = |1 : \varphi\rangle |2 : \varphi\rangle \dots |N : \varphi\rangle \quad (9.1)$$

In this case, as opposed to the situation where one single particle is in state $|\varphi\rangle$ (§7.1.2), it becomes possible to determine the state, or equivalently its wave function $\varphi(\mathbf{r})$; one sometimes says that this wave function becomes analogous to a classical field. But, in classical physics, two wave packets of the same field may give rise to interference effects. By analogy, one can ask the question: if two Bose–Einstein condensates made of identical atoms are prepared independently, is it possible to observe interference effects between them? The answer to the question involves interesting quantum measurement effects, and has been clarified by Javanainen and Yoo [361] as well as by other authors (see for instance references given in [100]).

9.4.1 Interferences between independent condensates

Assume that N identical bosons occupy state $|\varphi\rangle$ and P other bosons (identical to the former) state $|\chi\rangle$. When wave functions $\varphi(\mathbf{r})$ and $\chi(\mathbf{r})$ overlap, the first idea that comes to mind is to assume that the wave functions interfere exactly as a classical field, so that the probability to find a particle at point $\mathbf{r}$ is proportional to:

$$|\varphi(\mathbf{r}) + \chi(\mathbf{r})|^2 \quad (9.2)$$

In classical optics for instance, the intensity at point $\mathbf{r}$ contains an interference terms between the electric fields of two overlapping beams; here we have crossed interference terms in $\varphi^*(\mathbf{r})\chi(\mathbf{r})$ and $\varphi(\mathbf{r})\chi^*(\mathbf{r})$. But, in quantum mechanics, the phase of each wave function is arbitrary: no physical property associated with the wave functions is changed if they are replaced by $e^{i\alpha}\varphi(\mathbf{r})$ and $e^{i\beta}\chi(\mathbf{r})$, while this obviously changes the crossed terms in (9.2). This argument shows that the quantum probability cannot be given by this classical formula. An explicit quantum calculation confirms this point; when one detects the first particle, its probability of presence is given by:

$$|\varphi(\mathbf{r})|^2 + |\chi(\mathbf{r})|^2 \quad (9.3)$$

The probability is therefore merely the sum of the probabilities corresponding to a detected particle being initially either in state $|\varphi\rangle$ or in state $|\chi\rangle$, without any interference term. As far as the detection of the first particle is concerned, the relative phase of the two condensates is therefore non-observable, because actually it does not exist: nothing in the definition of the initial state can be used to determine this phase. The same result can also be interpreted in terms of an uncertainty relation between the phase and the number of particles (similar to the time-energy uncertainty relation): if the number of particles is perfectly determined, as is the case here, the phase is completely undetermined.

But, as the authors of [361] have shown, the situation becomes more interesting when the positions of several particles are measured. For a given realization of the experiment, one can study the correlations between these positions. The quantum prediction is then that, while more and more particle positions are measured, the relative phase becomes better and better defined. Already the first position measurement creates some information on this phase¹, which plays a role for the second measurement; the phase distribution, instead of being completely independent of the phase, is given by a sinusoid having maxima and minima at positions depending on the measured position. Then, while measurements accumulate, the phase

¹ For instance, the measurement shows that the phase cannot take a value that puts a destructive interference anti-node at the measured position (we assume that the two intensities are the same).

distribution becomes a product of more and more sinusoids and exhibits a sharper and sharper maximum, making this phase better and better determined; one finally practically reaches a classical situation where the phase is perfectly determined. This provides an interesting process where, initially, the phase did not exist, but where the successive projections due to quantum measurement make it emerge, and give it a better and better defined value. Interestingly, the observations would be exactly the same if the phase had existed from the beginning for each realization, but was totally unknown. Nevertheless, from one realization of the experiment to the next, the new phase value that emerges is totally different, with no correlation with the preceding value.

Experiments performed in Ketterle's group at MIT with condensates made of Sodium atoms have confirmed these predictions [362]. Two condensates were independently prepared, and then released from their separate traps to allow them to overlap spatially. Optical absorption was then used to measure the positions of atoms in the overlap region. Figure 9.7 shows the result obtained in a given run of the experiment: indeed one observes fringes with a well-defined phase, even if quantum mechanics predicts that it did not exist before measurements; this provides a case where the state vector reduction is very directly visible. From one experiment to the next, the phase is completely uncorrelated; in other words, if one sums the results obtained in several runs, the fringes disappear, as predicted by theory.

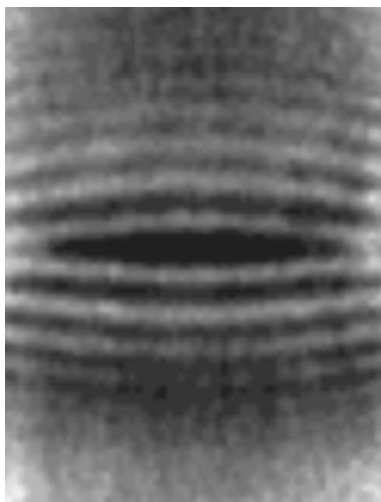


Figure 9.7. Observation of interference fringes when two Sodium condensates are prepared separately, and then released from their traps to spatially overlap. For each run of the experiment, the fringes have a well-defined phase, but the phase is completely uncorrelated from one run to the next (figure kindly provided by W. Ketterle).

A specificity of the experiment is that the physical quantity that emerges from the measurement (the phase) directly fixes the value of a macroscopic quantity (the density of the gas at each point of the interference region), which provides a direct observation of the effects of state vector projection.

9.4.2 An additional variable?

The question that then arises very naturally is the same as for any quantum measurement: should we really think, as standard quantum mechanics invites us to do, that it is the act of measurement that creates the value of the phase? Or should we consider, on the contrary, that it only reveals a value that existed before the measurement? When a variable is macroscopic, as can be the case for the phase of an interference pattern, it does not seem very natural to consider that it can be created by the measurement, due to some uncontrolled perturbation of the measurement apparatus. This raises once more the question of the existence of additional variables, this time in a macroscopic context. Moreover, this experiment has some analogy with the discussion of the Schrödinger cat, where quantum uncertainty is transferred to the macroscopic world.

Leggett and Sols [363] have discussed a similar situation, where a Josephson current appears between two superconductors; the value of this current is fixed by the difference between the two superconducting phases. These authors ask whether such a phase, with all its consequences on a macroscopic current, can really appear under the influence of a measurement apparatus, which can be very small: “Does the act of ‘looking to see’ whether a Josephson current flows itself force the system into an eigenstate of current and hence of relative phase? ... Can it really be that by placing, let us say, a minuscule compass needle² next to the system, with a weak light beam to read off its position, we can force the system to ‘realize’ a definite macroscopic value of the current? Common sense certainly rebels against this conclusion, and we believe that in this case common sense is right”.

9.4.3 Non-locality of phase

One can add another component to the argument by assuming that the interference takes place, not in a single region of space, but in two remote regions where Alice and Bob operate. The situation is then analogous to that discussed with measurements performed on two remote spins (§4.1.1), but the correlations now occur between the phases of the interference patterns instead of the directions of spins. To push

² The purpose of this needle is to measure, through its rotation, the magnetic field created by the macroscopic Josephson current.

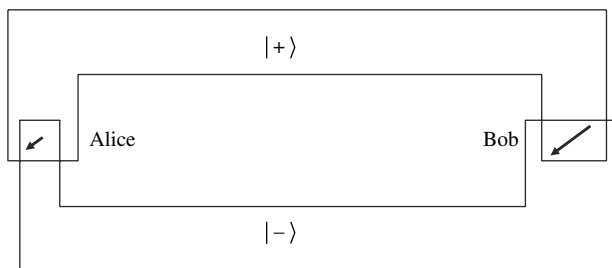


Figure 9.8. Two condensates, one corresponding to particles with spin state $|+\rangle$, the other to particles with spin state $|-\rangle$, have wave functions (represented in the figure) that overlap in two regions of space. In these two regions, two operators Alice and Bob measure a transverse component (in the plane that is perpendicular to the quantization axis) of the spins of the particles. Initially, the relative phase of the two condensates is completely undetermined, so that Alice's first measurement gives a completely random result. But, when she accumulates measurements, their effect on the quantum system is to determine the relative phase better and better; after some time, the phase is practically fixed, and one can say that the system has given transverse spin polarization in the overlap regions.

What is surprising is that, under the effect of Alice's measurements, the same direction of spin polarization appears in Bob's region, whatever the distance between the two regions is and without any mutual interaction. This is very similar to the usual situation of the EPR argument, but now with macroscopic EPR elements of reality if the number of particles in each region is macroscopic.

the analogy further, it is more convenient to assume that the condensates are spin condensates, with a relative phase controlling the average value of a transverse³ spin component in all overlap regions. The corresponding situation is schematized in Figure 9.8: two condensates, one containing spin $+$ particles, the other spin $-$ particles, extend in space and overlap in two remote regions, where Alice and Bob perform measurements of the transverse components of the spins. The predictions of quantum mechanics in such a case are a direct generalization of what was described in §9.4.1: when Alice performs a first measurement, the result is completely random; but, while measurements are accumulated, the successive results make a relative phase emerge with a better and better definition – in other words a transverse spin polarization emerges in a well-defined direction. After a sufficient number of results, the results of spin measurements along this direction are practically certain. If the overlap region in which Alice operates contains a macroscopic number of particles, the emerging spin polarization is macroscopic.

Still, according to standard quantum mechanics, another effect of Alice's measurements is the appearance of a parallel orientation in Bob's region, without any time of propagation of the influence between the two regions. But is it really possible

³ Here "transverse" means: in the plane perpendicular to the quantization axis defining the initial condensates.

that a measurement performed by Alice on a small number of spins, a hundred for instance, will create an arbitrarily large spin orientation in Bob’s region, at an arbitrary distance, especially if the distance between the two regions is extremely large? The question is strongly reminiscent of the EPR argument, but in a case where the “elements of reality” are macroscopic quantities [100], which makes the argument even stronger. Can we really imagine that a macroscopic spin orientation is created in Bob’s laboratory without any local interaction, and appears (so to say) from nothing? Or should we follow EPR and say that this orientation existed from the beginning of the experiment, so that the measurements do nothing but reveal the existence of a phase that was initially fixed at a random value? If so, the phase becomes an additional variable that should be inserted into quantum mechanics to make it complete. Since the discussion deals with macroscopic quantities, which presumably are directly accessible to human experience, it is difficult to guess precisely what Bohr would have replied to this form of the argument.

Moreover, the conservation of angular momentum raises delicate questions in such a case. How can the angular momentum contained in Bob’s region be changed instantaneously as a result of measurements performed in an arbitrarily remote region⁴? Of course, if one considers only the region where Alice performs her experiment, no particular difficulty occurs. The measurement apparatus she uses, in order to be able to measure the angular momentum of the particles, must interact with them with a coupling Hamiltonian that contains their angular momentum, as well as its own angular momentum; this makes a transfer between them possible. One can then assume that the transfer depends on the result of measurement in a way that ensures perfect angular momentum conservation. Any paradox is avoided with the help of, so to say, a recoil effect of the measurement apparatus. But the coupling Hamiltonian in question certainly commutes with all operators associated with physical quantities localized in Bob’s region; its action cannot change angular momentum in that region of space. It is therefore difficult to understand how Alice’s measurement apparatus can create such an angular momentum at a distance. Moreover, if Alice performs her measurement on a small number of spins it seems paradoxical to consider, in order to conserve total angular momentum, that her apparatus gets a recoil effect that corresponds to the very large number of spins in Bob’s region. But if we decided, rather, that her measurement apparatus cannot get a larger recoil than the maximal angular momentum allowed by quantum for a small number of spins, then we would have to give up angular momentum conservation. Here also, it is tempting to follow EPR and to avoid any difficulty by considering that the angular momentum in Bob’s region already existed before

⁴ Here we discuss a single realization of the experiment. With a large number of measurements, the angular momentum may take all transverse directions, so that its average vanishes; no paradox then occurs.

Alice's measurement, which amounts to complete quantum mechanics with the addition of a phase.

Even if one modifies the theory to include this phase, this is not sufficient to obtain a completely classical behavior and to restore locality in the quantum predictions. Actually, the situation is relatively similar to the usual case with two spins, where the Bell theorem forces the EPR elements of reality to have a non-local evolution if one wishes to reproduce the predictions of quantum mechanics in all cases. Here also, when the relative phase is measured in two different regions, a model where this phase existed from the beginning (while being unknown) cannot reproduce all predictions of quantum mechanics [102]. This impossibility appears when the spin components along several different directions are measured, and takes the form of BCHSH inequalities that are violated by quantum mechanics. In addition, and in opposition to what one could have expected, these violations do not tend to zero for systems with a very large number of particles⁵, but remain constant.

Indeed, Fock states (or Bose–Einstein condensate) provide interesting transpositions of the EPR argument and the Bell theorem to macroscopic systems; this sheds a new light onto the arguments.

⁵ The angles between the directions of the spin components to be measured become smaller and smaller when the number of particles increases.

10

Various interpretations

Long ago, and almost in parallel with the “orthodox” Copenhagen/standard interpretation, other interpretations of quantum mechanics were proposed. Giving an exhaustive discussion of all points of view that have been put forward since then would probably be an impossible task. Moreover, while one can distinguish big families among the interpretations, it is also possible to combine them in many ways, with an almost infinite number of nuances. The Copenhagen/standard interpretation itself, as we have seen, is certainly not a monolithic construction, but can be declined in various forms. In this chapter, we will therefore limit ourselves to a general description of the major families of interpretations.

We will begin with a brief description of some frequent attitudes observed among scientists in laboratories, who do not necessarily pay extreme attention to the foundations of quantum mechanics, even when they do experiments in quantum physics. In practice, they often use pragmatic rules, which are sufficient to interpret all their experiments, and prefer to avoid difficult questions about the very nature of the measurement process. For instance, one popular point of view is the “correlation interpretation”, which can be considered as a “minimal interpretation” – minimal but sufficient for all practical purposes; it is accepted as a valid rule by a large majority of physicists, even those who prefer to supplement it with other elements in order to reach a more precise interpretation for the whole theory. We will then proceed to discuss various families of interpretations that are less common, including additional variables, modified Schrödinger dynamics, consistent histories, Everett interpretation, etc. All of them tend to change the status of the postulate of state vector reduction; some interpretations incorporate it into the normal Schrödinger evolution, others make it a consequence of another physical process that is considered as more fundamental, still others use a formalism where the reduction is hidden or even never takes place. But the general purpose remains always the same: to solve the problems and questions that are associated with quantum measurement,

the coexistence of two postulates for the evolution of the state vector, and the emergence of macroscopic uniqueness.

10.1 Pragmatism in laboratories

Until now the fundamental difficulties of quantum mechanics have had little impact in the practical application of quantum mechanics in laboratories. Fortunately, in a sense, physicists know how to avoid these difficulties by just applying common sense (or physical intuition), even if the logics behind the method may sometimes remain rather vague at a fundamental level. This attitude has proved fruitful, opening the way to many discoveries, which would not necessarily have been possible if the physicists had remained blocked at the level of the foundations; applying the theory as it is to concrete examples can be more productive. Moreover, for reasons that we have given in the introduction of Chapter 9, in practice there is often no need to apply the quantum postulates of measurement (Born rule and state vector reduction), since the calculation of quantum averages with the Schrödinger equation is sufficient. It is then rather natural that, in laboratories, the problems associated with the conceptual difficulties of quantum mechanics are not always given first priority.

Conceptually, of course, even very pragmatic physicists would not mind having some explanation for the uniqueness of macroscopic observations; logical consistency requires to be able to somehow incorporate it into quantum mechanics. In a first step, we will limit ourselves to the description of two popular strategies. The first (§10.1.1) is to break the Von Neumann chain by hand, when it “obviously” goes too far; the appearance of uniqueness is applied, so to say, above the formalism of the theory. Another popular strategy (§10.1.2) is to use the “correlation interpretation”, a point of view where macroscopic events are indeed considered as unique by definition (or by common sense), and where the role of the theory is just to give probabilities relating preparation events to measurement and observation events. Other scientists prefer to emphasize the central role of information in a theory such as quantum mechanics (§10.1.3).

10.1.1 Common sense: breaking the Von Neumann chain by hand

When physicists put an end to the infinite Von Neumann chain, they often do it in an implicit way, based on physical intuition. Here we attempt to make the process explicit by discussing two examples: modified macroscopic decoherence, and the effect of consciousness.

10.1.1.a Modified decoherence

One possible empirical rule to apply is to consider that, as soon as “significant” decoherence takes place, the Von Neumann chain automatically stops: all

its branches spontaneously disappear, except one. Nature chooses this branch by some unknown physical process. In other words, one systematically associates emergence of uniqueness with macroscopic decoherence. For instance, as soon as a measuring apparatus is inserted in the experiment, if it is able to register results, one considers that it actually registers a single result, whether or not a human being observes it. As we have already seen (Chapter 6), there is no hope to ever observe the physical effects of coherent superposition when they have propagated too far in the environment; assuming that they have merely vanished therefore creates no risk of contradiction with experimental observations. The difficulty is, of course, to define the precise meaning of the word “significant” in this context.

Breaking the Von Neumann chain in this way “by hand” is, after all, not very different from applying the state vector reduction postulate in a slightly different way: instead of the conscious act of measurement, one considers that decoherence triggers the reduction. In other words, one believes in the Schrödinger equation, but not too far: only until it contains correlations with the environment that become too macroscopic. Somewhere, decoherence sets the border between Schrödinger’s world and Born’s (introduction of §2). The “postulate of macroscopic decoherence” is not very different from Bohr’s point of view, since it also involves a macroscopic world that is accessible to our human experience and is unique. It contains inherent non-locality to explain Bell-type experiments: when two spins in a single state are measured at very distant places, the correlation with the macroscopic world and the modified decoherence will occur simultaneously at these places in a way that is non-local.

One could nevertheless see this view more as phenomenology than as a full interpretation of the theory; questions such as “Precisely at what degree of entanglement is the Von Neumann chain resolved into a single branch?” are left to common sense and personal judgment; in this sense the theory is not complete. Theories with modified Schrödinger dynamics (§10.8) are precisely built to rationalize this approach by introducing a precise physical mechanism to stop the Von Neumann chain. They provide precise answers to such questions, but the mechanism allowing the emergence of a single branch is more intrinsic to the system (it may involve the masses of the particles) than necessarily related to the environment.

10.1.1.b Effect of consciousness

We have already mentioned in §2.3 what is sometimes called the “Wigner interpretation” of quantum mechanics [55], where the origin of state vector reduction is related to consciousness. The idea is similar to breaking the Von Neumann chain when it reaches a measurement apparatus, but here the chain stops as soon as the linear superposition involves different states of consciousness. Similar views have been discussed by London and Bauer in 1939 [49], who emphasize that state vector reduction restores a pure state from a statistical mixture of the measured

sub-system (see §6.3.3), as well as “the essential role played by the consciousness of the observer in this transition from a mixture to a pure state”; they then explain this special role by the faculty of introspection of conscious observers. Others prefer to invoke “special properties” of the electrical currents which correspond to perception in a human brain.

In any case, common sense suggests that consciousness provides at least an upper limit, a boundary beyond which the Von Neumann chain certainly does not propagate; whether or not this propagation stops exactly at this limit, or before, is another question. Indeed, Wigner may have seen the introduction of an influence of consciousness only as an extreme case (exactly as the Schrödinger cat was introduced by Schrödinger as a “ridiculous case”), just for illustrating the necessity of a non-linear ingredient in order to predict definite results of experiments (we come back to modified Schrödinger dynamics in §10.8). In any event, the merit of the idea is also to show how the logic of the notion of measurement in the standard interpretation can be pushed to its limits: indeed, how is it possible to ascribe such special properties to the operation of measurement without considering that the human mind also has very special properties?

10.1.2 Correlation interpretation

As already noted, the correlation interpretation is, more or less, a component that is common to all other interpretations. Those who would consider that it is not really sufficient to be a full interpretation may even call it “minimalistic”. But, precisely, the fact that it is a sort of common core makes it especially interesting. After all, anyone who finds it insufficient may add more elements to it, for instance more realism. Here we will use the words “correlation interpretation” to describe this point of view, since it puts the emphasis on the correlations between successive results of experiments.

In this interpretation, the uniqueness of macroscopic events (including acts of measurements) is not questioned, but postulated; one also postulates that they occur in a stochastic way. The purpose of the theory is just to provide probabilities associated with all possible situations of preparation, evolution, and measurements performed on the quantum system.

10.1.2.a Calculating the probability for a sequence of results in successive measurements

We now evaluate the probability associated with any sequence of measurements, performed at different times. Let us assume that a measurement¹ of a physical

¹ We assume that all measurements are ideal; if non-ideal measurements are considered, a more elaborate treatment is needed.

quantity associated with operator M is performed at time t_1 , with possible results $m_i, m_j, \dots$; this is followed by another measurement of observable N at time t_2 , with possible results $n_k, n_l, \dots$, of observable P at time t_3 , etc. The system is initially described by a normalized pure state $|\Psi(t_0)\rangle$. Between time t_0 and time t_1 this state evolves from $|\Psi(t_0)\rangle$ to $|\Psi(t_1)\rangle$ according to the Schrödinger equation:

$$|\Psi(t_1)\rangle = U(t_1, t_0) |\Psi(t_0)\rangle \quad (10.1)$$

where $U(t', t)$ is the evolution operator between times t and t' ; we can expand this new state into its components corresponding to the various results that can be obtained at time t_1 :

$$|\Psi(t_1)\rangle = \sum_m |\Psi_m(t_1)\rangle \quad (10.2)$$

where $|\Psi_m(t_1)\rangle$ is obtained by applying to $|\Psi(t_1)\rangle$ the projector $P_M(m)$ on the subspace corresponding to result m (§1.2.2.a):

$$|\Psi_m(t_1)\rangle = P_M(m) |\Psi(t_1)\rangle \quad (10.3)$$

The terms inside the sum over m of (10.2) are all mutually orthogonal; they will never give rise to interference effects in the future, since they correspond to different results of measurement. Actually, each term becomes correlated to an orthogonal state of the environment (the pointer of the measurement apparatus for instance) and decoherence will cancel any interference effect (cf. §6.3.3). Instead of making a complete calculation including the state of the environment, here we take the simpler point of view where we ignore the environment and consider the various orthogonal components $|\Psi_m(t_1)\rangle$ of $|\Psi(t_1)\rangle$ as independent from each other.

From time t_1 to time t_2 , state $|\Psi_m(t_1)\rangle$ evolves under the effect of the Schrödinger equation and becomes a state $|\Psi_m(t_2)\rangle$ given by:

$$|\Psi_m(t_2)\rangle = U(t_2, t_1) |\Psi_m(t_1)\rangle \quad (10.4)$$

For the second measurement, the procedure repeats itself; we expand this new state according to:

$$|\Psi_m(t_2)\rangle = \sum_n |\Psi_{m,n}(t_2)\rangle \quad (10.5)$$

where $|\Psi_{m,n}(t_2)\rangle$ is obtained by the action of the projector $P_N(n)$ on the subspace corresponding to result n :

$$|\Psi_{m,n}(t_2)\rangle = P_N(n) |\Psi_m(t_2)\rangle \quad (10.6)$$

Again, we consider that the evolution of each $|\Psi_{m,n}(t_2)\rangle$ is independent and, if a third measurement is performed at a later time t_3 , it will generate one more decomposition, and so on. One can then check (Appendix G) that the probability of any given sequence of measurements m, n, p , etc. is nothing but the square of the norm of the final state vector:

$$\mathcal{P}(m, t_1; n, t_2; p, t_3; \dots; q, t_r) = \langle \Psi_{m,n,p,\dots,q}(t_r) | \Psi_{m,n,p,\dots,q}(t_r) \rangle \quad (10.7)$$

The probability is therefore merely the square of the norm of the “branch” of the state vector, after the action of all projections at various times corresponding to the whole series of results.

This formula can also be written in terms of the initial density operator $\rho(t_0)$:

$$\rho(t_0) = |\Psi(t_0)\rangle \langle \Psi(t_0)| \quad (10.8)$$

In the Heisenberg picture (for more details, see Appendix G), the projectors $P_M(m)$ and $P_N(n)$ become time-dependent operators $\hat{P}_M(m, t)$ and $\hat{P}_N(n, t)$. For two measurements, the probability of obtaining result m followed by result n can then be written as²:

$$\mathcal{P}(m, t_1; n, t_2) = \text{Tr} \{ \hat{P}_N(n_2, t_2) \hat{P}_M(m_1, t_1) \rho(t_0) \hat{P}_M(m_1, t_1) \hat{P}_N(n_2, t_2) \} \quad (10.9)$$

Relation (10.9) is sometimes called the Wigner formula³. It can easily be generalized to more than two measurements by inserting additional projectors on both sides in the reverse time order, and to situations where $\rho(t_0)$ describes a statistical mixture instead of a pure state.

10.1.2.b Getting rid of state vector reduction

Equations (10.7) and (10.9) can be seen as a consequence of the postulate of state vector reduction. Conversely, it is also possible to take either of these equations as a starting point, which then becomes a postulate in itself giving the probability of any sequence of measurements in a perfectly unambiguous way. The postulate of state vector reduction then becomes superfluous, since the generalized (multi-time) Born rule given by (10.7) or (10.9) is sufficient to predict the probabilities of any sequence of measurements – one could still argue that the state vector reduction is also contained in some way in the trace operation of (10.9), but no explicit reference to it is indispensable. From this perspective, the projection of the state vector resulting from measurement is no longer a postulate, but just seen

² Using circular permutation under the trace, one can in fact suppress one of the extreme projectors $\hat{P}_N(n_2; t_2)$ in formula (10.9), but not the others.

³ See equation (12) of [54, 364].

as a convenient rule to make calculations, which can be derived from another postulate. As for the Schrödinger evolution, it is contained in the Heisenberg evolution of projection operators, so that the evolution of $|\Psi\rangle$ is no longer directly apparent.

It is of course still necessary to postulate that the results of any measurements can give only one of the eigenvalues of the corresponding operator, and that the result is fundamentally random, which is essentially the content of the Born rule. The advantage is that, if one just uses formula (10.9) and ignores state vector reduction, the problems associated with a difficult coexistence of two different evolution postulates disappear; no discontinuous jump of any mathematical quantity ever occurs in the formalism. Why not then give up entirely the other postulates and just use this single formula for all predictions of results?

This is indeed the best solution for some physicists: if one accepts the idea that the purpose of physics is only to correlate the preparation of a physical system, contained mathematically in $\rho(t_0)$, with all possible sequence of results of measurements (by providing their probabilities), it is true that nothing more than (10.9) is needed. Why then worry about which series of results is realized in a particular experiment? There is no need in physics to do more than just give rules for the calculation of probabilities associated with the various preparation and measurement procedures (see for instance Dirac's quotation in §2.5). No attempt is made to describe the physical system itself during each realization of an experiment.

The “correlation interpretation” is perfectly consistent; it goes well with the view where a state vector expresses a preparation procedure (§1.2.3) rather than a physical property of the measured system. On the other hand, it is the complete opposite of the EPR reasoning, since it shows no interest whatsoever in questions related to physical reality as something “in itself”. Questions such as: “How should the physical system be described when one first measurement has already been performed, but before the second measurement is decided” are dismissed as superfluous or meaningless⁴. Needless to say, the notion of the EPR elements of reality itself becomes completely irrelevant to physics, which automatically solves all potential problems related to Bell, GHZ, and Hardy type considerations. The same is true of the emergence of a single result in a single experiment; in a sense, the Schrödinger cat paradox is eliminated by putting it outside of the scope of physics, because the paradox is not

⁴ Suppose for instance that the polarization of a photon is measured non-destructively somewhere in Europe, that it propagates along a polarization preserving optical fibre, and reaches America where another polarization measurement is performed. In practice, most physicists consider intuitively that “something having physical properties” has propagated from one site of measurement to the other, and that physics has something to say about these properties. In the pure correlation interpretation, this intuition should be rejected: nothing real that physics could describe propagates along the fiber.

expressed in terms of correlations. An interesting feature of this point of view is that the boundary between the measured system and the environment of the measuring devices is flexible; an advantage of this flexibility is that the method is well suited for successive approximations in the treatment of a measurement process, for instance the tracks left by a particle in a bubble chamber as discussed by Bell [47].

10.1.2.c Discussion

In practice, most physicists who favor the correlation interpretation do not feel the need for making it very explicit. Some nevertheless do; see for instance the article by Mermin [365], which starts from the statement: “Throughout this essay, I shall treat correlations and probabilities as primitive concepts”. In a similar context, see also a recent “opinion” in *Physics Today* by Fuchs and Peres [86] who emphasize “the internal consistency of the theory without interpretation”. On the other hand, the correlation interpretation is seen by some physicists as minimalistic because it leaves aside, as irrelevant, questions they find important; the best example is probably the notion of physical reality seen as an entity that should be independent of measurements performed by human beings. As we have already mentioned, this interpretation can easily be supplemented by other elements to make it more specific. Experience shows that defenders of the correlation point of view, when pressed in a discussion to describe their point of view with more accuracy, often express themselves in terms that come very close to the Everett interpretation (§10.11); in fact, they may sometimes be proponents of this interpretation without realizing it!

Let us finally mention in passing that formula (10.9) may be the starting point for many interesting discussions, taking it as a convenient formula, not necessarily basic in the interpretation. Suppose for instance that the first measurement is associated with a degenerate eigenvalue of an operator, in other words that $\hat{P}_M(m; t_1)$ is a projector over a subspace of more than one dimension:

$$\hat{P}_M(m; t_1) = \sum_{i=1}^n |\varphi_i\rangle \langle \varphi_i| \quad (10.10)$$

(for the sake of simplicity we assume that $t_1 = t_0$, so that no time dependence appears in this expression). Inserting this expression into (10.9) immediately shows the appearance of interference terms (or crossed terms) $i \neq j$ between the contribution of the various $|\varphi_i\rangle$. Assume, on the other hand, that more information was actually obtained in the first measurement, so that the value of i was also determined, but that this information was lost, or forgotten; the experimenter ignores which of two or more i results was obtained. Then, what should be calculated is the sum of

the probabilities associated with each possible result, that is a single sum over i from which all crossed terms $i \neq j$ have disappeared. In the first case, interference terms arise because one has to add probability amplitudes; in the second, they do not because one has to add the probabilities themselves (exclusive events). The contrast between these two situations may be understood in terms of decoherence: in the first case, several states of the system correlate to the same state of the measuring apparatus, which plays here the role of the environment; they do not in the second case, so that by partial trace all interference effects vanish. This remark is useful in the discussion of the close relation between the so-called “Zeno paradox in quantum mechanics” [366] and decoherence; it is also basic in the definition of consistency conditions for decoherent histories, to which we will come back later (§10.10).

10.1.3 *Emphasizing information*

Another popular point of view is to emphasize the role of information [287, 367–373]; with the rise of the field of quantum information, the informational point of view is natural. Information can be about the whole experimental setup, including preparation and measurement devices; in this case, the idea becomes very similar to Bohr’s emphasis on the relevance of the whole experimental setup (§3.3.2). Information can also include components that are acquired when experimental results become known, and undergo sudden changes when the results are observed; this property can be invoked to explain the Von Neumann state vector reduction, seen as a “purely mental process” (Appendix A).

An interesting point is that any measurement process gives rise to a flow of information from the location where the measurement takes place to the environment. This is because the interaction between the measured system and the measurement apparatus, and then the environment, initiates a Von Neumann chain, during which entanglement progresses further and further in the environment [372]. Consider a fixed volume containing all the measurement apparatus; as long as the entangled chain does not leave this volume, since any Hamiltonian evolution preserves pure states as well as entropy, the amount of entropy contained in the volume remains constant; but, as soon as entanglement proceeds beyond the fixed volume, the properties of the physical system contained inside the volume must be obtained by a partial trace (over the system outside of the volume), and the entropy contained in the volume increases (§6.2.2). To the local observer, the effect of this leak of entanglement towards the outside world appears as an entropy production – as an illustration, see for instance Peres’s quotation in §10.3.2 and his discussion of the flux of information in an EPR experiment, and [374].

This informational point of view can be applied in a more or less strict way. One may focus the interest on the information content of the state vector⁵, or consider that the nature of the state vector itself is mostly informational, or even take the more extreme view where it is purely informational (§10.3.2). In any case, this point of view has obvious relations with the correlation interpretation, and can actually be used to complement it. It seems to explain the discontinuities arising from state vector reduction rather naturally, but of course the difficulties already discussed in §§1.2.3.a and 10.1.2 remain (for instance, questions about the division of the world into systems providing information and those on which information is acquired, or the description of independent reality during the experiment). Wigner's friend paradox is not problematic since, as long as the friend outside the closed laboratory has less information than the other, he continues to use a non-reduced state vector, while the friend inside has already reduced his.

10.2 Statistical interpretation

The statistical interpretation⁶ does not discard the idea of describing physical systems, but specifies that the description given by the state vector applies only to ensemble systems prepared in identical conditions, not single systems (or single experiments). Einstein supported this point of view; for instance, in a letter to Schrödinger [53], he writes: "The Ψ function does not describe the state of a single system but (statistically) an ensemble of systems". The EPR argument leads to the statistical interpretation. A review of this interpretation has been given by Ballentine [375], who writes: "Several arguments are advanced in favour of considering the quantum state description to apply only to an ensemble of similarly prepared systems, rather than supposing, as is often done, that it exhaustively represents an individual physical system. Most of the problems associated with the quantum theory of measurements are artifacts of the attempt to maintain the latter interpretation". This author distinguishes two classes of interpretations:

- (i) "The statistical interpretation ... according to which a pure state ... need not provide a complete description of an individual system.

⁵ Here is for instance how Fuchs [370] describes the program of informational quantum theory: "The quantum system represents something real and independent of us; the quantum state represents a collection of degrees of belief about *something* to do with that system ... The structure called quantum mechanics is about the interplay between these two things – the subjective and the objective. The task before us is to separate the wheat from the chaff. If true quantum state represents subjective information, then how much of its mathematical support structure might be of the same character? Some of it, maybe most of it, but surely not all of it."

⁶ This interpretation is also sometimes called "ensemble interpretation".

- (ii) Interpretations which assert that a pure state provides a complete and exhaustive description of an individual system. This class contains ... several versions of the Copenhagen interpretation”.

He then adds that “hypothesis (ii) is unnecessary for quantum theory, and moreover leads to serious difficulties”. Other authors have expressed similar points of view; see for instance [98], where an alternative to the orthodox interpretation of quantum theory is proposed in the same line. For an example of explicit disagreement with the statistical interpretation, see for instance [45].

Once the statistical interpretation is accepted, two different attitudes become possible logically:

- (i) either one considers as satisfactory a theory that describes only ensembles of systems. If a single experiment is performed, one decides that a fundamentally random process takes place and makes a single result appear; no attempt is made to describe this process. This is, in a sense, a modern version of the historical “quantum jumps”. Pushing this attitude to the extreme leads to what Leggett [376] calls the “extreme statistical interpretation, according to which no physical meaning should be attached to the QM formalism either at the microscopic or at the macroscopic level”.
- (ii) or one considers that the fact that a description is valid for ensembles only, but not for a single system, signals that the description cannot be complete; more variables are needed to specify which particular system is considered inside the ensemble. This leads to introducing new elements of description that come in addition to the state vector⁷, which leads to theories with additional (hidden) variables.

Not all physicists favoring this interpretation make an explicit choice among these two possibilities. Nevertheless, with or without explicit reference to additional variables, one can find a number of authors who support the idea that the quantum state vector should be used only for the description of statistical ensembles. An interesting connection between the two points of view has been proposed by Aharonov *et al.* in [377], under the title “What is the meaning of the wave function?” After remarking that “Since the discovery of quantum theory, a very fundamental question has haunted physicists: what is the physical meaning of the wave function?” the authors explain how it is possible to extend the usual interpretation of the wave function (with a physical meaning by ensemble averages of a large number of identical systems) to a single system. This is done by considering

⁷ In his article, Ballentine remarks that “the introduction of hidden variables is fully compatible with the statistical predictions of quantum theory” and discusses the properties of these variables at the end of his article.

a measurement that lasts a long time, a “protective measurement”, during which the wave function is prevented from changing appreciably by means of another interaction which it undergoes at the same time.

10.3 Relational interpretation, relative state vector

Other possible interpretations of quantum mechanics emphasize the relative character of the state vector: several observers using the same Galilean frame may use different state vectors to describe the same sequence of events.

10.3.1 Relational interpretation

The relational point of view on quantum mechanics, introduced by Rovelli in 1996 [378, 379], is inspired by relativity, where different observers may use different times (if they use different Galilean frames). In the relational interpretation, the notion of an absolute, observer-independent, state of a physical system is rejected: different observers may give different accounts of the same series of events by using different state vectors. But the difference does not arise from the use of different Galilean frames; it arises from different informations available to the observers, or on a more concrete basis from the use of different measurement apparatuses to make observations. Physical properties of systems are then not seen as absolute, but as depending on the apparatus used to measure the property. Quantum mechanics is considered as “a theory about information”. In the words of Rovelli (§2.3 of [378]): “A quantum description of the state of a physical system *S* exists only if some system *O*, considered as an observer⁸, is actually ‘describing’ *S* or, more precisely, has interacted with *S*. The quantum state of a system is always a state of that system with respect to other systems”.

The emphasis is clearly put on the relations between physical objects, not the physical objects themselves. In an EPR experiment for instance, if the spin of one particle only is measured, this has no effect whatsoever on the state of the other spin, since it is not in relation with any measurement apparatus; nothing like a non-local effect of state vector reduction happens. In a second experiment, if the spins of the two particles are now measured along the same direction, one can then observe a property of the pair, namely that the results are always opposite; but this cannot be related to any property of the system in the first experiment, since the physical properties of the system with respect to a different apparatus may have

⁸ System *O* can be understood as a measurement apparatus. This sentence then implies that (in a given Galilean frame) a particle can have a definite position, or velocity, or spin direction, etc. only with respect to the specific apparatus used to measure this property, never in an absolute way.

completely changed. If finally the two spins are measured along different directions, one observes still other physical properties of the pair, with no special relation to those observed in the two first experiments.

In this interpretation, quantum state vector reduction (or collapse) becomes very different from what it is in the standard view. For instance, for a given quantum system, some observers may consider that state collapse has already occurred while, for others, collapse has not yet taken place – for them the system remains in a coherent superposition. Consider again an EPR experiment [380]: when Alice performs her experiment, a reduction occurs for her state vector describing her spin in relation with her measurement apparatus, and of course herself having observed a single result. Nevertheless Bob, who is far away, knowing that she has performed an experiment but not knowing the result, describes the same ensemble as a coherent superposition involving all possible results. In other words, for Bob, Alice has become a Schrödinger cat. Similarly, if Bob also performs a measurement along a different spin direction, this reduces his own state vector, but Alice does not know the result and considers that he has put himself into a coherent superposition with respect to her and her apparatus. Each experimenter considers that the other is a Schrödinger cat, and this remains true until they communicate to exchange their results; the cats then disappear and both operators use the same state vector again. This example illustrates that, in this interpretation, the state vector does not directly describe reality, but rather the information that is available on this reality (not necessarily available to someone, but possibly relative to a measurement apparatus for instance).

In Bohr's point of view, the role of the experimental apparatus is also emphasized, but it is essential to consider it as a whole. Here, various part of the apparatus may be distinguished, and may actually introduce different state vectors. Moreover, in the Copenhagen interpretation, measurement apparatuses and observers play a central role, as well as differences between the microscopic world and the world that is directly accessible to our human experience. In the relational interpretation, this becomes unnecessary. All physical systems are considered equal, with no mention of the necessarily macroscopic character of measurement apparatuses: the measurement process is just an ordinary interaction process, with no special role in the theory. The process can even take place in the absence of any observer, since the state vector and its reduction can be defined in relation with the measurement apparatus only.

As Rovelli summarizes in [378], “Quantum mechanics is a theory about the physical description of physical systems relative to other systems, and this is a complete description of the world”. Since the relational interpretation considers that state vectors provide a complete description of physical reality, it is of course in complete opposition with the EPR argument, additional variables, etc.

10.3.2 Pure informational point of view

As we saw in §10.1.3, some physicists emphasize the informational content of the state vector. One can go further in this direction and take an extreme informational point of view, where $|\Psi\rangle$ is completely disconnected from physical reality; what is then considered is only the information content of the state vector $|\Psi\rangle$, which describes the information that one observer has concerning a given physical system (§10.1.3), but says nothing about the system itself. The state vector $|\Psi\rangle$ then relates to mental information only (Appendix A) and may take almost arbitrary values for the same system, depending on the amount of knowledge one ascribes to an arbitrary observer. This is analogous to a classical probability distribution, which also expresses a relation between some amount of knowledge and some independent reality; for instance, such a distribution may sometimes describe the system with perfect accuracy, but sometimes also contains no information at all (if the observer knows nothing about the system).

Consider an experiment which, in ordinary language, one would describe as: “A device at point A produces a particle, which propagates in space to another remote point B, where its interactions with a target are observed and recorded” – it can be a proton that is injected into the big CERN accelerator and collides with a target kilometers away. During the experiment, a wave packet associated with the state vector propagates from A to B. Nevertheless, from a strict informational point of view, one should consider that “nothing real” travels along within this wave packet: it is just a transfer of abstract information. The real world includes only the experimental devices and the observations; what propagates in between is just information about what the future experimental observations with the measurement apparatuses can be.

In [94], the words of Peres nicely illustrates this point of view, applied to the discussion of EPR experiment: “When Alice measures her spin, the information she gets is localized at her position, and will remain so until she decides to broadcast it. Absolutely nothing happens at Bob’s location . . . It is only if and when Alice informs Bob of the result she got . . . that Bob realizes that his particle has a definite pure state. Until then, the two observers can legitimately ascribe different quantum states to the same system . . . Quantum states are not physical objects: they exist only in our imagination . . . The question raised by EPR ‘can quantum mechanical description of physical reality be considered complete’ has a positive answer. However, reality may be different for different observers”.

Most physicists are prepared to accept the idea of observer-dependent density operators: if different observers have different amounts of information about the same physical system, they may describe it with different statistical mixtures, and therefore by different density operators ρ . The analogy with a classical distribution

function is then obvious, since ρ does not correspond to the most precise possible description of a system within quantum mechanics; it actually attributes probabilities to several such descriptions (with state vectors, or pure states). But an observer dependence of state vectors $|\Psi\rangle$, or equivalently with density operators ρ associated with pure states, seems much more delicate to accept: pure states do not leave room for a more precise quantum description. Moreover, most physicists would probably agree that, at least sometimes, the wave function does contain elements of reality; for instance, a system described by a ground state wave function is really in the ground state and has the properties of this state, not only in the mind of humans. The electrons of a superconductor at very low temperature are really in a BCS (Bardeen–Cooper–Schrieffer) ground state; the BCS wave function is not a mental process, but a description of the system itself. Similarly, chemists think that molecular orbitals describe real properties of molecules such as shape, size, etc. and not just knowledge of these properties. The standard view is rather that the Schrödinger evolution contains at the same time a physical evolution for those properties that quantum mechanics does attribute to the system (observables that have $|\Psi\rangle$ as an eigenvector), but also an evolution of probabilities (for all other observables) that represent only our knowledge of the system, and therefore can indeed be seen as mental.

We have already discussed the difficulties of the pure informational point of view in §§1.2.3.a and 10.1.3. If $|\Psi\rangle$ is only information, what is this information about? How can we distinguish between systems that produce information and the systems on which information is acquired? If we accept the idea of an independent physical reality, how should we then describe the system itself? Does this mean that this very idea should be abandoned, and that physics completely gives up any idea of saying something about independent reality? In some early writings, Heisenberg seemed not to be very far from the pure information point of view, but later he took a more moderate point of view. As for Bohr, he was not a positivist but a realist, one could suggest in his own way (see §2.5, for example Von Weizsäcker’s quotation at the end); the pure information point of view is usually considered as beyond the standard interpretation (§1.1.3).

10.4 Logical, algebraic, and deductive approaches

In this section, we discuss a family of interpretations where the state vector $|\Psi\rangle$ is not considered as the starting point for describing a quantum physical system. It is rather seen as a mathematical tool that can be derived from more basic considerations concerning the possible statements that can be made about the properties of a physical system. We will briefly review three classes of theories: quantum logic, which is related to logic as a branch of philosophy applied to quantum phenomena,

but also involves mathematics; algebraic and C^* algebra theories, which are more purely mathematical; axiomatic theories, which also borrow from both logic and mathematical formalism. In the last part of this section, we discuss a general theorem, the Gleason theorem, which has interesting applications in these theories and, more generally, in all quantum mechanics.

10.4.1 Quantum logic

Quantum logic studies the formal structure of deductive reasoning and of propositional statements within quantum mechanics, in order to provide a set of rules that are compatible with its principles. Within science in general, this defines a broad problem, with a rich history, actually too broad and philosophically oriented to be discussed here, despite its intrinsic interest. The roots of the subject go back to Greek philosophy (Aristotle), but the application to quantum mechanics is of course more recent. Nevertheless, even within this more restricted frame, one can still find a large variety of contributions; we will therefore limit ourselves to a brief summary, giving just a flavour of the variety of ideas in quantum logic. A general historical presentation of the field can be found in Chapter 8 of Jammer's book *The Philosophy of Quantum Mechanics* [50], with many references. More specialized books are for instance *Quantum Logic* by Mittelstaedt [381] or *The Logic of Quantum Mechanics* by Beltrametti and Casinelli [382]; many useful references can also be found in [383].

Von Neumann, in his famous 1932 treatise [4], remarked that “the relations between a physical system on one hand, and the projections on the other, make possible a sort of logical calculus with these. However, in contrast to the concepts of ordinary logic, this system is extended by the concept of ‘simultaneous decidability’, which is characteristic of quantum mechanics”. In 1933, an astrophysicist, Fritz Zwicky, suggested that non-classical logic should be used in microphysics [384]; he proposed a “principle of flexibility of scientific truth”, which “must be many-valued”. In 1936, Birkhoff and Von Neumann attempted to reconcile the apparent inconsistency of classical boolean logic with the rules applying to the measurement of incompatible observables in quantum mechanics [385]; they pointed out that the characteristic difference between the logical structures of classical and quantum mechanics is that distributive identities hold in the former, but not in the latter. Jordan published a few articles on the subject around 1950 [386]. Soon after the publication of the Birkhoff and Von Neumann article, Strauss criticized it in his Ph.D. dissertation (1938); he then introduced another version of logic, “complementarity logic” [387]. Later, in 1968, Popper also criticized the Birkhoff and Von Neumann article, but this time by considering that it is logically inconsistent [388]; a debate ensued, in particular with Ramsay and Pool, which

will not be reported here. For a simple introduction of quantum logic, see for instance [389].

Many valued logic corresponds to a version of quantum logic that has been proposed by Reichenbach in the early 1940s [390]. In his “probability theory of meaning”, he considers that “a proposition has a meaning if it is possible to determine its weight, i.e. a degree of probability, for the proposition” as well as similar rules for groups of propositions. Later, Von Weizsäcker [391] proposed a different version of complementarity logic, where “every elementary proposition can have, apart from 1 and 0, a complex number as its truth value”. The square modulus of this value gives, as in standard quantum mechanics, the probability that an experimental test will verify the logical proposition; the phase of the complex number is related to the complementary logical alternative. This logical construction is analogous to an infinite valued logic.

10.4.2 Algebraic theory

The axiomatic approach and the “ C^* algebra theory” is a mathematical view of quantum mechanics, sometimes involving elaborate theorems, algebras of operators, and their representations. It arose initially from the need to develop a quantum theory treating rigorously systems with an infinite number of degrees of freedom, especially in quantum field theory but also in statistical mechanics. Von Neumann also initiated the field by considering algebras of operators [4], and continued his work with Jordan and Wigner [392]. In addition to the “Von Neumann algebras”, Gelfand and Naimark [393] introduced another class of algebras, now named “ C^* algebras”, which were also used in the contributions of Segal [394] and later of Haag and Kastler [395].

Murray and Von Neumann had given a classification of the various types of algebras [396], with type I, type II, etc. Type I algebras are appropriate to quantum mechanics with a space of states of finite dimension only. Von Neumann used type II algebras, but it turns out that type III factors are actually necessary for quantum field theory (for a brief review, see for instance §§9.2 and 8.3 of [383]). A important clarification in this field was introduced by Connes, who gave a classification of algebras of type III [397] and further introduced the notion of non-commutative geometry [398].

10.4.3 Formal axiomatic theory

In the line initiated by Von Neumann, Hilbert, and Jordan, Mackey [399] provided a formalism for relating observables, system preparation filters, and states, and introduced an axiomatization of quantum mechanics based on two fundamental notions, observables and states. This was the first complete system of axioms for a

formal theory. The work was then expanded by Piron [400] who, with Jauch [401], showed that one can reconstruct a Hilbert lattice (and not only an orthomodular lattice) by introducing an appropriate “covering property”. For a more detailed review of the role of the Geneva school in this field, see §8.5 of [50]. A more recent and important contribution is the Solèr theorem [402] which, on the basis of algebraic assumptions, proves that the Hilbert space must be built either from real numbers or complex numbers (as in standard theory), or quaternions.

The above theories are not free of difficulties, sometimes illustrated by counterexamples. One of them is that, while the structure of a Hilbert space provides a natural way to treat the composition of several physical systems, with the notion of the tensor product (§11.2), one can not merely multiply Hilbert lattices. In fact, combining two or more physical systems may introduce new mathematical structures. To treat this problem, it has been proposed to introduce convex sets and the language of categories [403–405]; for a review, see [406].

10.4.4 Gleason theorem

A theorem due to Gleason [407] is useful in the context of this family of interpretations. Consider a space of states $\mathcal{E}$ and the ensemble of the orthogonal projectors P_j onto all its subspaces⁹. Any orthonormal basis $\{|u_i\rangle\}$ of $\mathcal{E}$ immediately provides examples of such projectors, for instance projectors onto single states:

$$P_j = |u_j\rangle\langle u_j| \quad (10.11)$$

Any sum of these projectors (without including the same operator twice or more) also gives a possible P_j :

$$P_j = |u_i\rangle\langle u_i| + |u_k\rangle\langle u_k| + \dots |u_k\rangle\langle u_k| \quad (10.12)$$

since the $|u_i\rangle$ are mutually orthogonal. Moreover, any basis $\{|u_i\rangle\}$ can be chosen, so that it is clear that the ensemble of the P_i is large; it is neither limited to projectors onto single states, nor to projectors that are mutually orthogonal.

Each P_j is an Hermitian operator with two eigenvalues, 0 and 1. Physically, one may associate a measurement to P_j , with results 1 or 0. If the state $|\Psi\rangle$ of the system is invariant under the action of P_j , the result 1 is obtained with certainty; if $|\Psi\rangle$ is cancelled under the action of P_j , the result 0 is obtained with certainty. In all other

⁹ To define P_j , one chooses a subspace $\mathcal{E}_S^j$. Any ket $|\Psi\rangle$ of $\mathcal{E}$ can be written as the sum of a vector $|\Psi\rangle_j$ belonging to $\mathcal{E}_S^j$ and a vector $|\Psi\rangle_j^\perp$ orthogonal to $\mathcal{E}_S^j$. The definition of P_j is then $P_j |\Psi\rangle = |\Psi\rangle_j$.

Orthogonal projectors are Hermitian. If the product $P_j P_k$ of two operators vanishes, since the Hermitian conjugate $P_k P_j$ also vanishes, P_j and P_k are commuting operators. Orthogonal projectors are a special case of the ensemble of projectors acting inside a space (oblique projectors), satisfying the relation $P^2 = P$.

situations, namely when the state is changed but not cancelled under the action of P_j , the result is randomly either 1 or 0; the probability $\mathcal{P}_j(1)$ of obtaining 1 is given by the average of P_j in $|\Psi\rangle$:

$$\mathcal{P}_j(1) = \langle \Psi | P_j | \Psi \rangle \quad (10.13)$$

A particular operator P_j is the identity operator $I(\mathcal{E})$ in $\mathcal{E}$, corresponding to a measurement having a result 1 with certainty, whatever the state of the system is.

The sum of operators P_j is not necessarily another operator P_n . It is a projector if all the P_j project over subspaces that are orthogonal, or equivalently if all the binary products $P_i P_j$ vanish¹⁰. We then call them “mutually exclusive projectors” or just “exclusive projectors” for simplicity. Mathematically, they are exclusive because it is impossible to find any ket $|\Psi\rangle$ that does not vanish under the successive action of the two orthogonal projectors. Physically, the projectors correspond to mutually compatible observables that can both be measured at the same time; when they are exclusive, no state exists for which two results 1 are certain.

Adding a sufficient number of orthogonal P can then lead to a point where their sum projects over the whole space $\mathcal{E}$ and gives the identity operator:

$$P_{j_1} + P_{j_2} + \dots P_{j_p} = I(\mathcal{E}) \quad (10.14)$$

One then says that this series of projectors provides a “resolution of unity”. An obvious example is provided by the series of projectors over all kets $|u_i\rangle$ of any orthonormal basis of $\mathcal{E}$, but one can also group these projectors in any way to obtain other decompositions of identity.

Suppose now that a real function f is defined over all possible P_j : a real number is assigned to each of these projectors. Moreover, we assume that f has the following properties:

$$\begin{aligned} 0 &\leq f(P_j) \leq 1 && \text{for any } P_j \\ f[I(\mathcal{E})] &= 1 \\ f(P_{j_1} + P_{j_2} + \dots P_{j_m}) &= f(P_{j_1}) + f(P_{j_2}) + \dots f(P_{j_m}) && \text{for exclusive } P_j \end{aligned} \quad (10.15)$$

¹⁰ If two operators P_i and P_j project over orthogonal subspaces, one can choose orthonormal basis in each of these subspaces, and complete an orthonormal basis in $\mathcal{E}$ by adding a sufficient number of normalized vectors that are mutually orthogonal and orthogonal to the two subspaces. All these vectors are common eigenvectors of P_i and P_j with eigenvalues 1 or 0. Since by construction none of them has eigenvalue 1 twice, the product $P_i P_j$ vanishes.

Conversely, if this product vanishes, the projectors necessarily commute (footnote 9), and again one can build a common basis of eigenvectors in $\mathcal{E}$. The product $P_i P_j$ vanishes only if none of these eigenvectors has a double eigenvalue 1, which means that P_i and P_j project onto orthogonal subspaces.

In both cases, it is then easy to see that the sum $P_i + P_j$ is the orthogonal projector over the subspace given by the direct sum of the two initial projection subspaces. The generalization to more than two orthogonal projectors is straightforward by recurrence.

(the second and third conditions imply immediately that the probability associated to any decomposition of identity is 1). These mathematical conditions have a simple physical interpretation: they mean that $f(P_j)$ can be considered as the probability of obtaining result 1 in the quantum measurement associated with any P_j . The third relation expresses that the probability of the union of exclusive events is the sum of their probabilities.

The Gleason theorem then states that, if the dimension of $\mathcal{E}$ is more than 2, there exists a non-negative Hermitian operator with unit trace ρ acting in $\mathcal{E}$ such that:

$$f(P_j) = \text{Tr} \{ \rho P_j \} \quad \text{for any } P_j \quad (10.16)$$

In other words, all the values of f can be derived from a single operator ρ by a simple product and trace operation¹¹. The mathematical proof of this theorem is not trivial, and we refer the reader to the original publication [407].

If f is interpreted as giving the probabilities associated with the measurements defined by the projectors, this theorem shows that these probabilities can all be obtained from one single operator ρ , which can then be considered as the density operator of the system. This is an interesting result, since it derives mathematically the quantum formalism of the density operator and of the trace from general considerations concerning the necessary conditions satisfied by probabilities.

The Gleason theorem requires a lower limit of 3 for the dimension of the space of states; it does not apply to a spin 1/2. Nevertheless, Bush has shown [408] that, if one extends conditions (10.15) to POVM (positive operator valued measures, see §8.2.2), this limitation of the theorem can be removed. POVM define a broader class of operators, and lead to more possible decomposition of identity, than orthogonal projectors. As a consequence, the Bush form of the theorem requires more assumptions than the Gleason form. This is the price to pay to obtain a more general result (also valid in two dimensions), as well as a simpler mathematical proof.

10.5 Veiled reality

The interpretation of the “veiled reality” was introduced by d’Espagnat [24, 25]. As far as the mathematical formalism is concerned, it does not differ from the standard interpretation, but it proposes a conceptual and philosophical frame that is different. This frame is realist, as for Bohr, but here the definition of reality does not involve human perception relayed by measurement apparatuses; the existence of reality is considered as fundamental, with no need to refer to humans and their perceptive structure. Analyzing the theory of measurement, the EPR and Bell arguments, a

¹¹ There is a one-to-one correspondence between the projectors and the subspaces over which they project, so that it is equivalent to consider that the function f is defined over the projectors or over the subspaces of $\mathcal{E}$. In mathematics, one often defines f as a “measure over all possible subspaces of $\mathcal{E}$ ”.

study of the relation between counterfactuality and realism, and the consequences of the inter-subjective agreement, d’Espagnat concludes that quantum mechanics cannot lead to descriptive interpretations of individual objects. He distinguishes between independent reality, of which at best the general structures are accessible, and empirical reality (perceived phenomena). This leads him to conclude that the ultimate reality is a “veiled reality”, only marginally accessible to discursive knowledge.

Other interpretations of quantum mechanics also distinguish between two levels of reality. This is for example the case of the de Broglie–Bohm theory (§10.6.1), with a physical field that can be manipulated (the wave function) and the particle positions (which are observable, but not manipulable). But it remains a very different approach, in particular because the two levels of reality appear in the mathematical formalism of the de Broglie–Bohm theory itself.

10.6 Additional (“hidden”) variables

With additional/hidden variables, elements that clearly do not belong to the orthodox interpretation are explicitly introduced. Additional variables are added to the quantum state vector in order to obtain a more precise description of a single system. We have already seen that the EPR theorem itself can be seen as a strong argument in favor of the existence of additional variables. These variables are sometimes still called “hidden”, even if this word is somewhat paradoxical, since they are often more visible than the complex state vector¹²; Bell proposed to use the word “be-able” instead [410], which tends to be more and more used in recent articles. Here we will use the word “additional”, because of its generality.

Theories with additional variables are often built mathematically to reproduce the predictions of orthodox quantum mechanics exactly. If they give the same probabilities for all possible measurements, it is clear that there is no hope to disprove experimentally orthodox quantum mechanics in favor of these theories, or the opposite. In this sense, they are not completely new theories, but rather variations on a known theory (an exception is nevertheless mentioned in footnote 17). They nevertheless have a real conceptual interest: they can restore realism and solve the difficulties related to the presence of two kinds of evolution of the state vector (Schrödinger cat paradox). They can also restore determinism, although this is not necessarily the case: one can also build theories with additional variables that remain fundamentally non-deterministic.

¹² They appear directly in the results of measurements so that, instead of being hidden, the additional variables are actually visible. In [409], Bell writes “Absurdly, such theories are known as ‘hidden variables’ theories. Absurdly, for there it is not in the wave function that one finds an image of the visible world, and the results of the experiments, but in the complementary ‘hidden’ (!) variables”.

The general idea is that the wave function does not directly represent a quantum particle but, rather, is a wave that guides the motion of the particle. These theories started in 1926–1927 with the early work of de Broglie [411, 412], in the line of his thesis [30]. L. de Broglie first elaborated his “theory of the double solution”, which he named so since the same wave equation has two solutions: the usual continuous wave function $\Psi(\mathbf{r})$, and a solution with mobile singularities $u(\mathbf{r})$ representing the physical particle itself. The particle is then considered as an energy concentration in the singularity region of the new field $u(\mathbf{r})$; it therefore remains perfectly localized as in classical physics. It is also considered as a small clock having internal vibrations, and a motion determined by a synchronization condition between these vibrations and the external wave: “the particle glides on its wave in such a way that its internal vibrations always remain in phase with the vibration of the wave at the point where it sits”¹³. Mathematically, this leads to a “guidance equation” for $u(\mathbf{r})$ in terms of the wave $\Psi(\mathbf{r})$. Since the latter is subject to diffraction effects on external obstacles, one can then recover the predictions of standard quantum mechanics, provided one assumes that the statistical distribution of singularities is given by $|\Psi(\mathbf{r})|^2$. Nevertheless, when invited by Lorentz to report on his work at the 1927 Solvay meeting, it seems that de Broglie was too worried about the mathematical difficulties of this theory to report on it in front of this audience [77, 413]. He preferred to discuss a simpler version¹⁴, the “pilot wave theory”, where the localized singularity of $u(\mathbf{r})$ is directly replaced by a position of the particle, as in classical theory, while its motion remains totally different from classical physics since it is determined by the pilot wave.

Actually, also in 1926, Born had envisaged the possibility of introducing “additional parameters” to the theory, in the course of the famous article where he introduced the probabilistic interpretation of the wave function [34]; he even worked on this subject in more detail with his assistant Frenkel [50] but, unfortunately, this work seems to have been lost. Chapter 8 of [414] discusses early attempts at causal theories in more detail.

¹³ In de Broglie’s view, a single particle is really represented by a wave $u(\mathbf{r})$, while the wave function $\Psi(\mathbf{r})$ gives only statistical information on an ensemble of particles, so that it does not have the same status of reality. He describes this as a “curious mixture of single events and statistics” [77].

¹⁴ At the Solvay meeting, Pauli made objections to the pilot wave theory, in particular that it could not reproduce the results of standard theory for inelastic collisions. We now know that it is perfectly possible to build a pilot wave theory that provides exactly the same predictions as standard quantum mechanics, so that this objection is not valid.

Later, de Broglie regretted [77] to have “weakened his point of view” by presenting a “truncated version” of his double solution theory, from which the singularity of the wave had been suppressed and useful features disappeared. He was particularly interested in the similarities with the theory of general relativity, where the motion of singularities of the gravitational field follows geodesics of space-time. He also considered the idea of introducing a non-linear equation of motion for the wave $u(\mathbf{r})$, which spreads the singularity over a tiny mobile region of space.

In 1952, Bohm [11, 415] independently elaborated a more complete version of the pilot wave theory. As in the de Broglie pilot wave theory, he added positions to the wave function and assumed that the motion of the particle in space is guided by the gradients of the wave function, but also introduced new elements as we will see below. Both points of view thus share common basic concepts, and one often speaks of the “de Broglie–Bohm” theory. Bohmian theory is probably the best known among those that add variables to standard quantum mechanics. Another well-known example is nevertheless the work of Wiener and Siegel [12], who elaborated a mathematical formulation of quantum mechanics in terms of probabilities (or probability densities), instead of probability amplitudes, while the results remain exactly equivalent to those of the standard calculation of probabilities. Such a point of view completely eliminates the need for a special postulate concerning measurement processes.

10.6.1 Bohmian theory

In standard quantum mechanics, a single quantum object may sometimes behave as a particle, sometimes as a wave, sometimes combine both, depending on the experiment considered. In the de Broglie–Bohm theory, the particle and the wave always coexist; they actually go together and are constantly coupled. The elementary system called “quantum particle” in standard theory is replaced by an inseparable couple made of a particle and a field. We begin with a brief discussion of the general framework of this theory, and then discuss trajectories for one or two particles as well as quantum measurements. We will then outline a few objections that have been made to the theory, and conclude with a discussion putting the achievements of the theory into perspective.

10.6.1.a General framework

We first introduce the essential components of the de Broglie–Bohm theory, the guiding formula and the quantum equilibrium condition.

Guiding formula None of the usual ingredients of standard quantum mechanics disappears in theories with additional variables such as the Bohmian theory. In a sense, these ingredients are even reinforced, since the status of the wave function is promoted to become a field that is physically real¹⁵, instead of an object with

¹⁵ This field has actually two components, the real and the imaginary part of the wave function (for simplicity, we assume that the particle is spinless). Similar real fields are, for instance, the components of the electromagnetic field (we consider components since these fields are vectors while, in Bohmian theory, we are dealing with scalar fields). The analogy is nevertheless imperfect since the wave function is defined only within an irrelevant phase factor, which implies that the two components of the associated real field are also only defined modulo a constant phase. The real field then depends only on the spatial changes of the phase of the wave function, not of the value of the phase at a given point of space.

intermediate status of reality (see §1.2.3). The Schrödinger equation itself remains strictly unchanged but a completely new ingredient is added to it: in addition to its wave function/field, each particle now gets an additional variable λ , more precisely three additional variables that are the three components of a vector $\mathbf{Q}$. The evolution of $\mathbf{Q}$ is coupled to the wave function field through a “guiding formula” (or “quantum velocity term”) that, for a single particle with wave function $\Psi(\mathbf{r}, t)$, is given by¹⁶:

$$\frac{d}{dt}\mathbf{Q} = \frac{1}{m|\Psi(\mathbf{Q}, t)|^2} \operatorname{Re} \left[\frac{\hbar}{i} \Psi^*(\mathbf{Q}, t) \nabla \Psi(\mathbf{Q}, t) \right] = \frac{\hbar}{m} \nabla \xi(\mathbf{Q}, t) \quad (10.17)$$

where m is the mass of the particle and $\xi(\mathbf{r}, t)$ the phase of the complex number $\Psi(\mathbf{r}, t)$; the right-hand side of this equation is nothing but the velocity of the hydrodynamic version of the Schrödinger equation introduced by Madelung in 1927 [416]. We will call $\mathbf{Q}$ the Bohmian position of the particle and its time derivative (10.17) the Bohmian velocity, but for simplicity the word “Bohmian” will often be omitted.

For a system made of N particles, the λ variable summarizes the $3N$ components of $\mathbf{Q}_1, \mathbf{Q}_2, \dots, \mathbf{Q}_N$, which evolve according to a direct generalization of (10.17) in configuration space: the time derivative of each $\mathbf{Q}_i$ is obtained by replacing the gradient ∇ by the gradient $\nabla_{\mathbf{r}_i}$ of the partial derivative with respect to the corresponding variable in the wave function – see for instance (10.19). An important point is that there is no retroaction of the additional variables onto the wave function; the coupling goes one way only. From the beginning, the theory therefore introduces a marked asymmetry between the two mathematical objects that are used to describe the system; we will see later that they also have very different physical properties.

The quantum velocity term depends only on the gradient of the phase of the wave function, not on its modulus. Therefore, vanishingly small wave functions may have a finite influence on the position of the particles. With a Gaussian wave packet for instance, the influence of the wave packet on the velocity of the particle is comparable near the center of the wave packet or at arbitrarily large distances, where the wave function is vanishingly small. Of course, situations where the position of the particle is extremely far from the center of the wave packet are very rare, but when they occur by chance, the position is guided exactly with the same efficiency in all space; this can be seen as a sort of non-local effect (we come back to more pronounced non-locality effects in §10.6.1.c).

In order to make his theory more similar to classical mechanics, initially Bohm did not start directly from (10.17), but introduced the notion of a “quantum potential”; this potential depends on the wave function and adds its effect to the

¹⁶ We assume that no potential vector $\mathbf{A}(\mathbf{r}, t)$ is contained in the Schrödinger equation. In the presence of a potential vector, and for a particle of charge q , some changes are necessary: in the right-hand side of (10.17), $(\hbar/i)\nabla$ should be replaced by $(\hbar/i)\nabla - q\mathbf{A}(\mathbf{r}, t)$, so that $\hbar\nabla\xi$ becomes $\hbar\nabla\xi - q\mathbf{A}(\mathbf{r}, t)$.

usual potential $V(\mathbf{r})$ acting on the particle, and its value is:

$$V_{\text{quantum}}(\mathbf{r}) = -\frac{\hbar^2}{2m} \frac{\Delta |\Psi(\mathbf{r}, t)|}{|\Psi(\mathbf{r}, t)|} \quad (10.18)$$

(where Δ is the Laplacian). Using the Schrödinger equation for $\Psi(\mathbf{r}, t)$ and the usual Newton equation for the acceleration of the particle, one can then recover the quantum velocity term. The two points of view are equivalent: one can either directly postulate that (10.17) is always valid, or assume that it is valid at time $t = 0$ and add $V_{\text{quantum}}(\mathbf{r})$ to the usual potential. Subsequent versions of Bohmian mechanics have nevertheless often discarded the quantum potential in favor of a quantum velocity term only.

Quantum equilibrium condition

The theory also postulates an initial random distribution of the position variables $\mathbf{Q}_1, \mathbf{Q}_2, \dots, \mathbf{Q}_N$ that reproduces exactly the initial quantum probability distribution $|\Psi(\mathbf{Q}_1, \mathbf{Q}_2, \dots)|^2$ of standard theory for position measurements. This distribution is often called the “quantum equilibrium distribution”. The distribution is not due to a preparation of the system that is not sufficiently accurate, but postulated as fundamental. For a given realization of an experiment, there is no way to select which value of the position is actually realized inside the distribution; from one realization to the next, a new completely random choice of position is spontaneously made by Nature. This assumption conserves the fundamentally random character of the predictions of quantum mechanics [417].

The additional variables then depend on the wave function in two ways, through both their initial values and their evolution. Combining the Schrödinger equation with the form of the “quantum velocity term” (10.17) one can show that, if at time t the distribution of positions is equal to $|\Psi(\mathbf{Q}_1, \mathbf{Q}_2, \dots)|^2$, the equality also holds at time $t + dt$. This ensures that the property continues to be true for any time, and automatically provides a close contact with all the predictions of quantum mechanics¹⁷. In particular, under the effect of the quantum velocity term, the Bohmian positions are constantly dragged by the wave function and can never move away from it;

¹⁷ In 1953, Bohm nevertheless considered the possibility of theories that are not equivalent to quantum mechanics by introducing distributions differing from the quantum equilibrium distribution [415]. He then showed in a simple model that, under the effect of random collisions, an arbitrary probability density will ultimately tend to a distribution given by $|\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots)|^2$, so that its difference with quantum equilibrium is likely to remain very small at all times. In 1954 he and Vigier treated a more general case, by assuming irregular fluctuations of the fluid that guides the positions of the particles [418]. Later, Valentini took a different approach [419] inspired from the H-theorem in statistical mechanics, and introduced a “subquantum entropy” H to characterize any distribution differing from the quantum equilibrium condition; he then showed that a coarse-grained value of H can only increase in time, and that the maximal value is obtained for the quantum equilibrium. The latter then appears as an attractor. The reasoning can be seen as a derivation of the Born rule, which no longer appears as an independent postulate, but as a consequence of the dynamics. See [420] for more recent numerical simulations along the same line.

they remain in regions of space where it does not vanish, which ensures that neither the guiding formula (10.17) nor the quantum potential (10.18) become indeterminate. Another important consequence of this postulate is to ensure compatibility with relativity [421], since arbitrary distributions would introduce the possibility of superluminal signaling (Appendix H). Since the Born rule is a consequence of the quantum equilibrium, one can then consider that this rule is not an independent postulate of quantum mechanics, but merely a consequence of relativity.

One can then restore determinism, and assume that the results of measurements merely reveal the initial pre-existing value of the positions, chosen among all possible values in the initial probability distribution (we come back in more detail to this point in §10.6.1.c). This assumption solves many difficulties, for instance those related to understanding why quantum systems can manifest both wave and particle properties in interference experiments. The system always contains two objects, a wave and a particle; the wave may produce interference effects and guide the particle in a way that forces its trajectory to reproduce the interference pattern – nothing especially mysterious conceptually (below we study Bohmian trajectories in more detail). Similarly, in the negative experiment with a Mach–Zehnder interferometer discussed in §2.4, in all events a wave propagates in the two arms, uninfluenced by the position of the particle; the interference effect it produces at the output beam splitter is different, whether one of its components in one of the arms is absorbed by the object, or whether it is not, and the particle is guided accordingly. The final outcome of the experiment is just a result of an initial choice for the path of the particle and of its guiding by the wave at the output beam splitter; no paradox at all! The same is also true for the Schrödinger cat: depending on the exact initial position of a many-dimension variable λ , which belongs to an enormous configuration space (including the variables associated with the radioactive nucleus as well as all variables associated with the cat), the cat remains alive or dies. Nevertheless, decoherence will act in exactly the same way as in standard quantum mechanics, and make it impossible in practice to observe interferences with macroscopic objects in very different states (this is related to the notion of “empty waves, see §10.6.1.c); the reason is that the theory is built to be equivalent to ordinary quantum mechanics.

Description of physical reality

For anyone who is very familiar with standard quantum mechanics, but not with the concept of additional variables, at first they may look somewhat mysterious, since they require a drastic change in the way we usually reason in quantum mechanics. This may explain why they are sometimes called “hidden”, but this is only a consequence of our much better familiarity with standard quantum mechanics! In fact, these variables are less abstract than wave functions. The additional variables are those that are directly “seen” in a measurement, as opposed to the state vector which

remains invisible; it actually plays mostly an indirect role, through its effect on the additional variables. In the example of a particle creating a track in a bubble chamber, what we see directly on the photograph is the successive values of an additional variable, which is actually nothing but the position of the particle. By contrast, who has ever obtained a photograph of the wave function of a single quantum system?

To summarize, we have in this context a description of physical reality at two different levels:

- (i) First, one corresponding to the elements associated with the wave function (or state vector), which are not directly visible but can be influenced in experiments by applying fields (or moving walls), since the evolution of the state vector depends on a Hamiltonian that can be controlled by applying fields (and on boundary conditions). This evolution takes place in a space with high dimensionality: for instance, for a system made of N spinless particles, the wave function evolves in a space with dimension $3N$. The state vector is not sufficient to give a complete description of a physical system.
- (ii) Second, another corresponding to the additional variables, the Bohmian positions, which are visible in experiments and move in ordinary three dimension space. It is nevertheless impossible to manipulate them directly, and for instance to change their distribution from quantum equilibrium. The reason for this impossibility is fundamental, otherwise it would become possible to send superluminal signals, in contradiction with relativity [421]. The Bohmian positions can only be manipulated indirectly, through actions on the state vector, which then guides their velocities.

The two levels together are necessary and sufficient for a complete description of reality¹⁸. We have already mentioned that there is no retroaction of the additional variables onto the state vector, which creates an unusual situation in physics: usually, when two physical quantities are coupled, they mutually influence each other¹⁹. Another unusual feature is that the effect of the field on the particle position does not depend on the intensity of the field, but only on its relative variations on space. Interestingly, we are now facing another sort of duality, which distinguishes between direct action on physical systems (or preparation), directly related to the state vector, and results of observations performed on them (results of measurements), determined by the additional variables.

¹⁸ Some authors prefer to consider that Bohmian positions only are real, and not the wave functions. But, since the wave functions act directly on the positions, and even mediate our actions onto them, it seems delicate to consider that they are less real than the positions. See also Bell’s quotation in §10.6.1.e.

¹⁹ Variants of the Bohmian theory including a retro-action of the positions onto the wave functions have nevertheless been studied [422].

10.6.1.b Bohmian trajectories

As soon as particles get a time-dependent position, they also get a velocity, an acceleration, etc. and a trajectory as in classical physics. By studying these trajectories, one obtains a variety of interesting, sometimes unexpected, results. For instance, even in the simplest case, a single particle in free space, the trajectories are not necessarily simple straight lines [423, 424]; they may be curved in a complicated way. We now study a few situations leading to characteristic Bohmian trajectories.

One particle

A first remark is that (10.17) gives zero velocity for any wave function that is real; a non-zero velocity can occur only if the phase of the wave function varies in space. As a consequence, for instance in the ground state of the Hydrogen atom, the Bohmian position of the electron does not turn around the proton as one could expect, but remains static at the same place. The effect of the quantum potential exactly cancels the attraction of the proton when the electron is in its ground state, so that it experiences no force. Similarly, for an harmonic oscillator in the ground state (or in any stationary state), the position of the particle does not oscillate in the potential, but remains at the same place.

This property is general: each time the Hamiltonian is invariant under time reversal, one can choose a basis of stationary wave functions that are real, and for which the corresponding Bohmian velocity always vanishes²⁰. Of course, with arbitrary wave functions that are coherent superpositions of stationary states, the situation is different: under the effect of the phase changes induced by the Schrödinger equation, the position and the associated velocities become functions of time. An example is an oscillator in a coherent quasi-classical state, where the time evolution of the position reproduces the classical oscillation in the potential well. Moreover, even in real stationary states, the correlation functions of positions are time dependent because, in Bohmian theory, the effect of the measurement on the wave function has to be included (we come back to correlations between measurements at different times in the discussion of §10.6.1.e).

We now consider a usual interference experiment, as shown schematically in Figure 10.1: a source S emits, one by one, a series of single particles that reach a screen D through which two apertures have been created, so that interferences can be observed in region R on the other side of the screen. The experiment is symmetrical

²⁰ The hydrogen atom also has stationary wave functions that are not real, with a phase factor $e^{im_l\varphi}$, where φ is the azimuthal angle and m_l the quantum number associated with one component of the orbital momentum. For these states, the Bohmian trajectories do turn around the proton, as in the classical image. Nevertheless, in the absence of a magnetic field, the values $\pm m_l$ have the same energy and, by linear combination, one can build a basis of stationary wave functions that are real and correspond to zero Bohmian velocity. So, for a given degenerate energy, depending on the wave function one chooses, this velocity vanishes or not.

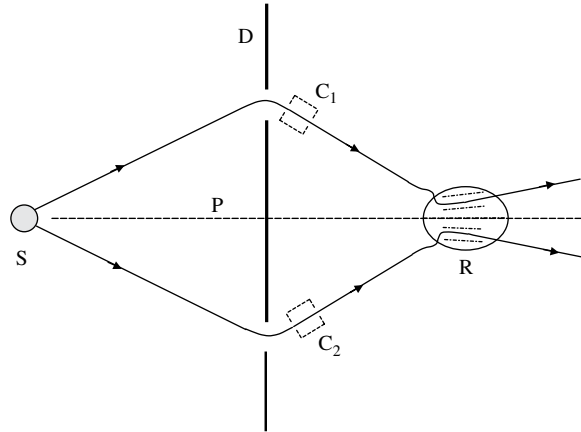


Figure 10.1. Bohmian trajectories associated with a particle emitted by a source S and crossing a screen D with two holes. In §10.6.1.e, two cavities C₁ and C₂ will be added into the scheme, but for the moment they can be ignored. Quantum mechanics predicts that the particle interferes with itself and produces an interference pattern in the observation region R; the dotted lines in this region symbolize the bright fringes. In Bohmian mechanics, the particle trajectories are not necessarily straight lines in free space; they bend inside the two holes under the effect of diffraction, and moreover oscillate in the interference region. The final result is that the predictions of quantum mechanics are exactly reproduced. One notices a “no-crossing rule”: the trajectories never cross the symmetry plane P (central horizontal dashed line).

with respect to an horizontal plane drawn as a broken line in the figure. We know that Bohmian mechanics is built to reproduce exactly the predictions of standard quantum mechanics concerning the measurement of the positions of particles; how will it manage to reproduce the interference pattern observed on the other side of the screen if one measures the position of the particle?

Bohmian theory solves the problem because (10.17) predicts that free particles may have curved trajectories; this feature is actually indispensable for the statistics of the positions to reproduce the usual predictions of quantum mechanics [425]. Two typical trajectories are shown in the figure. A first interesting effect occurs when the particle crosses the screen through a hole that is sufficiently small to diffract the wave function (its size is smaller than, or comparable to, the de Broglie wavelength): without touching the walls, the particle is deflected because the wave function that pilots its velocity is subject to diffraction. Moreover, after the screen and in the region where the states coming from the two holes overlap, bending effects again occur, since the velocity of the particle is affected by interference effects. The trajectories are then modified in a way that ensures that, when the experiment is repeated many times, the series of measured positions reproduces the quantum interference pattern exactly.

One can easily see by symmetry that, at points in the plane of symmetry of the experiment (central horizontal dashed line of Figure 10.1), the Bohmian velocity always lies within this plane: no trajectory of the particle can cross the symmetry plane. This result is sometimes called the “no-crossing rule” of Bohmian trajectories. Consider now two wave packets, each coming from a different hole, and crossing each other in the region of the symmetry plane. The no-crossing rule makes the trajectories “bounce” on the symmetry plane so that, after the wave packets have crossed each other, the trajectories that followed one wave packet have “jumped” to the other. As a consequence, after the screen, any trajectory that is above the symmetry plane necessarily went through the upper hole, and conversely.

See [47] for a discussion of the trajectory of a particle that can be reconstructed from the observation of its successive positions in a cloud chamber, and Chapter 5 of [424] for a discussion of Bohmian trajectories through a potential barrier (tunnel effect). Similar effects occur when particles propagate inside a Mach–Zehnder interferometer; in the absence of a final beam splitter, their position jumps from one wave packet to the other at the output, making their trajectory look like a zig-zag [426].

Two particles For systems of two particles or more, the situation becomes even more interesting. The velocities of the two particles are given as a function of the two-particle wave function $\Psi(\mathbf{r}_1, \mathbf{r}_2; t)$ by:

$$\begin{aligned} \frac{d}{dt}\mathbf{Q}_1 &= \frac{1}{m_1 |\Psi(\mathbf{Q}_1, \mathbf{Q}_2; t)|^2} \operatorname{Re} \left[\frac{\hbar}{i} \Psi^*(\mathbf{Q}_1, \mathbf{Q}_2; t) \nabla_{\mathbf{r}_1} \Psi(\mathbf{Q}_1, \mathbf{Q}_2; t) \right] \\ \frac{d}{dt}\mathbf{Q}_2 &= \frac{1}{m_2 |\Psi(\mathbf{Q}_1, \mathbf{Q}_2; t)|^2} \operatorname{Re} \left[\frac{\hbar}{i} \Psi^*(\mathbf{Q}_1, \mathbf{Q}_2; t) \nabla_{\mathbf{r}_2} \Psi(\mathbf{Q}_1, \mathbf{Q}_2; t) \right] \end{aligned} \quad (10.19)$$

where m_1 and m_2 are the masses. Since the Schrödinger equation remains unchanged in Bohmian theory, the wave function still propagates in the configuration space, as in standard quantum mechanics; on the other hand, the Bohmian positions propagate in ordinary three-dimensional space. As a result, non-local propagation effects may appear.

An interesting feature of the theory is that, even for interacting particles, their two positions are not directly coupled. The interaction Hamiltonian affects the wave function, which in turn guides the positions, but the world of Bohmian positions remains completely free of mutual interactions.

If the wave function $\Psi(\mathbf{r}_1, \mathbf{r}_2; t)$ is a product:

$$\Psi(\mathbf{r}_1, \mathbf{r}_2; t) = \varphi(\mathbf{r}_1, t) \chi(\mathbf{r}_2, t) \quad (10.20)$$

it is easy to see that (10.19) simplifies into:

$$\begin{aligned}\frac{d}{dt}\mathbf{Q}_1 &= \frac{1}{m_1 |\varphi(\mathbf{Q}_1, t)|^2} \operatorname{Re} \left[\frac{\hbar}{i} \varphi^*(\mathbf{Q}_1, t) \nabla_{\mathbf{r}_1} \varphi(\mathbf{Q}_1, t) \right] \\ \frac{d}{dt}\mathbf{Q}_2 &= \frac{1}{m_2 |\chi(\mathbf{Q}_2, t)|^2} \operatorname{Re} \left[\frac{\hbar}{i} \chi^*(\mathbf{Q}_2, t) \nabla_{\mathbf{r}_2} \chi(\mathbf{Q}_2, t) \right]\end{aligned}\quad (10.21)$$

Each particle then propagates independently, guided locally by its own wave function.

If the wave function is not a product, the velocities have to be evaluated at a point of configuration space that depends on the positions of both particles; the result for the velocity of particle 1 may then depend explicitly on the position of particle 2, and conversely. Suppose for instance that the two-particle wave function can be written as a sum of two products:

$$\Psi(\mathbf{r}_1, \mathbf{r}_2; t) = \alpha \varphi(\mathbf{r}_1, t) \chi(\mathbf{r}_2, t) + \beta \varphi'(\mathbf{r}_1, t) \chi'(\mathbf{r}_2, t) \quad (10.22)$$

and let us study the Bohmian velocity of particle 1; when do non-local effects occur? This depends on the overlap of the single-particle wave functions in the two terms of (10.22).

- (i) Consider a particular point $(\mathbf{Q}_1, \mathbf{Q}_2)$ of a trajectory in configuration space. If, at time t , one of the two wave functions $\chi(\mathbf{Q}_2, t)$ or $\chi'(\mathbf{Q}_2, t)$ vanishes, only one term of (10.22) is relevant in the expression of the velocity of particle 1 at point $\mathbf{Q}_1$. The wave function of particle 2 disappears from this expression, as in (10.21); particle 1 then propagates locally.
- (ii) If both wave functions $\chi(\mathbf{Q}_2, t)$ and $\chi'(\mathbf{Q}_2, t)$ are non-zero, then the same simplification does not take place, and non-local effects take place in general. Of course, this is not necessarily the case; if, for instance, it turns out that at point $\mathbf{Q}_1$ the two wave functions of the first particle have the same relative variation:

$$\frac{\nabla_{\mathbf{r}_1} \varphi(\mathbf{Q}_1, t)}{\varphi(\mathbf{Q}_1, t)} = \frac{\nabla_{\mathbf{r}_1} \varphi'(\mathbf{Q}_1, t)}{\varphi'(\mathbf{Q}_1, t)} \quad (10.23)$$

again a simplification occurs and the velocity at point $\mathbf{Q}_1$ is independent of position $\mathbf{Q}_2$. But, in general, this relation does not hold, so that the velocity of particle 1 is explicitly dependent on the position of the other particle.

To summarize, non-local effects of particle 2 on particle 1 may occur when the wave functions $\chi(\mathbf{Q}_2, t)$ and $\chi'(\mathbf{Q}_2, t)$ overlap, and if the position $\mathbf{Q}_2$ falls in the overlap region. If the two wave functions do not overlap, whatever trajectory is selected, one of the two components always vanishes; this component then plays

no role whatsoever in the evolution of the positions, and the particles move independently. The vanishing component²¹ of the wave function is what Bohm calls an “empty wave” [11].

Empty waves In Bohmian theory, empty waves do not occur only when two particles are entangled as in the example above. In the single-particle experiment described in Figure 10.1 for instance, each time a particle is emitted by the source, the initial position of the Bohmian variable determines which path it will follow and which hole it will cross. One component of the wave function then guides the particle while the other becomes an empty wave, playing no role as long as the two components do not overlap. Nevertheless, when later both waves recombine in the interference region R, they both contribute to the guiding of the particle and create the observed fringe pattern. During the interference phenomenon, the empty wave has, so to say, recovered its particle. Moreover, as we have seen in §10.6.1.b, the empty wave “catches” the particle while the wave packets cross; after the crossing, the wave packet that was empty becomes non-empty and conversely.

Something similar may happen with wave function (10.22), when two particles are entangled but functions χ and χ' have no overlap: if an appropriate measurement is performed on particle 2, standard quantum mechanics predicts that interference effects between the two components of the wave function can be restored, affecting subsequent position measurements performed on particle 1. The condition to recover interference effects²² is that the measurement should project the wave function of particle 2 onto a function $\chi''(\mathbf{r}_2, t)$ having non-zero scalar products with both functions $\chi(\mathbf{r}_2, t)$ and $\chi'(\mathbf{r}_2, t)$. Similarly, in Bohmian mechanics, if a measurement performed on particle 2 drives its position $\mathbf{Q}_2$ in a region of space where the two components contribute, a wave that was empty can play a guiding role again.

If eventually particle 2 becomes entangled with a large number of other particles, for instance those contained in a measurement apparatus (§10.6.1.c), the Von Neumann chain propagates too far and it becomes impossible in practice to restore interference effects. When decoherence becomes irreversible, empty waves remain empty forever.

Consider now a series of experiments providing a chain of successive results. For each realization of the series, the chain of results that is actually observed corresponds to a non-empty component of the state vector, which is called the “effective wave function” in [417]. All other chains of results (those that have not been obtained) are associated to empty waves, which still exist, but can have no

²¹ Of course, depending on the trajectory considered, it may be one or the other component that is empty.

²² This is sometimes called “two-particle interference” or “quantum eraser” effect. It is also the basic mechanism of the Bell inequality violation experiments discussed in Chapter 4, where particle 1 is in an incoherent statistical mixture before any measurement is performed on 2, but reaches a coherent superposition when correlated with a given result for particle 2.

effect on the future evolution of any Bohmian position. These empty waves never disappear from the solution of the Schrödinger equation (as the Everett interpretation, §10.11), but their fate is to remain forever inactive, without any observable effect. If the number of successive experiments is increased, the number of empty waves becomes very large, while only one wave remains non-empty (the effective wave function) and may have future physical effects. If for instance we consider the state vector of the universe, we clearly have to face an absolutely fantastic number of empty waves. Of course, they can be discarded without affecting the physical predictions in any way, but their status of reality is not obvious (we come back to this point in §10.6.1.f).

For other discussions of the role empty waves in the context of Bohmian theory or other interpretations, see for instance [58] and [427].

Constraints on Bohmian velocities The agreement between the predictions of standard quantum mechanics and the statistics over Bohmian trajectories rests on a relation where only the divergence of the velocity defined in (10.17) plays a role. If one adds to the velocity any vector with no divergence, any curl, the agreement remains unaffected. This remark gives the impression of an almost too large flexibility for Bohmian theories, and even to some extent of an insufficiently defined framework. But in fact more constraints exist than this simple agreement to build reasonable trajectories. See [428] for a discussion of the requirements introduced by Galilean invariance, which creates drastic limitations (while still leaving several possibilities for defining the velocities).

10.6.1.c Quantum measurement, non-locality

We now study in more detail how of the mechanism of empty waves ensures that the uniqueness of results of quantum measurements appears naturally in Bohmian theory. This success is obtained at the price of an explicit non-locality, which we will also discuss in this section.

Measurement in Bohmian theory In Bohmian theory, there is no need to introduce a special postulate for state vector reduction; it is already contained in the theory as a consequence of the mechanism of “empty waves”, which in turn is a consequence of the equations of motion and of the notion of trajectory. This solves the difficulties related to the definition of the border between the Schrödinger equation and state reduction. In other words, the Schrödinger equation applies equally well during all stages of a measurement process and at any other time.

To understand in more detail why, we consider a physical system after it has interacted with a measurement apparatus. It is then necessarily entangled with the

apparatus (§8.1), so that the wave function of both systems takes the form:

$$\Psi(\mathbf{r}, \mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N; t) = \sum_j \varphi_j(\mathbf{r}, t) \chi_j(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N; t) \quad (10.24)$$

where the $\varphi_j(\mathbf{r}, t)$ are mutually orthogonal states of the measured system associated with the various possible results. The $\chi_j(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N; t)$ are the corresponding states of the measurement apparatus (including the “pointer”), which are also orthogonal. Actually, they are not only orthogonal but also without spatial overlap – otherwise the observation of the position of the pointer would not provide an observable measurement. This form of the state vector is similar to (10.22), in the case where there is no overlap, so that the same discussion applies. We have seen (§10.6.1.a) that the Bohmian variables cannot “leave” the wave function (they cannot reach points of configuration space where it vanishes). After measurement, the variables $\mathbf{Q}_1, \mathbf{Q}_2, \dots, \mathbf{Q}_N$ associated with the measurement apparatus necessarily belong to one of the domains of configuration space where one of the wave functions χ_j does not vanish; since these functions have no overlap, they cannot belong to more than one of these domains. Therefore, for any possible trajectory, only one term of the sum (10.24) can play a role, all the others being “empty waves”. We have also seen at the end of §21 that, because of the very large number of variables associated with a measurement apparatus and its environment, it is impossible in practice to restore an overlap of all the components and to drive each of the many Bohmian positions to its specific overlap. These empty waves can therefore no longer influence the guiding of the position of the particles, and have become completely irrelevant to determine the results of future position measurements; they remain empty forever. It is then possible to retain only this non-empty term, without affecting in any way the future dynamics of the combined system. Moreover, one can normalize this term since, in Bohmian theory, the velocities are unchanged if the wave function is multiplied by a constant. The net result is that, for any realization of the experiment, corresponding to a given trajectory:

- a single result is obtained
- only one component of (10.24) plays a role in the calculation of the future motion of the Bohmian positions in configuration space (as well as in the calculation of probability of any possible future measurement performed on the system). State vector reduction is then reconstructed.

From the point of view of Bohm, the result of any experiment is just a consequence of the random initial position of the system in configuration space. The uniqueness of this result follows from the impossibility of a single point in configuration space to belong at the same time to more than one of the domains

associated with the components of (10.24). This automatically makes all components of the wave function vanish, except the one corresponding to the result of measurement. Determinism is restored in principle but, since this additional variable cannot be controlled in the preparation stage of the experiment – see above and point (ii) in §10.6.1.a – in practice quantum experiments give random results. While in standard quantum mechanics, the mechanism of decoherence (correlation with the environment) is not sufficient to explain the emergence of a single result in a single experiment, in Bohmian theory it is, thanks to the introduction of the Bohmian variables of the measurement apparatus and the mechanism of empty waves. This is a great achievement!

Non-locality In §20, we have seen that Bohmian positions evolve according to explicitly non-local equations. Non-locality is of course a surprising feature for any physical theory, but we should keep in mind that this feature is not artificially added into a theory that otherwise would be completely local: standard quantum theory is not perfectly local either. Actually, in configuration space, the equations of both theories are local, but non-localities may appear when one comes back to ordinary three-dimensional space.

Consider for instance two spinless particles and their six-dimensional configuration space. The formalism of standard quantum mechanics offers no local description of the physics taking place at the position of one single particle, even if the particles are contained in remote regions of space. The standard description is actually contained in the wave function²³, which has values that depend on the positions $\mathbf{r}_1$ and $\mathbf{r}_2$ of the two particles. If the wave function is not a product, the only proper characterization of the system therefore occurs in configuration space. For instance, the time evolution of the phase of this wave function depends on the values of the single particle potentials $V(\mathbf{r}_1)$ and $V(\mathbf{r}_2)$. The same is true of the probability current $\mathbf{J}$, which is defined in the same configuration space, and has a time derivative containing both potentials in general (nevertheless, for both theories, non-locality disappears from the equations as soon as a partial trace over the other particle is taken; in Bohmian theory, this involves an integration over all possible positions of the second particle).

The difference between the two theories is therefore not so much in the mathematical equations, which in neither case are really local. It is more that, in Bohmian theory, one ascribes physical reality to variables that evolve non-locally (positions),

²³ This is true in the Schrödinger picture. In the Heisenberg picture, the dynamics is contained in the evolution of operators; their matrix elements belong to a space with dimension equal to the square of that of the configuration space, which is even larger. If one uses second quantization or field theory, the dimensions become infinite (in field theory, coordinates in ordinary space time appear in the formalism as parameters defining the operators, but the space in which these operators act has infinite dimension).

while in standard quantum mechanics non-local evolutions occur only for physical quantities that are more indirectly related to physical reality.

10.6.1.d Spin and field theory

Bohmian theory can include treatments of spin and of quantum fields. Here we give only a brief summary and limit ourselves to a few simple examples.

Spin

Spins can easily be included in Bohmian theory within a non-relativistic treatment based on Pauli spinors. A simple method does not add any specific Bohmian variable associated with the spin, but just keeps the usual Bohmian position of the particle. Even within this simple context, interesting effects occur, but still in a way that avoids any contradiction with the predictions of standard quantum mechanics.

For a single particle with spin S , the quantum state can be defined in terms of a spinor having $2S + 1$ components $\Psi_\mu(\mathbf{r}, t)$. If $S = 0$ (spinless particle), the single component of the spinor is the wave function $\Psi(\mathbf{r}, t)$ already appearing in (10.17); if the particle has a spin $1/2$, it is described by two components $\Psi_\pm(\mathbf{r}, t)$ obeying the Pauli equation, etc. At each point of space, the Bohmian velocity of the particle is defined as the ratio of the local current of probability by the local density of probability, which are both obtained by a trace over the spin variable (sum over index μ). Equation (10.17) is then replaced by:

$$\frac{d}{dt}\mathbf{Q} = \frac{1}{m \sum_\mu |\Psi_\mu(\mathbf{Q}, t)|^2} \text{Re} \sum_\mu \left[\frac{\hbar}{i} \Psi_\mu^*(\mathbf{Q}, t) \nabla \Psi_\mu(\mathbf{Q}, t) \right] \quad (10.25)$$

Moreover, by calculating from the spinor in standard quantum mechanics the local average of the spin orientation, one can also define a local direction of the spin, which leads to an even more visual representation of the propagation of the particle.

(i) The effects taking place for a spinless particle have their counterpart for particles with spin. In particular, the direction of the spin does not necessarily remain constant along a trajectory, even in free space – for more details, see for instance Chapters 9 and 11 of [424], and [429]. As in [430], consider an atom entering a Stern–Gerlach apparatus that measures the component of the spins along Ox . If, initially, the spins of the particle point in a perpendicular direction Oy , their directions change progressively when the trajectory propagates in the magnetic field gradient, as shown schematically in Figure 10.2 (for clarity, the uniform precession of the spins around the homogeneous part of the magnetic field has been removed). At the output of the device, the spins are now either parallel or anti-parallel to Ox . The former case corresponds to the trajectories leading to spin result $+1$, the latter to those giving result -1 .

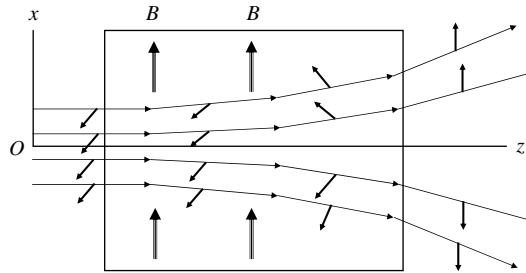


Figure 10.2. In a Stern–Gerlach experiment, atoms propagating along direction Oz enter a region (rectangle) where a high magnetic field B along direction Ox is applied. Its strong gradient acts on the magnetic moments of the atoms and creates a spin dependent deviation. Initially, the atoms have a transverse spin direction along Oy . While they progress in the gradient, the direction of their spin changes; when they exit the apparatus, two separate beams are formed with spins either parallel or anti-parallel to Ox . Four examples of Bohmian trajectories are shown (the directions of the arrows show the changes of the polar angle between the spin and Ox ; for clarity, the rapid precession of the azimuthal angle around B is not included).

Spin rotations along trajectories occur even for free particles; there is an equivalent of the “no-crossing rule”: if two wave packets associated with opposite spin directions cross in free space, along any Bohmian trajectory the direction of spin turns during the crossing and, after the position has jumped from one wave packet to the other, ends up in the opposite direction – an analogous effect to the exchange of wave packets by trajectories that we discussed above.

(ii) An interesting thought experiment considers one spin particle propagating in direction Oz that is sent through a series of spin analyzers (Stern–Gerlach systems for instance) with various orientations; for instance, the first analyzer measures the spin component along Ox , the second along a perpendicular direction Oy , the third along Ox again, etc. (at each step, the direction of measurement changes by 90°), as shown in Figure 10.3. The first analyzer divides the trajectories into two groups, those going to positive directions along Ox , and those going to negative directions, depending on the initial position of the particle. The second analyzer divides again each of these two groups of trajectories into two subgroups, depending again on the initial position of the trajectory. The same phenomenon repeats when more and more measurements are added. One could then hope to reach a situation where the determination of the initial position becomes sufficiently accurate to provide a determination of the trajectory of the particle in the next analyzer. Actually, this never happens: whatever number of analyzers is used, one never reaches a situation where the initial position becomes sufficiently well known to be able to predict the deviation observed in the next measurement with certainty: without any end, the result

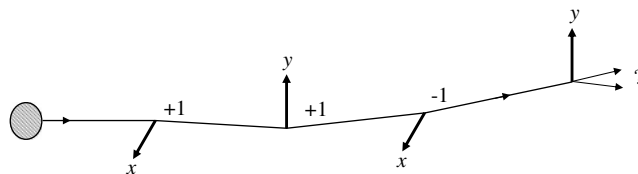


Figure 10.3. A particle with spin $1/2$ propagating in direction Oz crosses several Stern–Gerlach analyzers, which measure its spin component in succession along directions Ox , Oy , Ox , etc. Result $+1$ for the spin component is obtained if the deviation of the trajectory of the particle occurs in one direction, result -1 if it occurs in the opposite direction. These results are completely random, and provide information about the initial Bohmian position of the particle inside the wave packet associated with the wave function at the input of the system. Nevertheless, whatever the number of measurements is, this information is never sufficient to give any information on the result of the next measurement.

of the spin measurement remains completely random. The reason why this happens is that repeated measurements introduce an exponentially increasing sensitivity to the initial position of the particle, in analogy with chaotic situations in classical physics. At each step, the distribution of the Bohmian variables remains exactly the quantum equilibrium distribution. There is no way, even with many measurements, to determine the position sufficiently well to eliminate the fundamentally random character of quantum mechanics.

Wigner has pointed out the time irreversibility of the evolution of the additional variables in such situations [431]; Clauser has remarked that reversibility can be restored if one takes into account the spin polarization variables [432]. In addition, it is also possible to assume the existence of additional variables associated with the successive measurement apparatuses, as in “Von Neumann’s informal hidden variable argument” then, more randomization takes place at each measurement, so that it is no longer necessary to assume this extreme sensitivity to the initial conditions.

(iii) For two particles, consider an EPRB experiment of the type described in §3.3.1 and the evolution of the positions of the two particles when they are far apart. If particle 1 is sent through a Stern–Gerlach analyzer oriented along direction a , the evolution of its Bohmian position is obviously affected in a way that depends on a (the positions have to follow the quantum wave functions; in this case, it has the choice between two wave packets that are separated along axis a). But this will also change the position $(\mathbf{Q}_1, \mathbf{Q}_2)$ of the point representing the system in the six-dimensional configuration space. As we have seen above, if the orbital wave function is not a product, this changes the quantum velocity term for particle 2, in a way that then depends explicitly on a . Moreover, spins have to be taken into

account, including their correlations with positions. With a singlet spin state, the selection of an orbital wave packet for the first particle produces a projection of the spin state of the second particle onto a state that also depends on a (for more details, see Appendix I, second part). No wonder if such a theory has no difficulty in reproducing the non-local effects associated with violations of the Bell inequalities!

In this case, the advantage of introducing additional variables is to emphasize the effects of non-locality, while these effects often remain relatively hidden in the orthodox formalism (one more reason not to call these variables “hidden”!). Bell for instance wrote “it is a merit of the Broglie–Bohm interpretation to bring this (non-locality) out so explicitly that it cannot be ignored” – in fact, historically, he came to his famous inequalities precisely through this channel. As mentioned in §4.5.2, it is a matter of debate whether the standard interpretation of quantum mechanics just hides the non-locality it actually contains, or can be seen as a local theory despite the violations of the inequalities.

Bohmian field theory

In the second of his initial articles (Appendix A of [96]), Bohm included a brief discussion of field quantization, in the context of a study of the Compton effect. The close analogy between an harmonic oscillator and the dynamics of a free mode of the electromagnetic field in a cavity is a useful guide for introducing Bohmian trajectories within quantum electrodynamics – see also §§10.6 and 12.4 of [424]. Bohm introduces an additional variable $\mathcal{A}$ associated with the quantum operator A describing the vector potential, which plays the role of the “position” of the harmonic oscillator; the “wave function” in the representation where A is diagonal then guides $\mathcal{A}$ according to the usual pilot equation. To avoid questions related to gauge dependence, instead of introducing a Bohmian variable for the potential vector, one can also directly associate a Bohmian variable $\mathcal{E}$ to the electric field. This variable is guided by the “wave function” in the representation where the electric field is diagonal. Then $\mathcal{E}$ and its time derivative $d\mathcal{E}/dt$ directly describe the electric and magnetic (within a constant factor) fields associated with the mode considered. An inverse Fourier transform provides the values of the electromagnetic fields at every point inside the cavity. Similar methods may be applied to other bosonic fields²⁴.

For fermionic fields, a frequent method is to just add variables for particle positions, but not fields. The pilot equation remains exactly the same, all Bohmian positions being guided by a fully antisymmetric wave function. Combining this approach with the above description of bosonic fields, one then obtains a description of reality where bosons and fermions are treated in a very different way, the

²⁴ The difficulties associated with divergences in standard quantum field theory, and the necessity of introducing renormalization procedures, still occur in Bohmian theory, as one could expect.

former having Bohmian field variables and the latter position variables. The standard symmetry of second quantization is lost. But, after all, boson and fermion systems are indeed very different physically; there is no special reason why similar Bohmian variables should be used for both. Moreover, composite bosons (atoms for instance) made of an even number of fermions have a Bohmian position variable in this theory, the center of mass of the positions of the fermionic constituents.

Bell discussed another realist method to quantize bosonic and fermionic fields on a discrete space lattice [410]; it is actually on this occasion that he introduced the word “beable”. In this model, the deterministic guiding equation is replaced by stochastic transition probabilities, but a conjecture is that this stochastic feature should go away in some sense in the continuum limit. In [433], he extended this work to discuss the effect of annihilation and creation operators in Bohmian field theory, and the appearance of trajectories that can begin or end at points. For reviews on Bohmian field theories, see for instance [434] and [435] or, for relativistic theory, [436, 437].

10.6.1.e Objections and solutions

It is not possible here to discuss all aspects of Bohmian theory. It is not a “main stream” theory, and has certainly not been applied to all problems that have been treated successfully within standard quantum mechanics. These problems range from usual applications in condensed matter physics, quantum optics, etc. to quantum chromodynamics (theory of the strong interaction involving quarks), an essential component of our present physical description of the world. Since our purpose here is rather to focus on the interpretation, we will just discuss a few examples of applications of Bohmian theory. They have been selected because they involve situations that help clarify the physical content of the Bohmian point of view, as well as some frequent misunderstandings concerning this theory.

Are Bohmian trajectories real?

From one observation of the position of the particle, Bohmian trajectories can be reconstructed in a deterministic way for the past. For the future, it is nevertheless impossible to make more accurate predictions than those of standard quantum mechanics; the same fundamental indeterminism applies. In this sense, the explanations of quantum phenomena given by Bohmian theory are more retrodictive than predictive.

The authors of [429] propose to study an interference experiment in an interesting and intriguing case. The usual interference device is supplemented by electromagnetic cavities, which can store the energy of photons and be used as a “Welcher Weg” device (a “which way” device that tells the experimenter which hole the particle went through in an interference experiment). The first particle goes through

a screen having two holes, with two cavities sitting near the holes (C_1 and C_2 in Figure 10.1). If the particle goes through the upper hole, it leaves a photon in the first cavity²⁵; if it goes through the lower hole, it leaves it in the second cavity. Probing the number of photons in one of the cavities then allows one to know through which hole the particle went. A way to obtain the information is to observe the trajectory of a second particle, which is sent afterwards through this cavity, experiences the field it contains, and takes a trajectory that depends on the number of photons²⁶. At the end, one can compare the trajectories of the two particles, which can interact indirectly through a photon in a cavity, and to investigate the conditions under which they can influence each other.

The experiment is supposed to be symmetrical with respect to the horizontal plane P shown in Figure 10.1. The diffraction of wave function by each hole creates an angular spread so that, far beyond the screen, the component originating from any hole covers both sides of the symmetry plane. The authors of [429] then show that, in standard quantum mechanics, the first particle can leave a photon in the upper cavity, while the detection of its position gives a result that is below the symmetry plane. But, in Bohmian theory, the trajectories can never cross the symmetry plane (“no-crossing rule”): finding the particle below the symmetry plane P necessarily means that the corresponding trajectory went through the lower hole, even if a photon is left in the upper cavity. Eventually, when this photon is later detected by its influence on the trajectory of the second particle, the net result is that the two trajectories influence each other without ever coming close to each other. The first particle has left a photon in the cavity it never crossed! From this strange conclusion, the authors conclude that the Bohmian trajectories are “surrealistic”.

Of course, considering that trajectories may be called surrealistic is somewhat a matter of taste, and we will not discuss this issue. A first remark is that the paradox relies on the idea that the particles may be identified with their positions and trajectories only. This corresponds to a truncated form of Bohmian theory in which only the positions of the particles have physical reality, but not the wave function²⁷. But, in full Bohmian theory, one does not assume that only positions and trajectories contain all physical reality. On the contrary, the wave function becomes a real classical field, which exists objectively, being as real as a laser field for instance. As expressed by Bell [438]: “No one can understand this theory until he is willing to think of Ψ as a real objective field rather than just a probability

²⁵ We may assume that the particles are atoms in high Rydberg states, which can have very large dipole moments and can easily emit or absorb photons in superconducting cavities.

²⁶ A similar scheme is used in the experiment described in §9.3.

²⁷ This truncated theory would be intermediate between standard theory and Bohmian quantum mechanics, with a wave function that would keep the same status as in the standard interpretation (§1.2.3), combined with the addition of real positions.

amplitude”. Therefore, a “particle” always involves a combination of both a position and the associated field, and there is no reason whatsoever why the latter could not also influence its surrounding. Indeed, no fundamental argument limits mutual influences to the vicinity of the trajectory only.

So, within a full Bohmian theory, the simplest way to solve the apparent paradox is just to point out that the real field associated with the first particle can interact locally with the electromagnetic field in the cavity, leaving a photon in it; later this photon acted on the trajectory of the second particle. This explanation involves a crossed field–field–trajectory effect, and in these terms it is even perfectly local! Or, equivalently, if one prefers to use the notion of quantum potential, one can express the same idea by saying, as the authors of [439]: “the energy has been transferred non-locally to the cavity-field mode by the action of the quantum potential . . . a phenomenon no less real than the non-local correlations observed in EPR effects . . .” – see the figures of this reference for examples of plots of the trajectories.

After all, interaction between bodies mediated by fields occur in many domains of physics. The notion of action at a distance was introduced by I. Newton with his law of universal gravitation, which implies that the earth can attract the moon without the two mechanical bodies ever being in contact or even close. The earth creates a real gravitational field in all space, which then acts locally on the moon. In a similar way, here, the first particle creates a Bohmian field that crosses both cavities, and can create a photon in any of them, which, in turn, deviates the second particle²⁸.

We can also remark that, within the standard interpretation of quantum mechanics, “interaction-free measurements” are described in a way that is actually very similar to this Bohmian description. In §2.4, we have described an experiment where some events can occur only because the Schrödinger wave is absorbed in the arm of an interferometer, while the particle had only access to the other arm to propagate; the propagations of the wave and of the particle were also to some extent dissociated in these events. Standard interaction-free quantum measurement processes are not less “surreal” than Bohmian trajectories!

But we can also go further and make interesting connections with the appearance of correlations with the measuring apparatus and the famous Einstein–Bohr debate at Solvay conferences; this gives an idea of how partial traces (of density operators) appear in Bohmian theory. We now discuss how such correlation effects should be treated within this theory.

Measurement and entanglement

As we will see, the apparent paradox of “surrealistic trajectories” arises from an incomplete treatment of entanglement. This treatment assigns Bohmian positions

²⁸ Nevertheless, a curious property of this Bohmian field is that many trajectories are associated with the same field, while for the gravitational field different trajectories would lead to different values of the field.

to some entangled quantum systems, but not all of them, as would be required in a consistent Bohmian theory. The existence of the events studied by [429], where a particle leaves a photon in a cavity while its trajectory does not cross it, is derived from the “no-crossing rule” – Bohmian trajectories can never cross the horizontal symmetry plane of Figure 10.1. This is indeed true for a single particle, which propagates in a symmetrical environment with a symmetrical wave function. But, after the particle has crossed the cavities, the particle is entangled with the two fields in the cavities, and it is no longer possible to treat it as an isolated quantum system, with a single wave function propagating in three-dimensional space. A more precise investigation is necessary and shows that the no-crossing rule is no longer valid.

Two different cases can occur, depending whether a measurement is performed or not :

- If a measurement of the field in the cavities is performed, the analysis of §10.6.1.c directly applies: for each trajectory, one of the components of the wave function becomes “empty”, without any effect on the trajectory. The trajectory of the particle is then guided only by the component of the wave function that originates from one of the holes. The other component is empty and plays no role. Since the initial symmetry it then broken, the trajectories can then cross the symmetry plane, and the surreal events considered in [429] simply disappear, cancelling the paradox²⁹. See also the discussion given by Bell [425] of how consistent trajectories are generated in measurements within Bohmian theory. In Appendix I, we give a discussion of how correlations functions are affected by the first measurements in Bohmian theory.
- Even if no measurement is performed, when the particle propagates on the right side of the screen, we must take into account its entanglement with other physical systems. In Bohmian theory, when entanglement occurs, consistency requires that the position variables associated with all entangled sub-systems be included. We must therefore introduce position variables associated with the modes of the field in each cavity; since we do not wish to develop Bohmian field theory here, we will treat the modes of the field within a simple model: harmonic oscillators associated with massive particles in a potential. The position variable of the wave function of the initial particle is $\mathbf{r}$, that associated with cavity C_1 is $\mathbf{r}_1$, that associated with cavity C_2 is $\mathbf{r}_2$, corresponding in Bohmian mechanics to position variables $\mathbf{Q}$,

²⁹ The same conclusion applies to the spin version of the paradox also given in [429]; as soon as “Welcher Weg” (which way) detectors provide information about the path followed by the particle, one must include in the wave function the states of the pointer. The particles then go in straight lines in the region where the wave packets cross. They no longer bounce on the symmetry plane. The paradox therefore disappears again if a full Bohmian theory is used.

$\mathbf{Q}_1$, and $\mathbf{Q}_2$. The entangled situation is then described by a wave function:

$$\Psi(\mathbf{r}, \mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\sqrt{2}} [\varphi_{\text{upper}}(\mathbf{r}) \chi_e(\mathbf{r}_1) \theta_g(\mathbf{r}_2) + \varphi_{\text{lower}}(\mathbf{r}) \chi_g(\mathbf{r}_1) \theta_e(\mathbf{r}_2)] \quad (10.26)$$

where $\varphi_{\text{upper}}(\mathbf{r})$ and $\varphi_{\text{lower}}(\mathbf{r})$ are the wave functions of the particle that originate from the upper and lower hole respectively, $\chi_{g,e}(\mathbf{r}_1)$ the wave functions associated with the first harmonic oscillator in either its ground or first excited state, and $\theta_{g,e}(\mathbf{r}_2)$ the similar wave functions for the other oscillator.

The most probable locations of the positions $\mathbf{Q}$, $\mathbf{Q}_1$, and $\mathbf{Q}_2$ are those where $|\Psi(\mathbf{Q}, \mathbf{Q}_1, \mathbf{Q}_2)|^2$ takes relatively large values, which implies that at least one of the two components of (10.26) should have a significant modulus. A maximal value of the first component requires that $\mathbf{Q}_1$ should be chosen to give a maximal value to $|\chi_e(\mathbf{Q}_1)|^2$, while the second component requires that $|\chi_g(\mathbf{Q}_1)|^2$ should be maximal. But the functions $\chi_{e,g}(\mathbf{r})$ are different (they are actually orthogonal), and their maxima do not coincide; whatever the choice of $\mathbf{Q}_1$, at most one of these functions can be relatively large, while the other remains rather small³⁰. The same is true for variable $\mathbf{Q}_2$. Therefore, for any choice of $\mathbf{Q}_1$ and $\mathbf{Q}_2$, the probability can be significant only if one component of (10.26) dominates over the other. In other words, for a particular trajectory of the system (in nine-dimensional configuration space), there is a large probability that the two components should have different modulus. This breaks the initial symmetry and changes the interference pattern: the no-crossing rule is no longer valid. Some paradoxical trajectories (those where the trajectory goes through one cavity and excites the field in the other) still exist, but they require unlikely configurations of $\mathbf{Q}_1$ and $\mathbf{Q}_2$, which significantly reduces their probability.

But this is not the end of the Bohmian decoherence process: if we require some information on the state of the field in the cavities to be obtained, we need to couple it to other physical systems, a second test particle for instance. In turn, the position of this particle will be scattered in directions that depend on the state of the field in the cavities. The superposition written in (10.26) will then contain four functions and four variables, and by the same effect the unbalance between the two terms will be even more reduced for most trajectories (in 12-dimensional space). When this second particle reaches a measurement apparatus, the number of entangled degrees of freedom becomes so large that one reaches the same situation as above: one of the waves becomes empty forever, and the paradoxical trajectories disappear completely. A similar situation occurs if the test particle propagates in a bubble chamber; it then quickly becomes entangled with many other particles, and its Bohmian trajectory coincides with the trajectory detected by the excitation of these particles, in contrast with a statement in the introduction of [429].

³⁰ For instance, if $\mathbf{Q}_1$ happens to be at the maximum of the probability distribution for the ground state, the wave function of the excited state vanishes exactly, so that one of the terms in (10.26) completely disappears.

Connection with the Einstein–Bohr debate

It is interesting to also make a connection between this analysis and one of the famous paradoxes discussed by Einstein and Bohr at the Solvay meetings. Einstein raised an objection against the uncertainty principle, considering a interference experiment where the momentum recoil of the screen with the two interference holes is measured, so that “Welcher Weg” (which way) information is obtained (see for instance complement D_I of [440]). Bohr replied that one should take into account the quantum character of the screen, which is also subject to uncertainty relations; a proper calculation then shows that, if an interference can be observed, the state of the screen is necessarily such that the determination of the path of the particle is impossible. Einstein then agreed that the paradox was lifted. Here, we have a similar situation: momentum is also transferred from the particle to the screen, in particular while the trajectories are bent inside the holes³¹. Now, if the quantum state of the screen also changes depending whether the particle went through the upper or the lower hole, clearly it becomes necessary to include the corresponding Bohmian variables and the above argument applies again. One way to treat the problem³² is to consider that a phonon is emitted and propagates within the screen, and to calculate the microscopic changes of the positions of a macroscopic number of atoms in the screen (the recoil of each of these atoms is extremely small, but affects an extremely large number of terms in a product, so that the resulting effect is finite). We do not give more details here but, clearly, the fact that the Von Neumann chain inevitably propagates further than just the fields in the cavities reduces even more the weight of the unwanted component of the wave function. After a few steps, one expects that this component will vanish in practice, and that all paradoxical trajectories disappear. It is amusing to see that a historical argument due to Bohr should find another application in the context of Bohmian theory!

Our discussion of these examples can be summarized in two major points. The first is that one should always remember that Bohmian mechanics considers the wave functions as real physical quantities, not only the positions; as a consequence, the trajectories of two particles may influence each other, even if they never come close to each other (in an EPR type experiment for instance). This should not be considered as absurd, but just an illustration of the explicit character of non-locality in Bohmian theory – see the quotation by Bell above, as well as the discussion of this thought experiment by Griffiths [441]. There are many other

³¹ As for the fields contained in the cavities, they cannot recoil without transferring momentum to the wall of the cavities.

³² Another way to proceed would be to replace the two cavities by detectors, for instance bubble chambers, which contain particles having Bohmian positions that are changed if the test particle goes through the chamber. This change will immediately turn the other component of the wave function (the component where the test particle goes through the other detector) into an “empty wave”, which subsequently does not play any role. The non-crossing rule then disappears, as well as all paradoxical trajectories: the test particle can only leave a trace in the detector through which its trajectory passes.

examples illustrating that the quantum phenomena are indeed local in configuration space, but not necessarily in ordinary space. The second general remark is that, when entanglement is involved in the standard quantum mechanical treatment of the system, it becomes indispensable to take into account the role of all the Bohmian position variables associated with the various entangled systems; treating quantum entanglement correctly within standard theory, but in Bohmian theory forgetting to include all the corresponding variables and their effect – in other words treating the two quantum systems differently – is inconsistent.

Correlation between measurements at different times

Until now, we have only considered measurements made at one time, but quantum mechanics also gives predictions about measurements performed at different times; it is interesting to compare them to those of Bohmian theory [442]. In fact, at first sight, it is not obvious that the results will be the same; for instance, if one considers a system in a stationary state with a real wave function, the Bohmian velocity vanishes, so that the position $\mathbf{Q}$ remains constant. On the other hand, in quantum mechanics, the two-time correlation function in a stationary state has no reason to be time-independent: for a harmonic oscillator in its ground state for instance, it is well known that the two-time correlation function oscillates in time. One could then naively think that Bohmian theory predicts time-independent correlation functions, in contradiction with the time dependence obtained within standard quantum mechanics.

Actually, again a proper use of Bohmian mechanics shows that this contradiction does not exist. Within this theory, the effect of the first measurement must be taken into account – this is indispensable even if the two measured observables commute, and even if they correspond to independent systems. The first measurement correlates the position of the particle with that of a pointer; it becomes then necessary to take into account this second Bohmian position as well. Because the wave function of the whole system is no longer a product, situations similar to those discussed for two particles in §10.6.1.b then occur: the position of the measured system then takes a velocity that depends on the position of the pointer. This completely changes its future evolution, so that the correlation function is also modified. At the end of a correct calculation within full Bohmian theory, one recovers exactly the same results as in standard quantum mechanics – for more details, see Appendix I, which also discusses other experiments in the context of Bohmian mechanics [443–445].

The structures of the two theories are similar

The conclusion of this brief review of various objections to Bohmian mechanics is that, after all, its structure is more similar to that of standard quantum mechanics than one could think; all standard rules that are applied to the state vector have their

direct counterpart with the Bohmian velocities and positions. We should therefore not repeat mistakes that we have learnt to avoid in standard quantum mechanics. In standard theory, we know that, when the system becomes correlated with others, in particular measurement apparatuses, we must take this correlation into account in an entangled state vector. Similarly, in Bohmian theory, we must also take into account the correlations between the corresponding Bohmian variables – this introduces the phenomenon of “empty waves” (§10.6.1.c). We have learnt from Bohr’s historical arguments that the consistency of quantum mechanics requires that all parts of the experiments be treated as quantum systems, and for instance that one cannot ignore the recoil effect of parts of the apparatus (the screen). Similarly, the Bohmian description also requires that this recoil should be taken into account in terms of positions; if this principle is ignored, various inconsistencies appear.

But the similarities between the way standard quantum theory and the de Broglie–Bohm version should be used does not mean that they are also conceptually similar. There are actually deep conceptual differences. If they both share the same state vector, they attribute to it very different physical interpretations; moreover, the de Broglie–Bohm interpretation contains elements (positions, trajectories, local direction of the spin, etc.) that have no equivalent whatsoever in standard theory.

10.6.1.f Summary and discussion

One could summarize the main features of the de Broglie–Bohm version of quantum theory as follows:

- (i) it describes the same events in a richer way than standard theory, but without introducing any contradiction with its predictions. It often brings unexpected views on phenomena by introducing intricate arrays of trajectories, particles jumping from one wave packet to the other³³, etc., and leads to representations that may sometimes be found esthetically pleasing. It makes non-locality completely explicit but, as we have seen (§10.6.1.c), this non-locality is not totally absent from standard quantum mechanics either. This explicit character of non-locality may be the starting point of interesting new discoveries, as the Bell theorem has illustrated.
- (ii) it has often been rejected for bad reasons, starting from Pauli’s objections at the 1927 Solvay conference (inability to treat non-elastic collisions), which does not hold. Its predictions are indeed equivalent to those of standard quantum

³³ Of course, these pictures are not really specific of Bohmian theory. They can also be obtained within standard theory by drawing figures of the probability current. The difference is that Bohmian theory attributes more physical reality to these figures.

theory, but when entanglement is involved (or partial trace operations), for a correct treatment of the problem one must include the associated Bohmian positions.

- (iii) it brilliantly succeeds in eliminating the role of observers during measurements. With the mechanism of empty waves, measurements are considered as ordinary interaction processes; the uniqueness of the final result emerges without any external observation. State vector reduction is no longer a postulate, but is derived from this mechanism.
- (iv) concerning determinism, the situation is less clear-cut: the theory provides a scheme where the measurement process appears as deterministic, but where an average over an initial and uncontrollable “quantum equilibrium” distribution is necessary. The measurement reveals some pre-existing property of the Bohmian positions, in a retrodictive way. There is no possibility to control these positions in order to suppress the random character of the event, or even only reduce the indeterminism of standard theory. If one wishes, one can consider that the result of all experiments in physics is a consequence of the initial Bohmian positions of all particles of the universe, at the big bang, but one cannot say that indeterminism has really disappeared³⁴; it just gets a more concrete basis.
- (v) concerning realism, the theory does not succeed in re-introducing a simple, naive, realism. The reason is that, if one wishes to avoid incompatibility with relativity (avoid the prediction of possible superluminal signaling), one is forced to accept the existence of two levels of reality, as noted at the end of §10.6.1.a (see also Appendix H):
 - one level of reality (wave functions, or state vectors), described by a field that can be manipulated directly in experiments (by changing potentials, moving walls, etc.), but is not directly observable. It propagates in configuration space, which is much larger than the usual three-dimensional space (except for a single spinless particle).

This level includes empty waves, with a reality status that is intermediate and somewhat delicate³⁵ [427].

³⁴ In standard quantum mechanics, random events occur all the time – at least each time a measurement is performed somewhere in the universe. In Bohmian mechanics, a big single random event occurred at the creation of the universe; it is the source of all probabilities for future measurements

³⁵ As already discussed at the end of §21, one question is what level of reality should be attributed to empty waves associated with the results experiments that have not been obtained. These waves tend to constantly proliferate; for instance, the wave function of the universe develops an enormous number of empty branches, with a single non-empty branch that is the only one having any influence on future positions (the effective wave function). It is then tempting to decide that only the non-empty component of the state vector is a real Bohmian field. But this would amount to re-introducing the postulate of state vector reduction, while its disappearance was

- another level of reality (Bohmian positions) that is described by variables propagating in the usual three-dimensional space, and is directly observable. This level cannot be manipulated directly – it is actually completely interaction free. The corresponding variables are only guided by the field, which allows manipulating them indirectly, but without ever changing their quantum equilibrium distribution.

A sort of dualism is re-introduced at this stage, but an asymmetrical dualism since the former level of reality influences the latter, but not the reverse. Moreover, while positions propagate naturally in ordinary three-dimensional space, they are guided only by a field propagating in a completely different space, the abstract configuration space. The very fact that the two levels of reality are distinguished by the possible influence of experimental manipulations means that the observer, whose role was supposed to disappear, actually re-appears in the physical description of phenomena. In this sense, one could consider that the interpretation does not succeed in bringing the complete simplification of the interpretation that was initially intended. Moreover, we observe that the tension (without contradiction) between relativity and the standard theory has its counterpart in the de Broglie–Bohm interpretation: relativity is the source of the fundamental impossibility to manipulate Bohmian variables, which in turn ensures that the non-local evolution of these variables cannot be used for superluminal signaling.

Accepting or rejecting this interpretation therefore remains a matter of personal choice. It is not so surprising that even its inventors, de Broglie and Bohm, have themselves sometimes changed their views on the subject during their lives. A good point of comparison might be the theory of path integrals introduced by Feynman [446], which is usually considered as a branch of standard quantum mechanics. Most quantum physicists do not see any contradiction between this point of view and the use of the standard Hamiltonian quantification procedure and of the Schrödinger equation (or of the Heisenberg point of view); they are just complementary, and any of them may be used according to what is more convenient in every context. In the same way, Bohmian theory may be seen as still another possible point of view, complementary but not contradictory.

10.6.2 Nelson mechanics

Another formulation of quantum mechanics including additional variables was introduced in 1952 by Fenyès [447] and in 1966 by Nelson [448], in a somewhat different form. Instead of taking the Schrödinger equation for granted and adding

one main motivation of the theory; moreover, it would raise delicate questions concerning the exact stage at which an empty wave becomes empty forever.

positions to it, the idea is to derive the equation by introducing stochastic motions at a lower level. One considers point particles with mass m , each subject to a Brownian motion (Wiener process) with diffusion coefficient $\hbar/m$ and no friction; the effect of an external force deriving from a potential is included by simply using Newton's law. Non-linear equations are then obtained for two classical real variables, which can be grouped into one single complex variable. An appropriate change of variables³⁶ then results in a sort of “miracle”: a linear equation is obtained at the end. For appropriate values of the parameters, assuming a universal Brownian motion therefore leads to a natural derivation of the time-independent and time-dependent Schrödinger equations, including the possible role of an external electromagnetic potential.

One may consider this procedure as a purely mathematical method for quantization, without any particular physical consequence. But one may also attribute to it a more fundamental character, by considering the random positions and the Brownian motions of all particles as real. One then obtains another interpretation of quantum mechanics with additional variables, which is reminiscent of the Bohmian theory. One difference with this (and standard) theory is that the evolution of the wave function is no longer given by a postulate, but actually derived from other postulates that are considered more fundamental, including a “universal Brownian motion”. Another difference is that the Nelson point of view makes direct use of Newton's law, so that it remains much more classical than a theory where this classical law is modified. But, as in Bohmian theory, one introduces continuous trajectories for the particles, even if they are also not directly accessible by experimentation. Again, the formalism is built to lead exactly to the same predictions as standard quantum mechanics, so that a large part of our discussion concerning Bohmian theory can be transposed to Nelson mechanics. For the discussion of statistical mixtures in this context, see [449].

As in the de Broglie–Bohm theory, a central role is played by the Madelung hydrodynamic equations [416]. It has been pointed out [450] that these equations are not sufficient to ensure a single valued wave function, if this wave function has nodes; to ensure a single value, one has to add circulation quantization conditions around the nodes, in addition to the local equations of motion of the fluid. In stochastic mechanics as well, such additional constraints are necessary to ensure total compatibility with standard quantum mechanics, which makes the physical interpretation more complicated.

³⁶ The two classical real variables are the “current velocity” and the “osmotic velocity”, defined as the half sum and half difference of two velocities: the forward and backward time derivatives of the (non-differentiable) random position of the particle. The change of variables introduces Ψ , defined with a modulus equal to the square root of the probability density of the position, and with a phase gradient that is proportional to the current velocity.

Wiener processes have been used in many other domains of physics – we have already seen applications to quantum measurement (§8.3.2.c). They can actually provide a general quantization method. For instance, stochastic quantization has been applied in quantum field theory [451, 452], in particular by Parisi and Wu, who have proposed a quantization method with one additional time variable and a stochastic Langevin equation, and allowing gauge independent perturbation calculations [453, 454].

10.7 Modal interpretation

The words “modal interpretation of quantum mechanics” are used for a class of interpretations [455], introduced by Van Fraassen [456], which also attribute more properties to physical systems than standard quantum mechanics does – but, in this case, the position of particles plays no particular role. On the one hand, the additional properties are sufficiently restricted to avoid contradictions or inconsistencies with the predictions of standard quantum mechanics. On the other hand, they are sufficiently rich to lead to definite macroscopic events and results of measurement. A physical system may have a sharp value of a given observable even if its state vector in standard quantum mechanics is not an eigenstate of the corresponding operator. The state vector reduction postulate is then no longer necessary.

The main general idea, common to all this family of interpretations, is to associate two different states to any physical system:

- (i) the “dynamical state”, which is the usual quantum state (or density operator) of standard quantum mechanics. This state describes the evolution of the system, which, for an isolated system, is given by the Schrödinger equation. In the modal interpretation, the dynamical state never undergoes state vector reduction.
- (ii) a “value state”, which represents the physical properties of the system at any time (it may be tempting to call it “physical state”, but the tradition is to use the words “value state” in this context).

Consider a system S that is part of a larger system T , which, for the sake of simplicity, is assumed to be isolated. The dynamical and the value states describing S may differ if the dynamical state of S (density operator obtained by partial trace) is not a projector over a pure state (at least two of its eigenvalues do not vanish). In this case, the value state of S may contain more information about the physical properties of S than its dynamical state. A general limit is nevertheless put to how much information may be added in this way: the set of physical properties resulting from this more precise description should not exceed the maximum that is permitted by a standard description with any quantum state. In other words, one may well attribute to S a pure state even if this system is entangled with another quantum system (i.e. even if this system cannot be associated with a single pure

state in standard quantum mechanics), but it is not possible to go beyond this and obtain an even more precise description. One can then never attribute simultaneous sharp definitions to non-commuting operators, position and momentum for instance (as opposed to Bohmian mechanics, where a particle has a perfectly well defined position and velocity at every time). This provides a description with two state vectors that is somewhat reminiscent of the Bohm–Bub theory of measurement [7], to which we come back in the next section.

The question, now, is how to define the value state precisely; various possibilities have been suggested. Van Fraassen [456] initially proposed relatively general rules, assuming that the state value could be any state appearing in the decomposition of the dynamical state (it can be any linear combination of the eigenstates with non-zero eigenvalues of the partial density operator ρ_S defining the dynamical state of S). The definition of the value state then remains rather loose. Other authors have proposed to be more specific and to use the bi-orthonormal decomposition (Schmidt decomposition, §6.2.1) to write the entangled dynamical state of the whole system T as:

$$|\Psi\rangle = \sum_n c_n |\varphi_n\rangle \otimes |\Phi'_n\rangle \quad (10.27)$$

where the $|\varphi_n\rangle$ are normalized and mutually orthogonal, as well as the $|\Phi'_n\rangle$. The $|\varphi_n\rangle$ are then the possible value states describing system S . Of course, if all c_n but one are zero, within standard quantum mechanics system S is already in a pure state, and there is nothing new. But the modal interpretation postulates that, even when several c_n do not vanish, system S has the all properties associated with a single $|\varphi_n\rangle$. This postulate remains fundamentally non-deterministic: the only possible prediction is given by the probability $|c_n|^2$ for S to be in state $|\varphi_n\rangle$, which reconstructs the Born rule (at this stage, the use of the dynamical state cannot be avoided). But, even if it is impossible in advance to predict which of the accessible states will be reached, when it is reached, all sets of propositions about S that would be true if S was in state $|\varphi_n\rangle$ within standard quantum mechanics are indeed true. This point of view, where the dynamical state does not determine value states in an univocal way, but just provides possible values for them, is called “modal” because it leads to a modal logic of quantum propositions (§10.4).

Consider a measurement process involving system S and a measurement apparatus M . During the initial stage of interaction between S and M , both systems develop entanglement. This is sufficient to introduce a value state for S that differs from its standard description (with a density operator). From this moment, the interpretation guarantees that both sub-systems have all the properties associated with the emergence of a single result from the measurement (macroscopic uniqueness).

This program was later extended by Kochen [457] with more emphasis on the relational character of properties of physical systems: he considers that the value

states of S do not describe properties of S alone, but properties in relation to all the other systems that are entangled with it (including possibly measurement apparatuses). Dieks [458] introduces another view where systems do have intrinsic properties, and discusses how measurement processes and the existence of macroscopic behavior can be understood within the modal interpretation. He points out that this interpretation is realist in the sense that it assigns definite values to a set of magnitudes pertaining to a physical system (“beables” or “existents”), as opposed to results of measurement only in standard theory.

Healey [459] has proposed an “interactive interpretation”, which has similarities with a modal interpretation. He introduces a distinction between:

- (i) the “quantum state” of the system, which is the usual state of standard quantum mechanics.
- (ii) the “dynamical state” of the system, which contains information about all its current dynamical properties. “The dynamical state of a system at an instant may be identified with a truth value assignment to all sentences ascribing a quantum dynamical property to that system at that instant”. Nevertheless, the dynamical state does not generally also suffice to determine the future behavior of the system, nor even the probabilities of future behaviors.

The function of the quantum state is precisely to generate probabilities. It characterizes a system’s current probabilistic disposition, some of which will be manifested in future interactions, but does not suffice to determine its future behavior. Every quantum system always has a dynamical state, but not necessarily a quantum state (if it is entangled with another system). Note the difference in vocabulary: the dynamical state of the interactive interpretation is analogous to the value state of the modal interpretation, while the dynamical state of the modal interpretation is analogous to the quantum state of the interactive interpretation.

The role of environment is clearly essential in all these interpretations³⁷. For a measurement apparatus, the role of “pointer states” is important, see §8.1.2.b. Bacciagaluppi [460], Dickson [461], Berkowitz and Hemmo [462], and others have also made contributions to this interpretation, discussing for instance the role of imperfect measurements, the relation with the BKS theorem (§5.5), etc. Nevertheless, this interpretation has some difficulties, in particular with models of environmental decoherence assigning definite values to pointer positions [463] when they are continuous [460], with Lorentz invariance and contextuality [462, 464], and with applications in relativistic quantum field theory [465].

³⁷ Of course, the idea to create a correspondence between the physical properties of a system and its environment is not a specificity of this interpretation; it exists also, for instance, in the pragmatic views discussed in §§10.1.1 and 10.1.2.

10.8 Modified Schrödinger dynamics

Another way to resolve the coexistence problem between the two evolution postulates of quantum mechanics is to fuse them into a single postulate, by changing the Schrödinger equation itself in order to incorporate all features that are necessary to describe a measurement. One assumes that the equation of evolution of the wave function contains, in addition to the usual Hamiltonian terms, new terms (possibly non-linear and/or stochastic), which introduce the equivalent of state vector reduction when necessary [7, 13, 15, 344, 466]. The main objectives of this approach are:

- (i) Uniqueness of macroscopic reality: the new dynamics should be built in such a way that superpositions of states that are macroscopically distinguishable are never reached, in other words that Schrödinger cats (as defined in the original article) never occur. In particular, measurements made with macroscopic instruments must have definite outcomes.
- (ii) Compatibility with standard quantum mechanics at the macroscopic level: the probabilities of the different outcomes of measurements must reproduce the usual Born probability law.
- (iii) Compatibility with standard quantum mechanics at the microscopic level: the theory should remain consistent with the enormous amount of known experimental results where quantum mechanics has already been applied to various microscopic systems successfully.

Condition (i) implies that, for macroscopic superpositions (involving for instance pointers of measurement apparatuses), the new terms select one branch of the superposition, and cancel all the others. The postulate of state vector reduction of standard quantum mechanics then becomes useless; a single law of evolution is sufficient. In other words, macro- and micro-dynamics are unified. Condition (ii) adds that the selection of one branch must be (or seem to be) random in a way that exactly reproduces the probabilistic rules of standard mechanics and the effect of state vector reduction. Condition (iii) implies that the consequences of the change in the Schrödinger equation must remain extremely small in all situations involving microscopic objects only (atoms, molecules, etc.); this will immediately ensure that the enormous amount of successful predictions of quantum mechanics is preserved.

In other words, two extremes must be avoided: either to perturb the linear Schrödinger equation too much, and make interference effects disappear while they are still needed (for instance, recombination of the two beams at the exit of a Stern–Gerlach magnet if no decoherence has occurred), or too little and not ensure the complete disappearance of macroscopic indeterminacy (Schrödinger cats). The perturbation term must be efficient when (but not before) any microscopic system

becomes strongly correlated to a macroscopic environment, which ensures that significant decoherence has already taken place; the recovery of interference effects is then already impossible anyway within standard theory. When necessary, the process then reproduces the effect of the postulate of the wave function collapse, which no longer appears as an independent postulate, but as a consequence of the “normal” evolution of the wave function.

Including the emergence of macroscopic uniqueness in the fundamental dynamical equation of all physical systems solves the major conceptual difficulties of quantum mechanics. Obviously, this suppresses the problem of specifying the limits of application of two different postulates: measurement processes become ordinary interaction processes between two physical systems. Moreover, the state vector can then be said to directly describe physical reality, instead of being just a computational tool. The theory contains the elements that explain [14] “why events occur”! General discussions of the necessity of introducing the notion of events in quantum physics, and the “transmutation from possibilities to facts”, can be found in [467, 468].

To really define the theory, the above conditions are of course not sufficient; it is necessary to specify exactly what form the modified Schrödinger equation has, with no special guidance for doing so except the “border lines not to cross” (ii) and (iii) mentioned above. It is therefore not surprising that several versions of theories with modified Schrödinger dynamics should have been proposed.

10.8.1 Evolution of the ideas

It is possible to build versions of modified Schrödinger dynamics by changing the Schrödinger equation either without introducing additional variables, or by introducing them at the same time; both methods have been used.

10.8.1.a Early work

In 1966, Bohm and Bub [7] proposed a theory of measurement that contains a mechanism leading to state vector reduction. It belongs to the second category since these authors incorporate in their theory the additional variables previously considered by Wiener and Siegel [12], which are contained in a “dual vector”. The latter is a mathematical object that is actually very similar to the usual state vector $|\Psi\rangle$, but obeys an entirely different equation of motion – in fact, both vectors evolve with coupled equations. The probabilities of the results of measurement are just consequences of the initial random distribution of the dual vector. For some “normal” distribution of the new variables, the predictions of usual quantum mechanics are recovered; but it is also possible to assume the existence of “dispersion free” distributions that lead

to more precise, non-orthodox, predictions. As a consequence, this theory leads to a statistic that is richer than usual quantum mechanics. The authors discuss under which circumstances the effect of the new variables, which have not been observed yet, could be detected in the future.

In 1976, Pearle [13] introduced a version of the theory that belongs to the first category, where additional variables are not necessary. He considers the state vector of the combined system containing both the measured system and the apparatus, just after their interaction, when they are strongly entangled in a coherent macroscopic superposition (§8.1). The idea is to assume that the result of a measurement is determined by the relative phases of the components of this state vector; since these phases are not controlled experimentally with perfect accuracy, the final result of measurement then takes values that seem to be completely random. This behavior is obtained by adding a new non-linear term into the equation of evolution of the state vector, with a time rate constant λ ; this term drives the state vector to one of its macroscopic components only, in a way that depends on the initial relative phases. From this point of view, the randomness, postulated by standard quantum mechanics as fundamental, appears to be just apparent, being a mere consequence of experimental conditions that are never perfectly reproducible. Relatively little is added to the usual formalism of standard quantum mechanics, but the new term added in the equation of evolution should obey necessary conditions listed by the author:

- property 1: a single result is predicted by the new equation of evolution, meaning that, among a series of probability amplitudes, all of them go to zero after the interaction, except one that gets modulus unity – the equivalent of property (i) above.
- property 2: if the experiment is repeated many times, the selection of a single non-vanishing probability amplitude happens in a way that is consistent with the probabilistic predictions of quantum mechanics – the equivalent of property (ii) above.

In both the preceding theories, the reduction of the state vector is seen as a dynamical process which, as any dynamical process, has a finite time duration. We will see below examples of theories (GRW for instance, discussed in the next section) where this is not the case.

In 1979, Pearle proposed a different approach [14], this time by introducing stochasticity into the Schrödinger equation; the equation then becomes a stochastic differential equation with a term containing a randomly fluctuating operator, and acting on the state vector in a non-linear way. The presence of this new operator puts the theory into the category with additional variables. By contrast with the 1976

version, the selection of the outcome of the experiment is no longer due to uncontrolled phases, but to a randomly fluctuating operator. The mathematical process through which the new equation reduces the state vector to one component associated with a single result is described by a Fokker–Planck equation; it is analogous to a random walk between absorbing barriers – the corresponding statistical properties are sometimes described as those of the “gambler’s ruin game”. This version of Pearle’s theory is more general than the previous one; it nevertheless still has some problems, which he listed in a review article [469], including “preferred basis problem”, “interaction problem”, “trigger problem”, etc. In the Bohm–Bub theory, the complete reduction time was infinite but here, because the statistical process associated with the “gambler’s ruin game” has a finite mean duration, this time is finite [470].

10.8.1.b Spontaneous localization (“hits”)

In 1986, Ghirardi, Rimini, and Weber (GRW) introduced a new version of the theory, which provides an “unified dynamics for microscopic and macroscopic systems” [15]. They obtain this result by adding to the usual Schrödinger equation a random process, called “spontaneous localization” (SL), which suddenly changes the state vector by localizing the particles – as if the positions of the particles were measured at random times, but not with perfect accuracy (the localization remains approximate). Some features of this theory are actually reminiscent of the effect of continuous observations and measurements in standard quantum mechanics [471, 472] (§8.3.2), or that of approximate measurements. Spontaneous localization occurs at random times with a time constant that is adjusted so that, for macroscopic systems, the occurrence of superposition of far-away states is destroyed – condition (i) above; but this is true for macroscopic systems only, not microscopic systems – condition (ii) above. Earlier (1984), Gisin had also introduced modified Schrödinger dynamics with similar equations of evolution implementing the projection postulate [344], but where the collapse time of the state vector was infinite.

A remarkable property of the GRW approach is that it completely solves the “preferred basis problem” as well as the “trigger problem” mentioned above; the basis necessarily corresponds to position localized states. In this model, for individual systems³⁸ the localization processes are sudden (they are sometimes called

³⁸ For ensemble of systems, the discontinuities are averaged, and one recovers continuous equations of evolution for the density operators. Since most of the discussion of [15] is given in terms of density operators/matrices, and of the appearance of statistical mixtures (decoherence), one may get the (incorrect) impression that individual realizations are not considered in this work; but this is in fact not the case and “hitting processes” are indeed introduced at a fundamental level.

“hitting processes”), which of course makes them completely different from what happens within usual Schrödinger dynamics. Consider for instance a single particle described by a state vector $|\Psi(t)\rangle$. The effect of one of these processes is to suddenly replace $|\Psi(t)\rangle$ by a ket $|\Psi'(t)\rangle$ given by:

$$|\Psi'(t)\rangle = \frac{F_j |\Psi(t)\rangle}{\langle \Psi(t) | (F_j)^2 | \Psi(t) \rangle} \quad (10.28)$$

(the numerator of this expression ensures the conservation of norm of the state vector) where F_j is an Hermitian operator diagonal in the position representation. GRW assume that F_j localizes the particle around point of space $\mathbf{r}_j$ with an accuracy characterized by a free parameter α of the theory ($\alpha^{-1/2}$ is a length):

$$F_j = c e^{-\alpha(\mathbf{R}-\mathbf{r}_j)^2/2} \quad (10.29)$$

($\mathbf{R}$ is the position operator of the particle and c a real normalization factor). Now assume that all localization processes corresponding to the various $\mathbf{r}_j$ act in parallel, constantly and randomly localizing the particle at these points, each with a probability per unit time given by:

$$\lambda \langle \Psi(t) | (F_j)^2 | \Psi(t) \rangle \quad (10.30)$$

where λ is another free parameter of the theory (a rate, the inverse of a time). The condition:

$$\sum_j (F_j)^2 = 1 \quad (10.31)$$

ensures that the total probability for any sort of hits is independent of the initial state $|\Psi(t)\rangle$. The random effect of these localization processes is then added to the usual Schrödinger evolution. One then obtains a theory where, for each possible realization of the random hits at all times, the state vector follows a trajectory that is well defined, but different for each realization.

Consider the first localization process. If $\Psi(\mathbf{r}, t)$ is the wave function associated with state $|\Psi(t)\rangle$, according to (10.29) the probability of occurrence of the process with index j is:

$$\mathcal{P}_j = \lambda c^2 \int d^3r e^{-\alpha(\mathbf{R}-\mathbf{r}_j)^2} |\Psi(\mathbf{r}, t)|^2 \quad (10.32)$$

The localization processes are therefore more likely to be centered at values of $\mathbf{r}_j$ where the density of probability $|\Psi(\mathbf{r}, t)|^2$ is maximal; conversely, spontaneous localizations are very unlikely to occur around positions where the modulus of

the wave function is tiny. Whatever $\mathbf{r}_j$ is selected in the first localization process, the wave function after this process is multiplied by $e^{-\alpha(\mathbf{r}-\mathbf{r}_j)^2/2}$, which tends to restrict the wave function to a neighborhood of $\mathbf{r}_j$ with spatial extension $\alpha^{-1/2}$.

For the sake of simplicity, let us first assume that the wave function does not have the time to evolve before a second localization process takes place; because of the effect of the first localization, the second is then more likely to be centered at a point not too far from $\mathbf{r}_j$. Similarly, the third localization is likely to select a point that is in the neighborhood of the two preceding localization points, and so on: after a few localization processes, the wave function is well localized around a point $\mathbf{r}$ that is random, but well defined. This creates the spontaneous spatial reduction process of the wave function. Now, if the wave function evolves between the localization processes, the succession of the points of localization reconstructs a trajectory for the particle, as the track created in a Wilson or cloud chamber.

The scheme can immediately be generalized to a system made of N particles, by assuming that all the particles independently undergo localization processes. The operators F_j are then replaced by Hermitian operators F_j^i acting on the i -th particle, and their effect is summed over both indices i and j in the Schrödinger equation, but again for each realization one obtains a single trajectory for the state vector $|\Psi(t)\rangle$ describing the physical system. The effect of the random processes is to spatially localize the wave function around a single point of configuration space with $3N$ dimensions, with a spatial extension $\alpha^{-1/2}$ in every direction.

Consider now an ensemble of realizations of the same physical system, described by an ensemble density operator ρ . The average over all realizations of the hits provides the time evolution of ρ in the form:

$$i\hbar \frac{d\rho(t)}{dt} = [H(t), \rho(t)] + \lambda \sum_j [F_j \rho(t) F_j - \rho] \quad (10.33)$$

In the right-hand side of this equation, the first term containing the Hamiltonian $H(t)$ gives the usual Schrödinger–Von Neumann evolution. The second term describes the average effect of stochastic processes that replace the density operator by a new value $F_i \rho(t) F_i$ with a probability per unit time λ .

In quantum mechanics, localizing a particle is automatically associated with a change of its kinetic energy; the hitting process therefore “heats” the particle, necessitating a study of the role of energy conservation. In practice, this means that the time constant λ and localization constant α have to be adjusted to avoid the prediction of spontaneous heating effects that have never been observed. GRW propose choices of the parameters ($\lambda = 10^{-16} \text{ s}^{-1}$ and $\alpha^{-1/2} = 10^{-7} \text{ m.}$) such that the heating rate of a mole of ideal gas is approximately $10^{-15} \text{ K. per year}$, which is indeed compatible with known results.

Benatti *et al.* have checked that the main effect of the new term is to produce an actual reduction of the state vector [473]. A more general study of the consequences of introducing non-Hamiltonian evolutions of the state vector was given in [474], discussing in particular the possibility of introducing position and momentum localizations simultaneously; the conclusion of the study is that, in order to avoid unacceptable modification of the predictions concerning microscopic processes, one should limit oneself to spontaneous localization of positions only. A problem, nevertheless, of the GRW approach assumes independent localizations of all the particles, so that it does not preserve the symmetry requirements of quantum mechanics for identical particles, bosons, or fermions.

A general approach to derive spontaneous localization models has been proposed [475]: Event-Enhanced Quantum Theory (EEQT), where the quantum system S under study is coupled to a classical system C having a state described by a classical parameter α . In the simplest picture, α can take two values only, corresponding for instance to the two positions of the pointer of a measurement apparatus; but α may also take any series of discrete values, or even be a vector with many components having such discrete values. The value of α determines the evolution of the state vector $|\Psi(t)\rangle$ of S , which obeys an equation containing α -dependent non-Hermitian terms Λ_α (in addition to the usual Hamiltonian). The value of α changes randomly at times determined by the value of the average of Λ_α in state $|\Psi(t)\rangle$; as a consequence, the evolution of $|\Psi(t)\rangle$ depends on α and conversely. This model contains the GRW theory as a special case. Nevertheless, inasmuch as α is considered as describing a classical physical system, the model re-introduces the Bohrian notion of a frontier between classical and quantum worlds, while a motivation of theories with modified Schrödinger dynamics is precisely to suppress this frontier.

10.8.1.c Continuous spontaneous localization

In 1989, Pearle showed [476] that it is possible to solve this problem and to get rid of the discontinuous character of the hitting processes, while retaining the attractive features of the GRW model. This is obtained by adding terms of “continuous spontaneous localization” (CSL) to the usual Hamiltonian in the Schrödinger evolution. The new terms introduce Markov processes that depend on a series of random functions of time $w_j(t)$ with a broad frequency spectrum (white noise), and contain a time rate λ as well as a set of mutually commuting Hermitian operators A_j . For an appropriate choice of these operators, a full compatibility with the standard notion of identical particles in quantum mechanics is realized. One advantage of this point of view is that it introduces less radical changes with respect to the standard theory, since the state vector still evolves according to a differential equation. Nevertheless, this equation includes random functions of time as well as anti-Hermitian terms

(Itô stochastic differential equation); the norm of the state vector $|\Psi(t)\rangle$ is no longer conserved.

This non-conservation of the norm actually plays an important role in the theory, since the statistical properties of the random functions are defined precisely from this norm: one postulates (CSL probability rule) that the probability³⁹ for realizing any time dependence $w_j(t)$ that leads to a given value of $|\Psi(t)\rangle$ is nothing but $\langle\Psi(t)|\Psi(t)\rangle^2$. At this point, the non-linear character of the theory becomes obvious: the evolution of $|\Psi(t)\rangle$ depends on functions $w_j(t)$ which, in turn, have statistical properties constantly following the norm of $|\Psi(t)\rangle$ itself. Such a postulate strongly favors the realizations of the random functions that give a large norm to the state vector, while reducing the effect of all the others that give exponentially small values to the norm (even if mathematically they correspond to many more possibilities). One can then check that this choice of statistical properties remains consistent with the independent Markovian evolution of each realization of the state vector. A Fokker-Planck equation may be obtained for the time evolution of the probability densities.

In order to understand the mechanism of state reduction in this theory (described in more detail in Appendix J), let us first consider a simple case where one single operator A (with eigenvalues a_n) is introduced; we assume that A corresponds to the observable measured in a quantum measurement process such as that discussed in §8.1. Just after the end of the interaction between the measured system and the apparatus, both systems are entangled and described by the ket (8.7) of §8.1. Now, under the effect of anti-Hermitian terms controlled by the random functions $w_j(t)$, the modulus of each probability amplitude $c_n(t)$ fluctuates in time, instead of remaining constant. Among the large number of mathematically possible $w_j(t)$, according to the CSL probability rule, only a very small proportion may occur with a non-negligible probability: the proportion leading to a large sum over n of all $|c_n(t)|^2$. It turns out that, among these special functions, the most effective in providing a large norm for the state vector are those that give a large value to one $|c_n(t)|^2$ only. This is because, during its fluctuations, $w_j(t)$ can favor one value of n , but not several at the same time; situations where the fluctuations of $w_j(t)$ successively favor two (or more) of these coefficients lead to a dilution of the norm preservation effect and to an exponentially smaller value of the total norm at the end. We then obtain a process that selects a single outcome in a way that is reminiscent of the “gambler’s ruin game” of [14]. The fluctuations of the random functions break the symmetry between all possible measurement outcomes, so that this theory reproduces state vector reduction.

³⁹ This is somewhat reminiscent of postulating (10.30). Several realizations of the random function $w_j(t)$ may lead to the same value of $|\Psi(t)\rangle$ at time t ; the square norm of $|\Psi(t)\rangle$ then gives the probability of each of these realizations.

In the CSL theory, the operator A is actually not associated with any particular measurement process – this would make it measurement dependent and immediately destroy one of the achievements of the theory (giving no special role to measurements). One rather assumes that A is replaced by a series of position localization operators A_j , acting on all particles of the system, and localizing them at all possible positions in space (j then becomes an index for spatial positions, and thus may be continuous). As in GRW theories, perfect localizations of the particles would transfer an infinite amount of energy to them, which is physically unacceptable. One then postulates that the localization provided by each A_j is imperfect, over a spatial range $\alpha^{-1/2}$; all A_j are mutually commuting operators. Despite these changes concerning the definition of the operators, the essence of the localization process remains similar to that discussed above. It introduces a selection that eventually localizes each particle into a single random region of space – a spatial reduction process of the quantum state. For small quantum systems (single particle, atoms, molecules, etc.), the probability of occurrence of any collapse process remains extremely low for a very long time (λ is small). For macroscopic systems in quantum superpositions of two spatially distinct states, collapse is very likely to take place rapidly and to cancel one of the components; this is because all the particles involved are constantly subject to localization, while the localization of a single particle is sufficient to destroy one of the two components.

The functions $w_j(t)$ are considered as fundamentally random; as the Bohmian position variables, they cannot be manipulated directly, but they cannot be detected directly either (they could be called “hidden functions” instead of “hidden variables”). But, if one changes by hand the external parameters controlling the Hamiltonian (changing the magnetic field for instance), the state vector is affected, which may in turn change the statistical properties of the $w_j(t)$ and therefore affect these functions indirectly.

In 1990, the study of Markov processes and continuous spontaneous localization for identical particles was expanded by Ghirardi, Pearle, and Rimini [477]. They showed that, for an ensemble of systems, equation (10.33) is then replaced by an “Lindblad form” (§6.4.3) giving the time evolution:

$$i\hbar \frac{d\rho(t)}{dt} = [H(t), \rho(t)] + \frac{\lambda}{2} \sum_{j=1}^N \left[2A_j \rho A_j - (A_j)^2 \rho(t) - \rho(t) (A_j)^2 \right] \quad (10.34)$$

where the A_j are mutually commuting Hermitian operators (they may be arbitrary operators but, in this context, they are position localization operators). In general, discrete Markov processes in Hilbert space can be reduced, in the limit of small hits with infinite frequency, to a continuous spontaneous localization. For instance, if one chooses operators such that $A_j = F_j$ in (10.34), according to (10.31), one

recovers (10.33); this introduces an interesting correspondence between discrete (GRW type) and CSL models [477] in terms of the average evolution of density operators. But, conversely, the CSL theory does not require that the sum of squares of all operators should necessarily be 1, so that the most general Lindblad equation cannot be obtained from a hitting process. The CSL theories therefore offer a broader range of possibilities, which turns out to be useful for elaborating relativistic versions of the theory (§10.8.1.e) or in cosmology [478].

A problem nevertheless is that in CSL, again, a complete collapse of the wave function is never obtained in any finite time. Even for a macroscopic system, when most of the wave function goes to the component corresponding to one single outcome of an experiment, there always remain a tiny component on the others (even if extremely small and continuously decreasing in time). The existence of this component is not considered as problematic by the proponents of the CSL theory [479]. Nevertheless, Shimony has argued [480] that it is philosophically objectionable to associate physical reality with one single state, as long as the system remains in a quantum superposition of this state and another; he considers that this is true even if the weight of the former is enormous in comparison with that of the latter. This is often called the “tail problem” of modified Schrödinger dynamics; see [481] and for instance [482] for a proposed solution to the tail problem involving a combination of modified Schrödinger dynamics and Bohmian velocities, both stochastic.

An important common feature to all these theories is that they include new universal constants. They appear in the modified Schrödinger dynamics, and are adjusted to satisfy conditions (i) to (iii) of §10.8. For instance, we have seen that the GRW theory (SL) introduces a time scale λ^{-1} for the rate at which spontaneous localization takes place, and a length scale $\alpha^{-1/2}$ to characterize how narrow this localization is. These constants may in a sense look like *ad hoc* physical quantities, introduced only for technical reasons. But actually they have an extremely important conceptual role: they define the border between the microscopic and macroscopic world. The corresponding border line, which was ill-defined in the Copenhagen interpretation, is now introduced in a perfectly precise way. Conceptually, their role is therefore somewhat reminiscent of that of the Planck constant.

10.8.1.d Relation with gravity

In 1989, starting also from a study of quantum stochastic processes (“quantum Wiener processes” [12]) as models for state vector reduction [466], as well as from the treatment of continuous measurements [483], Diosi proposed an interesting modification of the GRW theory [484]. The new theory still provides a unification of micro- and macro-dynamics, but without requiring the introduction

of any new parameter. The general idea is to replace the previous spontaneous position localization processes by a stochastic process of mass localization, with a strength that is proportional to the Newton universal gravitational constant G ; this is the only parameter appearing in the new terms that are added to the usual Schrödinger dynamics. Diosi then introduced a treatment of the collapse of the wave function from a completely general law of density localization, which results in a parameter-free unification. One can then really speak of a universal mechanism for reduction!

Nevertheless, Ghirardi *et al.* [485] soon showed that this appealing approach also introduces serious problems at short distances, for which completely unrealistic predictions are obtained; for instance, atomic nuclei would receive so much energy that they could not stay in their ground state, but would rapidly be excited (or even dissociated). Ghirardi *et al.* propose a modification of Diosi's theory that solves these problems while still retaining the idea of implying gravitation. One of the two parameters of the GRW theory is eliminated with the help of a relation implying Newton's constant, but the other remains necessary, a constant having the dimension of a length. In this sense, the "universal" character of the mechanism is lost – the authors take this as a general indication that a new parameter is really necessary to solve the problem of the quantum theory of measurement.

Penrose has often invoked a relation between gravity and uniqueness of reality in quantum mechanics [486], in a more general and philosophical context including the notion of consciousness, and argued that "the quantum gravity threshold for self-collapse is relevant to consciousness".

10.8.1.e Relation with relativity

The localization process introduced in the GRW theory is not easy to describe in a relativistic way, since it assigns a special role to positions and their localization into a finite volume at a given time; the concept belongs more to Galilean, than Einsteinian, relativity. Nevertheless, relativistic versions of stochastic quantum dynamics and of the CSL theory have been developed [487, 488]. Moreover, in 2005 Pearle introduced a quantization of the classical random field of the CSL theory [489], leading to a "completely quantized theory of state vector collapse"; this resolves a problem related to the conservation of energy under the effect of collapse mechanism where particles gain energy, by assigning an energy to the random field, which loses energy so that the total energy remains constant. See also [490] and [491].

A general remark is that non-linearity and stochasticity must go together if one wants to avoid contradictions with relativity. For instance, if one introduces non-linear changes in the Schrödinger equation (at a microscopic level) while keeping

the standard reduction postulate of quantum mechanics (normally this postulate is no longer useful in these theories), faster than light signaling becomes possible, as pointed out for instance by Gisin [492]. Tumulka has proposed a relativistic version of the GRW theory [493], which he calls “flash ontology”, where the “local beables” (§10.6) are given by a discrete set of space-time points, at which the collapses are centered; these points have a random distribution that is determined by the initial wave function⁴⁰. For general or historical reviews on collapse and dynamical reduction theories, see [494] and [469].

10.8.1.f Relation with experiments

As we have seen, the fundamental motivation of theories with modified Schrödinger dynamics is to provide a unification of all kinds of physical evolution, including the emergence of a single result in a single experiment. They obtain this result by the introduction of new physical constants characterizing new physical mechanisms, which means that they are necessarily more specific about the conditions under which state vector reduction takes place. In other words, they are more predictive than standard quantum mechanics. This provides at the same time a strong constraint, namely compatibility with all known experimental results, but also interesting opportunities for testing new theory experimentally. This domain is therefore not limited to purely theoretical considerations.

Concerning possible conflicts with existing experiments, we have already mentioned the discussion of [485], which points out incompatibilities of Diosi’s universal form of the theory with known properties of microscopic objects. A similar case is provided by the generalization of quantum mechanics proposed by Weinberg [495], which he introduced as an illustration of a non-linearity that may be incompatible with available experimental data; see also [496] for an application of the same theory to quantum optics and [492] for a proof of the incompatibility of this theory with relativity, due to the prediction of superluminal communication (the proof is specific of the Weinberg form of the non-linear theory and does not apply to the other forms mentioned above).

Another possibility would be to detect the very weak spontaneous heating effect predicted by these theories. Under the effect of the collapse terms in the evolution equation, atoms and nuclei should become partly excited, and emit some weak radiation, which one could try to detect [497]. It turns out that several experiments of this sort have been performed for different purposes, for instance detecting radiation in Germanium crystals due to possible collisions with “dark matter”, or

⁴⁰ We recover one ingredient of the de Broglie–Bohm theory, but here we have collapse points in space-time instead of purely spatial position. Moreover, in Bohmian theory collapse does not occur.

two neutrino double- β decay [498]. For the moment, no experiment has been able to confirm or to exclude theories with modified Schrödinger dynamics.

Significant tests of these new theories could be given in the future by quantum interference experiments performed with objects containing many particles and going through two spatially distinct paths at the same time. While, in the experiments performed until now with small objects, the localization terms have a completely negligible effect, with larger objects they should reduce or even cancel the contrast of the observed fringes. With the presently proposed values of parameters λ and α , significant tests would be obtained with objects having 10^8 nucleons. Nevertheless, observing interference patterns with objects having the corresponding masses remains a very difficult experimental challenge with present technology. For a review of recent results obtained with clusters and molecules, and a discussion of future experiments, see [499].

10.8.2 Physical description of reality within modified dynamics

In these interpretations of quantum mechanics, the state vector no longer keeps the subtle intermediate status it had in standard quantum mechanics (§1.2.3): it now directly represents physical reality “in itself”. This reality evolves according to a unified dynamical theory, whether or not measurements, human observations, etc. are involved. Of course, this does not mean that the theory becomes more or less similar to classical mechanics. For instance, since the state vector evolves in a complex state space (Hilbert space), the description of reality should be made in this space, instead of the usual three-dimensional space; the two spaces are very different (not only because of their dimensions; even a classical configuration space with many dimensions remains very different from a Hilbert space). Nevertheless, the unification of the theory and of its dynamics is perfectly reached, which brings an enormous conceptual simplification.

Similar physical descriptions are obtained, whatever specific form of the non-linear dynamics theory is preferred. For instance, when a particle crosses a bubble chamber, the new terms create the appearance (at a macroscopic level) of a particle trajectory; they also select one of the wave packets at the measurement output of a Stern–Gerlach analyzer (and eliminate the other), but not before these packets become correlated with orthogonal states of the environment (e.g. detectors). Of course, a localization process of the wave function that operates in the space of positions, rather than in that of momenta, destroys to some extent the usual symmetry between positions and momenta in quantum mechanics. But this is actually not a real problem, since one can easily convince oneself that, in practice, what is measured in experiments is basically the positions of particles or objects (pointers, etc.); momenta are only indirectly accessible.

What about the Schrödinger cat and similar paradoxes? If the added non-linear term has all the required properties to properly mimic the state vector reduction, they are easily solved. For instance, a broken poison bottle must have at least some parts that have a different spatial localization (in configuration space) than an unbroken bottle; otherwise it would have all the same physical properties. It is then clear that the modified dynamics will resolve the components long before it even reaches the cat, which ensures the emergence of a single possibility. For a recent discussion of the effects of the modified dynamics on “all or nothing coherent states” (§5.2) in the context of quantum optics, and of the effects on perception in terms of the “relative state of the brain” (§10.11), see [500].

How is an EPRB experiment described in this context? In the case of Bohmian trajectories, we emphasized the role of the “quantum velocity term”, which has a value defined in configuration space, not in ordinary space. Here, what is essential is the role of the localization terms added into the Schrödinger equation and their action on the state vector. Intrinsically, these terms remain perfectly local; in the GRW theory for instance, at any time each particle may undergo a spontaneous localization process to any point of space where its probability density does not vanish. Consider two particles with spin propagating towards Stern–Gerlach magnets along orientations a and b . Nothing special occurs as long as each of the particles propagates towards a Stern–Gerlach analyzer, or even within its magnet, since the particles are microscopic and have an extremely weak localization probability. As in standard theory, they can perfectly well be in coherent superpositions, even if they are far away from each other. But as soon as particle 1 (for instance) hits a detector at the output of the magnet, the system develops correlations with the particles contained in the detector, an electronic current, the amplifier, etc. A Von Neumann chain then develops with two branches, each arising from one of the two wave packets split by the magnet. A macroscopic level is reached and the localization term becomes very effective, acting almost instantaneously. Assume for instance that the localization selects the branch where particle 1 has a $+$ component along direction a . Mathematically, this localization amounts to a projection of the two-particle state vector onto a component where the first particle is in the $+$ state along a ; because of the mathematical structure of the entangled state, this also projects particle 2 onto a spin state with component $-$ along the same direction⁴¹. As a consequence, the spin state of the second particle is projected exactly as if the state vector projection postulate had been applied. Non-locality is not introduced by the localization process alone, but by its action onto a state that already contains non-local entanglement. Since this point of view emphasizes the role of the detectors and not of

⁴¹ The equations are similar to those written in the second part of Appendix I in the context of Bohmian theory, except that they involve entanglement with the more than the other spin only.

the analyzers, it is clearly closer to the standard interpretation than the Bohmian interpretation. State vector reduction takes place later than in that interpretation, since entanglement with the experimental environment is required, and not only a magnetic splitting of one of the spin wave packets. Both interpretations emphasize the role of non-locality in a very explicit way, but with a different mechanism.

The program can be seen as a sort of revival of the initial hopes of Schrödinger, for whom all relevant physics should be contained in the wave function and its progressive evolution (see the end of §1.1.2); this is especially true, of course, of the versions of non-linear dynamics that are continuous (even if fluctuating extra quantities may be introduced), and less true of versions including “hits” that are somewhat reminiscent of the state vector reduction. Here, the state vector directly describes the physical reality, in opposition with our discussion of §1.2; we have a new sort of wave mechanics, where the notion of point particles is completely abandoned in favor of wave packets. The theory is different from theories with Bohmian additional variables because the notion of infinitely precise position in configuration space never appears. It is free of the difficulties mentioned in §10.6.1.f, since there is no need to introduce several levels of reality or a distinction between what can be observed and what can be manipulated; it nevertheless introduces in the evolution of the state vector stochastic functions (or stochastic hits) that cannot be controlled by human action. As we have seen, another important difference is that these theories with modified dynamics are really new theories: in some circumstances, they lead to predictions that differ from those of orthodox quantum mechanics, so that experimental tests might be possible. We should finally emphasize once more that, in these theories, the wave function can still not be considered as an ordinary field: it continues to propagate in a high dimension configuration space instead of the usual three-dimensional space.

A conclusion of our discussion of modified Schrödinger dynamics may be given by quoting Bell’s essay “Speakable and unspeakable in quantum mechanics” (Chapter 18 of [6]). Speaking of standard quantum mechanics, he writes: “The ‘problem’ then is this: how exactly is the world to be divided into speakable apparatus . . . that we can talk about . . . and unspeakable quantum system that we cannot talk about? How many electrons, or atoms, or molecules, make an ‘apparatus’? The mathematics of the ordinary theory requires such a division, but says nothing about how it is to be made⁴² . . . Now in my opinion the founding fathers (of quantum mechanics) were in fact wrong on this point. The quantum phenomena do *not* exclude a uniform description of micro and macro worlds . . . system and apparatus. It is *not* essential to introduce a vague division of the world of this kind”. Modified Schrödinger dynamics does provide an answer to these questions, and shows that a theory where the state vector directly describes the physical reality is perfectly

⁴² Compare, for instance, with the quotation by Landau and Lifshitz in §2.5

possible: quantum mechanics can indeed be modified without immediately introducing blatant contradictions with presently known results.

After all, if we truly believe that quantum mechanics is fundamentally non-deterministic, why should we request that its fundamental equation of evolution should remain deterministic, as is the standard Schrödinger equation? Is it not more natural to accept the existence of a small stochastic term in the evolution equation itself to account for this random character and unify all dynamics? The border between the situations where macroscopic uniqueness emerges, or does not emerge, is then no longer vague as in the standard interpretation, but precisely contained in the equations. Moreover, the theory opens the possibility for new physical phenomena occurring at this border. It is true that the precise form of the additional terms in the dynamics of quantum systems is not known; a whole class of possibilities exist, so that this is still a subject for exploration. But the only fact that one can build a theory without giving up realism is important, if only just as a proof of existence.

In particular, physics teachers do not necessarily have to introduce quantum mechanics to students with all the conceptual difficulties associated with the standard interpretation. In fact, if they wish they could take a simpler point of view, retaining the familiar idea that the properties of the physical world exist independently of observations, and considering that a physical system is really described by a wave function (or state vector) propagating in configuration space. They should admit that the evolution of this wave function obeys an equation which, for the time being, is still unknown in the most general case. What is nevertheless perfectly clear is that, in the limit of microscopic systems, this equation reduces to the Schrödinger equation, which describes these systems with an extreme accuracy. For larger systems, the equation contains terms that lead to the spontaneous emergence of macroscopic uniqueness; unfortunately, for the moment the precise form of these mathematical terms is still unknown. Several possibilities for these terms have been proposed, but the final form of the theory is still a subject of research and debate. But, after all, this situation is not worse than in the standard interpretation, where almost nothing can be said about the nature of state vector reduction! Eventually, the price to pay for this simplification of the general conceptual frame of quantum mechanics would be the introduction of more complicated mathematical equations, containing the reduction process. We can hope that, one day, experiments will be able to tell us which point of view is best.

10.8.3 Open quantum systems in standard quantum mechanics

One can also introduce stochastic terms into the Schrödinger equation without any fundamental purpose, just to obtain a convenient method of calculation [501–503].

These terms are used for the calculation of the evolution of a partial trace density matrix describing a sub-system, within an unchanged linear Schrödinger equation for the whole system. The method replaces a master equation for the partial density operator by the calculation of the evolution of a series of state vectors; these vectors are submitted to random perturbations, which may introduce sudden changes that mimic quantum jumps. Each time evolution of a state vector provides an individual quantum trajectory.

In some circumstances, it turns out that this approach saves computing time very efficiently. The method has found many useful applications, in particular in quantum optics. It is sometimes called “quantum trajectory method”, or “Monte Carlo wave function”, or still “quantum jumps simulation”; for a review, see [504]. In the limit of very small jumps, it becomes the method of “quantum state diffusion” [505, 506].

10.9 Transactional interpretation

The transactional interpretation of quantum mechanics also considers quantum states as real, rather than a mathematical representation of knowledge. It was proposed by Cramer in 1986 [507] as a generalization of previous work he had made in 1980 [508] on the possible role of advanced and retarded waves in EPR situations. It is known that, in classical electromagnetism, advanced waves (waves propagating from the future to the past) are solutions of the Maxwell equations, as well as the usual retarded waves. In non-relativistic quantum mechanics, since the Schrödinger differential equation is first order in time, it does not have this double type of solutions for a given energy. Nevertheless, in relativistic quantum mechanics, second order time equations replace the Schrödinger equation, so that in this case also advanced waves coexist with retarded waves. In transactional quantum mechanics, the microscopic exchange of a single quantum between a present emitter and a future absorber is described in terms of exchange of retarded and advanced waves between them.

This exchange is called a “transaction”, also described by Cramer as a “handshake” between the two participants of a quantum process. The emitter produces a retarded wave (called “offer wave”), which travels to the absorber, making it produce an advanced wave (“confirmation wave”), which travels back in time to the emitter and reacts on it. The cycle repeats itself until a standing wave regime is reached, in which destructive interference cancels the wave outside the time interval of the transaction (destructive interference between either the two retarded, or the two advanced, waves). The whole process has a finite extent in space and time, which means that in relativity it cannot be seen as an event (a point in space-time). As for the collapse of the state vector, it does not happen at any precise time in this

interpretation either; it takes place during the whole transaction, in a symmetric emission and absorption process. It is clear that the theory is explicitly non-local, so that it has no problem in explaining Bell-type correlations [508].

The mathematical elegance of this point of view makes it attractive. It nevertheless remains uncommon among physicists, probably because it requires to give up intuitive ideas about the past influencing the future and not the reverse, which may be considered as too high a price for a better understanding of quantum mechanics.

10.10 History interpretation

The interpretation of “consistent histories”⁴³ is also sometimes called “decoherent history interpretation”, or just “history interpretation” – here we will use the latter. As we will see, it proposes a logical framework to discuss the evolution of a closed quantum system, without any reference to measurements. The general idea was introduced and developed by Griffiths [17]; Omnès and Gell-Mann also contributed to it and sometimes adapted it [509–511]. The history interpretation is probably the most recent among those that are discussed in this book. We will remain within the limits of a non-specialized introduction; the reader interested in more precise information on the subject should go to the provided references – for a general presentation, see also an article in *Physics Today* [512], as well as the references contained therein, or the introductory review article by Hohenberg [513].

10.10.1 Histories, families of histories

Consider any orthogonal projector P on a subspace $\mathcal{F}$ of the space of states of a system; it has two eigenvalues, $+1$ corresponding to all the states belonging to $\mathcal{F}$, and 0 corresponding to all states that are orthogonal to $\mathcal{F}$ (they belong to the supplementary subspace, which is associated with the projector $Q = 1 - P$). One can associate a measurement process with P : if the state of the system belongs to $\mathcal{F}$, the result of the measurement is $+1$; if the state is orthogonal to $\mathcal{F}$, the result is 0 . Assume now that this measurement is made at time t_1 on a system that is initially (at time t_0) described by a density operator $\rho(t_0)$; the probability for finding the state of the system in $\mathcal{F}$ at time t_1 is then given by formula (10.9), which in this case simplifies into:

$$\mathcal{P}(\mathcal{F}, t_1) = \text{Tr} \{ \hat{P}(t_1) \rho(t_0) \hat{P}(t_1) \} \quad (10.35)$$

where $\hat{P}(t)$ is the projector P in the Heisenberg picture at time t . This result can obviously be generalized to several subspaces $\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3$, etc. and several

⁴³ The notion of consistency is essential at the level of families of histories, rather than at the level of individual histories.

measurement times t_1, t_2, t_3 , etc. (we assume $t_1 < t_2 < t_3 < \dots$). The probability that the state of the system belongs to $\mathcal{F}_1$ at time t_1 , then to $\mathcal{F}_2$ at time t_2 , then to $\mathcal{F}_3$ at time t_3 , etc. is, according to the Wigner formula (§10.1.2.a):

$$\mathcal{P}(\mathcal{F}_1, t_1; \mathcal{F}_2, t_2; \mathcal{F}_3, t_3 \dots) = \text{Tr} \{ \dots \hat{P}_3(t_3) \hat{P}_2(t_2) \hat{P}_1(t_1) \rho(t_0) \hat{P}_1(t_1) \hat{P}_2(t_2) \hat{P}_3(t_3) \dots \} \quad (10.36)$$

where, as above, the $\hat{P}_i(t_i)$ are the projectors over subspaces $\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3$ in the Heisenberg picture. We can now associate a “history” of the system with this equation: a history $\mathcal{H}$ is defined by a series of arbitrary times t_i , each of them associated with an orthogonal projector P_i over some subspace; its probability is given by (10.36) which, for simplicity, we will write as $\mathcal{P}(\mathcal{H})$. In other words, a history is the selection of a particular path, or branch, for the state vector in a Von Neumann chain, defined mathematically by a series of times and projectors. Needless to say, there is an enormous number of different histories, which can have all sorts of properties; some of them are accurate because they contain many times associated with projectors over small subspaces $\mathcal{F}$; others remain very vague because they contain only a few times with projectors over large subspaces $\mathcal{F}$ (one can even decide that $\mathcal{F}$ is the entire space of states, so that no information at all is contained in the history at the corresponding time).

There are in fact so many histories that it is useful to group them into families, or sets, of histories. A family is defined again by an arbitrary series of times $t_1, t_2, t_3, \dots$, but now we associate to each of these times t_i an ensemble of orthogonal projectors $P_{i,j}$ that, when summed over j , restore the whole initial space of states. For each time we then have a series of mutually orthogonal projectors that provide a decomposition of the unity operator:

$$\sum_j P_{i,j} = 1 \quad (10.37)$$

For each time t_i this gives the system a choice, so to say, among many projectors, and therefore a choice among many histories of the same family. It is actually easy to see from (10.37) and (10.36) that the sum of probabilities of all histories of a given family is equal to 1:

$$\sum_{\text{histories of a family}} \mathcal{P}(\mathcal{H}) = 1 \quad (10.38)$$

which we interpret as the fact that the system will always follow one, and only one, of them.

The simplest case occurs when a family is built from a single history: a trivial way to incorporate one history into a family is to associate, at each time

$t_i (i = 1, 2, \dots, N)$, in addition to the projector P_i , the supplementary projector $Q_i = 1 - P_i$; the family then contains 2^N individual histories. Needless to say, there are many other ways to complement a single history with “more accurate” histories than those containing the Q ; this can be done by decomposing each Q into many individual projectors, the only limit being the dimension of the total space of states.

10.10.2 Consistent families

The above definitions are in general not sufficient to ensure a satisfactory logical consistency in the reasonings. Having chosen a given family, it is very natural to also enclose in the family all those histories that can be built by replacing any pair or projectors, or actually any group of projectors, by their sum; this is because the sum of two orthogonal projectors is again a projector (onto a subspace that is the direct sum of the initial subspaces). The difference introduced by this operation is that, now, at each time, the events are no longer necessarily exclusive⁴⁴; the histories incorporate a hierarchy in their descriptive accuracy, including even cases where the projector at a given time is just the projector over the whole space of states (no information at all on the system at this time).

Consider the simplest case where two projectors only, occurring at time t_i , have been grouped into one single projector to build a new history. The two “parent” histories then correspond to two exclusive possibilities (they contain orthogonal projectors), so that their probabilities add independently in the sum (10.38). What about the daughter history? It is exclusive of neither of its parents and, in terms of the physical properties of the system, it contains less information at time t_i : the system may have either of the properties associated to the parents. But a general theorem in probability theory states that the probability associated to an event that can be realized by either of two exclusive events is the sum of the individual probabilities. One then expects that the probability of the daughter history should be the sum of the parent probabilities. But, in quantum mechanics, relation (10.36) shows that this is not necessarily the case; since any projector, $\hat{P}_2(t_2)$ for instance, appears twice in the formula, replacing it by a sum of projectors introduces four terms: two terms that give the sum of probabilities, as expected, but also two crossed terms⁴⁵ (or “interference terms”) between the parent histories, so that the probability of the daughter history is in general different from the sums of the parent probabilities. This difficulty was to be expected: we know that quantum mechanics is linear at the

⁴⁴ For these non-exclusive families, relation (10.38) no longer holds since it would involve double counting of possibilities.

⁴⁵ These crossed terms look very similar to the right-hand side of (10.36), but their trace always contains at some time t_i one projector $\hat{P}_{i,j}(t_i)$ on the left of $\rho(t_0)$ and one orthogonal projector $\hat{P}_{i,k}(t_i)$ on the right.

level of probability amplitudes, not probabilities themselves; interferences occur because the state vector at time t_i , in the daughter history, may belong to one of the subspaces associated with the parents, but may also be any linear combination of such states. As a consequence, a sum rule for probabilities is not trivial.

One way to restore the additivity of probabilities is to consider only families where the crossed terms vanish, which amounts to the condition:

$$\begin{aligned} \text{Tr} \left\{ \dots \widehat{P}_{3,j_3}(t_3) \widehat{P}_{2,j_2}(t_2) \widehat{P}_{1,j_1}(t_1) \rho(t_0) \widehat{P}_{1,j'_1}(t_1) \widehat{P}_{2,j'_2}(t_2) \widehat{P}_{3,j'_3}(t_3) \dots \right\} \\ \propto \delta_{j_1,j'_1} \times \delta_{j_2,j'_2} \times \delta_{j_3,j'_3} \times \dots \end{aligned} \quad (10.39)$$

Because of the presence of the product of δ in the right-hand side, the left-hand side of (10.39) vanishes as soon as at least one pair of the indices (j_1, j'_1) , (j_2, j'_2) , (j_3, j'_3) , etc. contains different values; if they are all equal, the trace merely gives the probability $\mathcal{P}(\mathcal{H})$ associated with the particular history of the family. In this way, we introduce the notion of “consistent family”: if condition (10.39) is fulfilled for all projectors of a given family of histories, we will say that this family is logically consistent, or consistent for short. Condition (10.39) is basic in the history interpretation of quantum mechanics; it is sometimes expressed in a weaker form, as the cancellation of the real part only; this, as well as other points related to this condition, is briefly discussed in Appendix K. We now discuss how consistent families can be used as an interpretation of quantum mechanics.

10.10.3 Quantum evolution of an isolated system

Let us consider an isolated system and suppose that a consistent family of histories has been chosen to describe it; any consistent family may be selected but, as soon as the choice is made, it cannot be changed (we discuss later what happens if one attempts to describe the same system with more than one family). This unique choice provides us with a well-defined logical frame, and with a series of possible histories that are accessible to the system and give information at all intermediate times t_1 , t_2 , ... Which history will actually occur in a given realization of the physical system is not known in advance: we postulate the existence of some fundamentally random process of Nature that selects one single history among all those of the family. The corresponding probability $\mathcal{P}(\mathcal{H})$ is given by the right-hand side of (10.36); since this formula belongs to standard quantum mechanics, this postulate ensures that the standard predictions of the theory are automatically recovered. For each realization, the system will then possess at each time t_i all physical properties associated to the particular projectors $P_{i,j}$ that occur in the selected history.

This provides a description of the evolution of its physical properties that can be significantly more accurate than that given by its state vector; in fact, the smaller

the subspaces associated to the projectors $P_{i,j}$, the more accuracy is gained⁴⁶. For instance, if the system is a particle and if the projector is a projector over some region of space, we will say that for a particular history the particle is in this region at the corresponding time, even if the whole Schrödinger wave function extends over a much larger region. Or, if a photon strikes a beam splitter, or enters a Mach–Zehnder interferometer, some histories of the system may include information on which trajectory is chosen by the photon⁴⁷, while standard quantum mechanics considers that the particle takes all of them at the same time. Since histories contain several different times, one may even attempt to reconstruct an approximate trajectory for the particle, while this is completely excluded in standard quantum mechanics (for instance, for a wave function that is a spherical wave); but of course one must always check that the projectors that are introduced for this purpose remain compatible with the consistency of a single family.

The physical information contained in the histories is not necessarily about position only: a projector can also project over a range of eigenstates of the momentum operator, or include mixed information on position and momentum (subject, of course, to Heisenberg relations, as always in quantum mechanics), information on spin, etc.. There is actually a huge flexibility in the choice of projectors; for each choice, the physical properties that may be ascribed to the system are all those that are shared by all states contained in the projection subspace, but not by any orthogonal state. A frequent choice is to assume that, at a particular time t_i , all $P_{i,j}$ are the projectors over the eigenstates of some Hermitian operator H : the first operator $P_{i,j=1}$ is the projector over all the eigenstates of H corresponding to the eigenvalue h_1 , the second $P_{i,j=2}$ the corresponding projector for the eigenvalue h_2 , etc. In this case, all histories of the family include an exact information about the value of the physical quantity associated at time t_i to H (for instance, the energy if H is the Hamiltonian). But, as already mentioned, we are not free to choose any operator H_i at any time t_i : in general, there is no reason why the consistency conditions should be satisfied by a family built just by choosing physical quantities in an arbitrary way.

Using histories, we obtain a description of the properties of the system in itself, without any reference to measurements, conscious observers, etc. This does not

⁴⁶ Obviously, no information is gained if all $P_{i,j}$ are projectors over the whole space of states, but this corresponds to a trivial case of little interest.

⁴⁷ Assume that, with a Mach–Zehnder interferometer, the family provides information about the path of the photon inside the interferometer. Then consistency requires that it gives no information on its output path (after the last beam splitter), and on which output detector it triggers at the end. This is because the probability of having a particle in each output path contains interference terms between the paths inside the interferometer, but that these interference terms vanish by sum over the two outputs.

One can also build other consistent families where the output channel is specified, as well as the excited output detector, but then consistency requires that no information about the path inside the interferometer is available. This is an illustration of how quantum complementarity applies within the history interpretation.

mean that measurements are excluded; but they can be treated merely as particular cases, by incorporating the corresponding physical devices in the system under study. Moreover, one attributes properties to the system at different times; this contrasts with the orthodox interpretation, where a measurement does not reveal any pre-existing property of the physical system, and moreover projects it into a new state that may be totally independent of the initial state. It is easy to show that the whole formalism of consistent families is invariant under time reversal, in other words that it makes no difference between the past and the future (instead of the initial density operator $\rho(t_0)$, one may use the final density operator $\rho(t_N)$ and still use the same quantum formalism [514]) – for more details, and even an intrinsic definition of consistency that involves no density operator at all, see §III of [515]. In addition, one can develop a relation between consistent families and semi-classical descriptions of a physical system; see [510] for a discussion of how classical equations can be recovered for a quantum system provided sufficient coarse graining is included (in order to ensure, not only decoherence between the various histories of the family, but also what the authors of this reference call “inertia” to recover classical predictability); see also Chapter 16 of [511] for a discussion of how classical determinism is restored, in a weak version that ensures perfect correlations between the values of quasi-classical observables at different times (or course, there is no question of fundamental determinism in this context). The history point of view undoubtedly has many attractive features, and seems to be particularly clear and easy to use, at least as long as one limits oneself to one single consistent family of histories.

10.10.4 Incompatibility of different consistent families

How does the history interpretation deal with the existence of several consistent families? They are a priori all equally valid, but they will obviously lead to totally different descriptions of the evolution of the same physical system; this is probably the most delicate aspect of the interpretation (we will come back to this point in the next subsection). The history interpretation considers that different consistent families should be considered as mutually exclusive (except, of course, in very particular cases where the two families can be embedded into a single large consistent family); any family may be used in a logical reasoning, but not combined together with others in general. The physicist is then free to choose any point of view in order to describe the evolution of the system and to ascribe properties to the system; in a second independent step, another consistent family may also be chosen in order to develop other logical considerations within this different frame; but it would be totally meaningless (logically inconsistent) to combine considerations arising from the two frames. This very important fundamental rule, somewhat

reminiscent of Bohr's complementarity, must be constantly kept in mind when one uses this interpretation. We refer the reader to [515] for a detailed and systematic discussion of how to reason consistently in the presence of disparate families, and to [516] for simple examples of incompatible families of histories (photon hitting a beam splitter, §II) and the discussion of quantum incompatibility (§V); various classical analogies are offered for this incompatibility, including a two-dimensional representation of a three-dimensional object by a draftsman, who can choose many points of view to make a drawing, but can certainly not take several at the same time – otherwise the projection would become inconsistent.

10.10.5 Comparison with other interpretations

In the history interpretation, there is no need to invoke conscious observers, measurement apparatuses, etc. The system has properties in itself, as in the Bohmian or modified Schrödinger dynamics interpretations. When compared to the others, a striking feature of the history interpretation is the enormous flexibility that exists for the selection of the points of view that can be chosen for describing the system; we have seen that all the times $t_1, t_2, \dots$ are arbitrary (actually their number is also arbitrary) and, for each of them, many different projectors P may be introduced. One may even wonder if the interpretation is sufficiently specific, and if this very large number of families of histories is not a problem.

What is the exact relation between the history interpretation and the orthodox theory? There is certainly a close relation, but several concepts are expressed in a more precise way with histories. For instance, complementarity stands in the Copenhagen interpretation as a general, almost philosophical, principle. In the history interpretation, it is related to mathematical conditions, consistency conditions, and orthogonality of projectors. It is impossible to come back to classical physics and to a simultaneous definition of all its observables: every projector cannot be more precise than the projector over a single quantum state $|\varphi\rangle$, which is itself obviously subject to the uncertainty relations because of the very structure of the space of states. Of course, it is still possible to make Bohrian considerations on incompatible measurement devices, or distinctions between the macroscopic and microscopic worlds, but with histories they lose part of their fundamental character. The history interpretation allows a quantum theory of the universe (compare for instance with quotation (viii) of §2.5); we do not have to worry about dividing the universe into observed systems and observers.

The bigger difference between the orthodox and the history interpretations is probably the way they describe the time evolution of a physical system. In the usual interpretation, we have two different postulates for the evolution of a single entity, the state vector, which may sometimes create difficulties; in the history

interpretation, the continuous Schrödinger evolution and the random evolution of the system among histories are put at very different levels, so that the conflict is much less violent. Actually, in the history interpretation, the Schrödinger evolution plays a role only at the level of the initial definition of consistent families (through the evolution operators contained in projectors in the Heisenberg picture) and in the calculation of the probability $\mathcal{P}(\mathcal{H})$; the real time evolution takes place between the times t_i and t_{i+1} and is purely stochastic. A purely non-deterministic evolution has now become the major source of evolution! There is a kind of inversion of priorities with respect to the orthodox point of view, where the major source of evolution is deterministic under the effect of the continuous Schrödinger equation. Nevertheless, and despite these differences, the decoherent history interpretation remains very much in the spirit of the orthodox interpretation; indeed, it has been described as an “extension of the Copenhagen interpretation”, or as “a way to emphasize the internal logical consistency of the notion of complementarity”. Gell-Mann prefers a more general point of view on the history interpretation [517], which makes the Copenhagen interpretation just “a special case of a more general interpretation in terms of the decoherent histories of the universe. The Copenhagen interpretation is too special to be fundamental ...”.

What about the relation with the “correlation interpretation”? It also seems to be very close, since both points of view give a central role to the Wigner formula. In a sense, this minimal interpretation is contained in both the Copenhagen interpretation and in the history interpretation. Some physicists favoring the correlation interpretation would probably argue that adding a physical discussion in terms of histories to their mathematical calculation of probabilities does not add much to their point of view: they are happy with the calculation of correlations and do not feel the need for making statements on the evolution of the properties of the system itself. Moreover, they might add that they wish to insert whatever projectors correspond to a series of measurements in (10.9), and not worry about consistency conditions: in the history interpretation, for arbitrary sequences of measurements, one would get inconsistent families for the isolated physical system, and one has to include the measurement apparatuses to restore consistency. We have already remarked in §10.1.2 that the correlation interpretation allows large flexibility concerning the boundary between the measured system and the environment. For these physicists, the history description appears probably more as an interesting possibility than as a necessity, but without introducing any contradiction either.

Are there also similarities with theories with additional variables? To some extent, yes. Within a given family, there are many histories corresponding to the same Schrödinger evolution and, for each history, we have seen that more information on the evolution of physical reality is available than from the state vector (or wave function) only. Under these conditions, the state vector appears as a

non-complete description of reality, and one may even argue that the histories themselves are indeed additional variables⁴⁸. In a sense, histories provide a kind of intermediate view between an infinitely precise Bohmian trajectory for a position and a very delocalized wave function. In the Bohm theory, the wave function pilots the position of the particles; in the decoherent history interpretation, the propagation of the wave function pilots rather the definition of histories (through a consistency condition) as well as a calculation of probabilities, but not the evolution between times t_i and t_{i+1} , which is supposed to be fundamentally random.

Two theories, one deterministic, the other completely indeterministic, are necessarily different conceptually. Nevertheless, it is not impossible to transpose some of the Bohmian ideas to the history interpretation, and make it deterministic. Consider a given consistent family for which, at time t_1 , the number of different projectors P_{j1} is Q_1 ; at time t_2 the number of different projectors is Q_2 , etc. One could for instance introduce an additional variable $x(t_i)$ that, by definition, always belongs to the $[0, 1]$ interval, and postulate that this variable provides a criterion to decide, among all projectors P_{ji} at each time t_i , which is “realized” (which projector defines physical properties that are associated with this history at this time). Mathematically, this can be done by dividing the $[0, 1]$ interval into Q_i smaller intervals I_{ij} (with $j = 1, 2, \dots, Q_i$) and associate the values of $x(t_i)$ within interval I_{ij} to the realization of projector P_{ji} . In this way, a “trajectory”, defined by the values of $x(t)$ at all discrete times $t_1, t_2, \dots, t_i, \dots$ defines a single history of the family. At the initial time t_1 , as in Bohmian theory one then could assume a probability distribution that reproduces the quantum predictions, and finally define a law of motion of the additional variable that ensures compatibility with the predictions of standard quantum mechanics⁴⁹ – in this case relation (10.36). By adding this variable $x(t)$ to the history interpretation, the latter could be made deterministic and, to some extent, compatible with the Bohmian ideas. Still, with the present states of the two theories, we can probably safely conclude this comparison between the Bohmian and history interpretations by saying that they give very different points of view on quantum mechanics.

Finally, what is the comparison with theories using a modified Schrödinger evolution? In a sense, they correspond to a completely opposite strategy, since they introduce into one single equation the continuous evolution of the state vector as well as a mechanism simulating the state vector reduction (when needed); by contrast, the history interpretation puts on different levels the continuous Schrödinger evolution and a fundamentally random selection of history selection by the system.

⁴⁸ They would then be family dependent, and therefore not EPR elements of reality, as we discuss later.

⁴⁹ This could be done either continuously by postulating some equation of evolution of $x(t)$, or by discrete steps by postulating a mapping of the interval $[0, 1]$ over itself corresponding to the translation in time from t_i to t_{i+1} (we note in passing a similarity with the notion of Poincaré maps in classical mechanics).

One might venture to say that the modified non-linear dynamics approach is an extension of the purely wave program of Schrödinger, while the history interpretation is a modern version of the ideas put forward by Bohr. Another important difference is, of course, that a theory with modified dynamics is not strictly equivalent to usual quantum mechanics, and could lead to experimental tests, while the history interpretation is built to reproduce exactly the same predictions in all cases – even if it can sometimes provide a useful point of view that allows to grasp its content more conveniently [316].

10.10.6 A profusion of histories; discussion

We finally come back to a discussion of the impact of the profusion of possible points of view, which are provided by all the families that satisfy the consistency condition. We have already remarked that there is, by far, no single way in this interpretation to describe the evolution of properties of a physical system – for instance, all the complementary descriptions of the Copenhagen interpretation appear at the same level. This is indeed a large flexibility, much larger than in classical physics, and much larger than in the Bohmian theory for instance. Are the rules that we have defined above (“no combination of points of view”) really sufficient to ensure that the theory is completely satisfactory? The answer to this question is not so clear for several reasons. First, for macroscopic systems, one would like an ideal theory to naturally introduce a restriction to sets corresponding to quasi-classical histories; unfortunately, the number of consistent sets is in fact much too large to have this property [325]. This is the reason why more restrictive criteria for mathematically identifying the relevant sets are (or have been) proposed, but no complete solution or consensus has yet been found; the detailed physical consequences of consistency conditions are still being explored, and actually provide an interesting subject of research. Moreover, the historical paradoxes are not all solved by the history interpretation. Some of them are, for instance the Wigner friend paradox, to the extent where no reference to observers is made in this interpretation. But some others remain unsolved, just with a reformulation in a different formalism and vocabulary.

Let us for instance take the Schrödinger cat paradox, which initially arose from the absence of any ingredient in the Schrödinger equation able to create the emergence of a single macroscopic result – in other words, for excluding impossible macroscopic superpositions of an isolated, non-observed, system. In the history interpretation, the paradox transposes in terms of choice of families of histories: the problem is that there is no way to eliminate the families of histories where the cat is at the same time dead and alive; actually, most families that are mathematically acceptable through the consistency condition contain projectors on macroscopic

superpositions, and nevertheless have exactly the same status as the families that do not. One would much prefer to have a “super-consistency” rule that would eliminate these superpositions; this would really solve the problem, but such a rule does not exist for the moment. At this stage, one can then do two things: either consider that the choice of sensible histories and reasonable points of view is a matter of common sense – a case in which one returns to the usual situation in the traditional interpretation, where the application of the postulate of wave packet is also left to the good taste of the physicist – or invoke decoherence and coupling to the external world in order to eliminate all these unwanted families – a case in which one returns to the usual situation where, conceptually, it is impossible to ascribe reasonable physical properties to a closed system without referring to the external world and interactions with it⁵⁰.

Finally, one may notice that, in the history interpretation, there is no attempt to follow “in real time” the evolution of the physical system; one speaks only of histories that are seen as complete, “closed in time”, almost as histories of the past in a sense. Basic questions that were initially at the origin of the introduction of the state reduction postulate, such as “how to describe the physical reality of a spin that has already undergone a first measurement but not yet a second”, are not easily answered. In fact, the consistency condition of the whole history depends on the future choice of the observable that will be measured, which does not make the discussion simpler than in the traditional interpretation, maybe even more complicated since its very logical frame is now under discussion. What about a series of measurements which may be, or may not be, continued in the future, depending on a decision that has not yet been made? As for the EPR correlation experiments, they can be re-analyzed within the history interpretation formalism [518] (see also [187] for a discussion of the Hardy impossibilities and the notion of “consistent counterfactuality”); nevertheless, at a fundamental level, the EPR reasoning still has to be dismissed for exactly the same reason already invoked by Bohr long ago: it introduces the EPR notion of “elements of reality”, or counterfactual arguments, that are not more valid within the history interpretation than in the Copenhagen interpretation (see for instance §V of [518] or the first letter in [517]). We are then brought back to almost the same old debate, with no fundamentally new element. We have nevertheless already remarked that, like the correlation interpretation, the history interpretation may be supplemented by other ingredients, such as the Everett

⁵⁰ For instance, in the context of histories, one sometimes invokes the practical impossibility to build an apparatus that would distinguish between a macroscopic superposition and the orthogonal superposition; this would justify the elimination of the corresponding histories from those that should be used in the description of reality. Such an argument reintroduces the notion of measurement apparatus and observers in order to select histories, in contradiction with the initial motivations of this point of view – see Rosenfeld’s quotation in §2. Moreover, this immediately opens again the door to Wigner friend type paradoxes, etc.

interpretation⁵¹ or, at the other extreme, EPR or deterministic ingredients, a case in which the discussion would of course become different.

For a more detailed discussion of this interpretation, see the references given at the beginning of this section. For a discussion of the relation with decoherence, the notion of “preferred (pointer) bases”, and classical predictability, see [325]. For a critique of the decoherent history interpretation, see for instance [519], where it is argued that consistency conditions are not sufficient to predict the persistence of quasi-classicality, even at large scales in the universe; see also [520], which claims that these conditions are not sufficient either for a derivation of the validity of the Copenhagen interpretation in the future; but see also the reply to this critique by Griffiths in [516]. The history interpretation is related to the theory of continuous measurements in quantum mechanics [521] (§8.3.2) as well as to the technique of stochastic quantum theories (§10.8.3). Finally, another interesting reference is an article published in 1998 in *Physics Today* [18], which contains a discussion of the history interpretation in terms that stimulated interesting reactions from the proponents of the interpretation [517].

10.11 Everett interpretation

A now famous point of view is that proposed in 1957 by Everett [522], who named it “relative state interpretation”. Other names also appear in the literature, such as “many-worlds interpretation” (MWI), “many-minds interpretation”, “universal wave function interpretation”, or “splitting universe interpretation” (the word “splitting”, or “branching”, refers here to the ramifications of the state vector of the universe into various branches, which we discuss below). These names correspond to different members of the same family of interpretations, which may differ slightly from each other, or even sometimes be significantly different. A common feature of all the family is that any possible contradiction between the two evolution postulates is removed by a simple but efficient method: the second postulate is merely suppressed.

10.11.1 No limit for the Schrödinger equation

In the Everett interpretation, the Schrödinger equation is taken even more seriously than in any other interpretation. One does not attempt to explain how, in a sequence of measurements, a well-defined result is obtained for each measurement. Instead, one merely considers that a single result never emerges: for each measurement,

⁵¹ Nevertheless, since the Everett interpretation completely suppresses from the beginning any specific notion of measurement, measuring apparatus, etc., the usefulness of completing it with the history interpretation is not obvious.

all possibilities are realized at the same time! The Von Neumann chain is then never broken; its tree of possibilities is left free to develop its branches *ad infinitum*. Everett writes [522] that, for a composite system of correlated sub-systems (observed system, measurement apparatus, and observer, after a measurement): “there does not exist anything like a single state for one sub-system. Systems do not possess states that are independent of the states of the remainder of the system . . . One can arbitrarily choose a state for one sub-system and be led to the relative state for the remainder” – this seems to be just a description of quantum entanglement, a well-known concept. But, now, the novelty is that the observer is considered as a purely physical system, to be treated within the theory on the same footing as the rest of the environment, microscopic or macroscopic. “As models for observers we can, if we wish, consider automatically functioning machines, possessing sensory apparatus and coupled to recording devices registering past sensory data and machine configuration”. Everett adds that “current sensory data, as well as machine configuration, is immediately recorded in the memory, so that all the actions of the machine at a given instant can be considered as functions of the memory contents only”; all relevant experience that the observer keeps from the past is then contained in this memory (magnetic tape, counter, even configurations of brain cells). From this Everett concludes that “there is no single state of the observer . . . with each succeeding observation (or interaction), the observer state branches into a number of different states . . . All branches exist simultaneously in the superposition after any sequence of observations”. He then checks that “experiences of the observer” (registered in the memory) “are in full accord with the predictions of the conventional ‘external observer’ formulation of quantum mechanics”. Much later, in a letter to D. Raub [523], he writes that this interpretation is “the only completely coherent approach to explaining both the contents of quantum mechanics and the appearance of the world”.

Consider for instance the physical system made of one microscopic system (or several systems) under study, a measurement apparatus and the observer, and then assume that this ensemble is isolated from the rest of the universe. Its state vector then represents reality itself, not our knowledge of this reality. While measurements are performed, the state vector ramifies into branches corresponding to all possible results of measurements, without ever selecting one of these branches; all remain real after measurement. The observer is part of this ramification process, and his mind becomes captured in an entangled state with the experimental apparatus and all the environment, reaching simultaneously different states corresponding to the registration of several different results (hence the words “many minds”). In other words, the observer plays the role of the Schrödinger cat in the historical paradox. Nevertheless, he cannot perceive several different results at the same time. Indeed, each “component of the observer” has no relation whatsoever with all the others, as

well as with the state vectors that are associated to them (hence the name “relative state interpretation”).

The emergence of macroscopic uniqueness is then just a delusion, a consequence of the way the observer perceives his own memory, of his abilities of introspection. This delusion appears as a consequence of the limitations of the human mind; since the universe itself evolves in a perfectly regular and deterministic way, randomness occurs only in the mind of the observer as a consequence of the specific properties of his memory (storage into the memory and reading its content). Independently of the observers, nothing puts a limit on the predictive deterministic power of the Schrödinger equation. In reality, the random process we usually call “quantum measurement” never takes place!

The “universal wave function” is the wave function of the universe, which contains many branches, in particular all those created by quantum experiments leading to several possible results. Nevertheless, a split of one branch into several others may also occur in interaction processes that were not designed specifically as measurements⁵². Different branches corresponding to macroscopically different situations, involving for instance the position of the pointers and their consequences, are independent. The reason is that no interaction Hamiltonian has matrix elements coupling states where a macroscopic number of elementary quantum systems are in orthogonal states; they cannot be made to interfere in practice, since this would require acting coherently on the individual microscopic states of too many particles. This is why one sometimes considers these branches as different “worlds” existing in parallel, which is the origin of the acronym MWI; others prefer to consider that one still has a unique world in a quantum superposition of very different states, sometimes called “multiverse” [524]. In any case, the universe has a single quantum state but its sub-systems, including the observers, are simultaneously in many states.

10.11.2 Consistency of the interpretation

In the Copenhagen interpretation, the consistency of quantum mechanics requires the existence of a classical world where external observers make and record observations; in this interpretation, observers are neither external nor classical. Of course, the postulate of state vector reduction disappears as well. Moreover, the Born probability rule is then seen, not as an independent postulate, but as a rule that should be

⁵² When Wigner’s friend (§2.3) communicates the result he has obtained to the physicist outside of the laboratory, for instance by giving him a telephone call, more branching takes place; the physicist outside becomes entangled with the branches resulting from the experiment, while he was not before he learnt the result. There is a relation between branching and the flow of information.

derived from the superposition principle. This means that one has to explain, within the formalism of the space of states and of the linear Schrödinger equation, why:

- (i) as we have already remarked, every observer has the impression that a single result has been obtained when he uses his memory to recall the experiments that were made in the past. Different observers of the same experiment agree when they compare the obtained results.
- (ii) why, when repeating the same experiment, the observer also has the impression that the frequency of occurrence of each possible result corresponds to the usual Born rule.
- (iii) it is also necessary to define precisely the appropriate basis of the space of states in which the preceding properties are true, and the conditions under which the branching (or splitting) of the observers take place.

Point (i) is a postulate related to the faculties of registering results and of introspection and the human mind. Even if they belong to a perfectly deterministic universe, when they perform experiments, observers have stochastic perceptions of the results they register in their memories. An essential component of the interpretation is that, in each state of his mind, the observer can have access only to the content of his memory in the corresponding branch of the state vector, without any influence of all the other components. Similarly, when exchanging information on results of experiments with other observers, he or she has access to only one branch associated with one single well-defined content of their memories. In other words, any group of observers is split into many completely independent components, without any possible communication between them.

Point (ii) requires that, without invoking the notion of external observations, probabilities should emerge from considerations on the mathematical structure of the entangled state vector, and that their value should be predicted [525, 526]. The Schrödinger equation is linear, and of itself does not imply any particular meaning to the square of the norm of its various components, which must be related to the frequencies of occurrence of the various results. When the same experiment is repeated many times, many branches appear in the state vector, corresponding to all possible series of results; the idea is to show that, in most cases, the observer is part of a branch where the relative frequency of the results corresponds to the Born rule⁵³.

⁵³ Consider the simplest case, an experiment providing two results, with probabilities p and $q = 1 - p$ according to the Born rule. When the experiment is repeated many times, the state vector splits into two components each time, and therefore develops a very large number of branches. One can then show that most of its norm goes to the components where the ratio between the two kinds of splits is very close to p/q , while the norm of the other components is negligible. This argument has sometimes been used as a derivation of the Born rule, but it is actually circular, since it assumes from the beginning that the weight of the various branches is given by their norm.

One may also consider situations where the nature of the experiment changes during the sequence, or for instance where several observers repeat different experiments in parallel.

Everett argues that a natural way to attribute measures to the various “trajectories” of the observers is to choose the usual Lebesgue norm, since it conserves the total probability. He then evaluates the measure associated with any particular sequence of results stored in a memory⁵⁴. More recent versions of the Everett interpretation are formulated in terms of information and flow of information (§10.1.3). For instance, Deutsch [525] has proposed an information-theoretic derivation of the Born rule by combining the Everett interpretation with game theory; see also the work of Wallace [527] and Saunders [528]. Zurek [326, 529] has discussed the relations between the Born probabilities and “environment-assisted invariance”, or “envariance”, which he compares to Laplace’s standard definition of probabilities based on the “principle of indifference”. Enviance can also be used to derive the dynamical independence of the branches of the state vector. Nevertheless, whether or not one can derive the Born rule from the Everett theory in a way that remains more economical than the rule itself remains somewhat controversial. For instance Peres writes (§12.1 of [287]), speaking of this family of interpretations: “None is satisfactory because they merely replace the arbitrariness of the collapse postulate by that of the no-communication hypothesis” (no-communication between branches of the state vector); see also [530] and [531].

In any case, the Everett notion of probability remains very different from any usual notion of probability, including that used within standard quantum mechanics. Normally probability relates to a situation of uncertainty where a set of possibilities exist, while only some will become actual. This is not the case in Everettian theory, where every outcome of a measurement actually occurs, so that there is no distinction between what actually happens and what does not. In reality, there is no uncertainty at all. Moreover, in the usual notion of probability, the observer samples an ensemble of possibilities among those of a statistical ensemble; in the Everett interpretation, the observer contains all possibilities in the various branches and therefore plays himself, in a sense, the role of the statistical ensemble. A new concept of probability is then introduced, see for instances the discussions of Refs. [532] and [533]. Curiously, instead of disappearing from the scene as one could have expected⁵⁵, the observer plays a central role in the theory again, by generating probabilities.

⁵⁴ There is an interesting relation between this work and Gleason’s theorem (§10.4.4), which was discovered independently and published almost simultaneously.

⁵⁵ See for instance Rosenfeld’s quotation in §2.5.

Point (iii) is also very important. In Everett interpretation, macroscopic uniqueness is not a property of the outside world, but rather of the states of this world that become entangled with the relevant basis of states for the memory of the observers⁵⁶ – a basis in which they store permanent information about the past and they use for instance to exchange consistent information on the results of the experiments with others. Each observer is always macroscopic, having no direct access to the microscopic world: an appropriate macroscopic device is necessary to transmit the microscopic information to his memory. Therefore, an observer can never become directly entangled with a microscopic system, for instance with the two trajectories that a single particle can take in an interferometer (which would destroy the interference effect). Such an entanglement can only occur indirectly, through an apparatus acting as an environment of the observed system. It then becomes natural to invoke macroscopic decoherence as the source of branching of the observer [534]; we have seen (§§6.3.3 and 10.1.1.a) that decoherence prevents the observation of any interference effect between the various components of a coherent superposition, whether the observer is involved or not in this superposition. Generally speaking, the “pointer state basis” (§8.1.2.b) created by the environment naturally appears as the best choice for expanding the Everett entangled state and applying properties (i) and (ii). This is a common point of view, even if it creates a somewhat indirect relation: both the memory of the observer and the external environment are coupled to the pointer, but the idea is to assume that the nature of the second coupling determines the relevant basis for the first interaction and entanglement. In other words, the relevant basis of the memory is not directly determined by its coupling with the observed object (pointer), but by its coupling with a third partner (considering that the memory of the observer is the only effective environment of the pointer seems rather unrealistic). One should then assume that both couplings share common properties concerning the preferred basis.

Another point of view is to consider (as is often the case in the variant called “many-world interpretation”) that the splits of the state of the universe take place only during the processes of measurement; the Copenhagen notion of measurement is then conserved. In the “many-mind interpretation”, the split takes place at the level of the minds of the observers, a point of view that can be seen as an elaborate version of Wigner’s interpretation involving human consciousness (§§10.1.1.b and 2.3).

⁵⁶ Consider macroscopic sub-ensembles of the universe that have interacted with observers. Macroscopic uniqueness is not an intrinsic property that could be attributed to them; rather, it corresponds to the perception of a property arising from the type of correlation occurring between them and the memory registers of the observers. In this interpretation, the observers therefore play a role that is at least as important as in the other interpretations of quantum mechanics. In the parts of the universe that have never interacted with observers, macroscopic uniqueness has no physical meaning at all.

Since each branch of the state vector involves a different history of the world, it is also relatively natural to combine the Everett interpretation with the decoherent history interpretation (§10.10).

10.11.3 Discussion

How is an EPRB experiment seen in this theory? In the Bohmian interpretation, we emphasized the role of Stern–Gerlach analyzers; in the modified Schrödinger dynamics, the role of the detectors and decoherence; here we have to emphasize the correlations with the external world, which take place for the memories in the minds of the two human observers. The state vector actually develops its Von Neumann chain through the analyzers and the detectors and, at some point, also includes the observers whose brains become part of the superposition. For each choice of the settings a and b , four branches of the state vector coexist; each branch depends explicitly on both parameters, and it is not possible to write mathematically a state vector describing locally the properties associated with a limited region of space and depending on a only, or on b only. Since in the Everett interpretation the state vector directly describes reality, the very expression of an entangled state vector then leads to a non-local description of reality. As for the observers, Alice and Bob, they are themselves embedded in this non-local Von Neumann chain, with components that depend on both a and b ; therefore, the choice of a has a distant influence on what Bob’s registers, and conversely. Clearly, the Everett interpretation implies a non-local description of reality⁵⁷.

A question that naturally then arises is whether or not the notion of external parameters remains valid within the Everett interpretation (cf. discussion and Bell’s quotation in §4.5.1.c). The memories of the observers are considered as registers functioning according to the deterministic Schrödinger equation. Should one then consider that these observers retain free will, so that they can arbitrarily choose the settings a and b , or rather that these choices are causally determined, exactly as their memories? In the latter case, one arrives at superdeterminism, and the notion of free parameter disappears altogether; the proof of the Bell inequalities becomes impossible, since the settings a and b may be considered as consequences of common causes in the past; all the discussion of locality we have given becomes meaningless.

⁵⁷ Nevertheless, if one considers the flux of information taking place during the Schrödinger (or Heisenberg) evolution of a system of qubits reproducing an EPR experiment, one is led to equations that remain perfectly local [535]; the authors of this reference conclude that a violation of the Bell inequalities means that one cannot attribute stochastic variables to experimental results, which is similar to rejecting counterfactuality (or realism).

We have seen above that different branches of the state vector of the universe are independent. Then, by construction, the Everett interpretation leads to experimental predictions that are identical to those of the standard interpretation; in this sense, it is not falsifiable. DeWitt [16] even considers that this interpretation is a mere consequence of the formalism: after asking the question “Could the solution to the dilemma of indeterminism be a universe in which all possible outcomes of an experiment actually occur?” he states that “the mathematical formalism of the quantum theory is capable of yielding its own interpretation” – see also the interesting debate [19] that was stirred by the publication of this point of view. DeWitt also states that the Everett interpretation is a necessity⁵⁸ in quantum cosmology: “Everett’s [‘many worlds’] interpretation has been adopted by the author (Bruce DeWitt) out of practical necessity: he knows of no other. At least he knows of no other that imposes no artificial limitations or fuzzy metaphysics while remaining able to serve the varied needs of quantum cosmology, mesoscopic quantum physics, and the looming discipline of quantum computation” (page 144 of [536]).

Considering the wave function of the universe is of course natural when studying it as a whole. Generally speaking, the idea of multiple universes is not uncommon in astrophysics and cosmology. It has sometimes been evoked to explain the existence of dark matter and dark energy; in the context of the Everett interpretation, one could propose an explanation involving the existence of some interaction between the branches of the state vector of the universe. Multiple universes have also been conjectured in the context of the “anthropic principle” (assuming that the state of the universe contains many branches where the conditions are such that the appearance of intelligent beings is not possible; nevertheless, mankind can only observe what is happening in the small proportion of branches where this appearance has been possible, and therefore very special universes). For a review of parallel universes and many worlds in the context of cosmology, see the articles by Tegmark [537], who emphasizes that “parallel universes are not a theory, but a prediction of certain theories”. A large proportion of physicists specialized in quantum cosmology seem to prefer the Everett interpretation [538].

At first sight, the Everett interpretation may look like a welcome unification in quantum physics, especially since it requires no change in the formalism. This makes it attractive to many physicists in a first contact; but on further study some find it difficult to really assimilate, and the interpretation has been criticized. For instance, Bell writes [539] that: “the elimination of arbitrary and inessential elements from Everett’s theory leads back to, and throws new light on, the concepts of de Broglie”. Indeed, while the Everett interpretation is often considered as the

⁵⁸ Other interpretations, for instance those involving modified Schrödinger dynamics (§10.8), are nevertheless also perfectly suitable for introducing the state vector of the universe.

exact opposite of de Broglie–Bohm interpretation, often seen as more naïve, both are actually rather similar. The ensemble of the Bohmian empty waves, together with the non-empty wave, reconstructs the same state vector as in the Everett interpretation. Moreover, when an Everettian observer examines the content of his memory after performing a series of experiments, he has access to a single branch of the state vector selected along a sequence of ramifications. But this branch can also be defined by a retrodictive Bohmian trajectory in configuration space; during each measurement, the trajectory directly indicates which branch has been selected by the Bohmian position to “surf” on it. The only difference is that the sequence of state vectors gives only a blurred view of the trajectory, while it is defined with infinite accuracy in Bohmian theory. Nevertheless, this difference has no consequence in practice since, in both cases, the observer remembers the same sequence of events, and nothing more. In [47], Bell adds that: “this multiplication of universes . . . serves no real purpose in the theory, and can simply be dropped without repercussions⁵⁹”. Peres calls it a “bizarre theory” [287] and considers the interpretation as non-economical (quotation above in §10.11.2). Leggett discusses this interpretation in the following terms [81]: “The branches of the superposition which we are not conscious of are said to be ‘equally really’, though it is not clear . . . what these words, ostensibly English, are supposed to mean”. One question is what we should expect from a physical theory; does it have to explain how we perceive results of experiments, and if so of what nature should such an explanation be? Since the emphasis is put, not on the physical properties of the systems themselves, but on the effects that they produce on our minds and memories, notions such as perception ([522] speaks of “trajectory of the memory configuration”) and their description in neurosciences (the properties of introspection) become part of the theory. What is clear, anyway, is that the whole point of view is exactly opposite to that of the proponents of the additional variables. The Everett interpretation is attractive aesthetically but remains somewhat mind boggling. Since the human population of the earth is made of billions of individuals, and presumably since each of them is busy making quantum measurements all the time without even knowing it, should we see the physical reality of the universe as constantly branching at an exponentially fantastic rate?

10.12 Conclusion

Quantum mechanics is, with relativity, the essence of the big conceptual revolution of the physics of the twentieth century. The progress, in theoretical and

⁵⁹ If one cuts by hand all branches of the various universes but one, one reaches an interpretation that is very similar to the “pragmatic” interpretation discussed in §10.1.1.a.

experimental physics, has been extraordinary. New breakthroughs have constantly occurred, ranging from pure theory to practical applications; they have given access to situations and devices that would not even have been conceivable before the advent of quantum mechanics. Our present understanding of the universe is only possible with the use of quantum mechanics at almost every step.

Now, do we really understand this marvellous theory? We understand its machinery pretty well, and we know how to use its formalism to make predictions in an extremely large number of situations, even in cases that may be very intricate. Quantum mechanics has provided many counterintuitive predictions that, initially, seemed hard to believe, but turned out to be verified by experiments – even if sometimes the experimental verification came long after. One striking illustration is the observation of Bose–Einstein condensation in dilute gases, which was predicted by Einstein in 1925 [540], but observed only in 1995 [541, 542]. The phenomenon takes place at extremely low temperatures (much lower than that of the cosmic radiation background), so that, presumably, it had never occurred anywhere in the universe before 1995! It was just in limbo, waiting in the equations of quantum mechanics until it was realized one day. This illustrates the extraordinary predictive power of the theory. Heinrich Hertz, who played such a crucial role in the understanding of electromagnetic waves in the nineteenth century (Hertzian waves), remarked that, sometimes, the equations in physics are more intelligent than the person who invented them: “It is impossible to study this amazing theory without experiencing at some times the strange feeling that the mathematical formulas somehow have a proper life, that they are smarter than we, smarter than their author himself, so that we obtain from them more than what was originally put into them” [543]. The remark certainly applies to the equations of quantum mechanics and to the superposition principle: they contain probably much more substance than any of their inventors thought, for instance in terms of unexpected types of correlations, entanglement, etc. It is really astonishing to observe that, in all known cases, the equations have always predicted exactly the correct results.

But, conceptually, whether or not we really understand the theory is less clear, as illustrated by the variety of interpretations used by physicists. True, the Schrödinger equation reaches the initial aim of quantum mechanics perfectly well: it explains why atoms, molecules, and matter in general are stable, which is impossible within classical physics. But the same equation, with its enormous number of successful predictions, seems to go too far when it extends its linear superpositions into the macroscopic world. It then raises difficult questions about the uniqueness of our observations and the very existence of any classical reality on a human scale. Some interpretations (Everett for instance) consider that this uniqueness is not part

of reality, but just a delusion arising from the way our memories register events. Others propose to change the Schrödinger equation to remove these problematic macroscopic superpositions and turn the wave function into a real wave, propagating in configuration space. Still other interpretations (Bohm for instance) propose to introduce additional variables to solve the problem of macroscopic uniqueness. In this context, a major issue is whether or not the present form of the theory of quantum mechanics is complete. If it is, it will never be possible in the future to give a more precise description of the physical properties of a single particle than its wave function; this is the position of the proponents of the orthodox interpretation. If it is not, future generations may be able to do better and introduce some kind of description that is more accurate.

We have seen why the EPR argument is similar to Gregor Mendel's reasoning, which led him from observations performed between 1854 and 1863 to the discovery of specific factors, the genes (the word appeared only later, in 1909). These genes turned out to be associated with microscopic objects hidden inside the plants that he studied. In such cases, one infers the existence of microscopic "elements of reality" from the results of macroscopic observations. Mendel could derive rules obeyed by the genes, when they combine in a new generation of plants, but at his time it was totally impossible to have any precise idea of what they really could be at a microscopic level (or actually even if they were microscopic objects, or macroscopic but too small to be seen with the techniques available at that time). It took almost a century before Avery *et al.* (1944) showed that the objects in question were contained in DNA molecules; later (1953), Franklin, Crick and Watson discovered how subtle the microscopic structure of the object actually was, since genes corresponded to subtle arrangements of nucleic bases hidden inside the double helix of DNA molecules. We now know that, in a sense, rather than simple microscopic objects, the genes are arrangements of objects, and that all the biological machinery that reads them is certainly far beyond anything that could be conceived at Mendel's time. Similarly, if quantum mechanics is one day supplemented with additional variables, these variables will probably not be some trivial extension of the other variables that we already have in physics, but variables of a very different nature. But, of course, this is only a possibility, since the histories of biology and physics are not necessarily parallel!

A natural comparison is with special relativity, since neither quantum mechanics nor relativity is intuitive; indeed, experience shows that both, initially, require a lot of thought from each of us before they become intellectually acceptable. But the similarity stops here: while it is true that the more one thinks about relativity, the more understandable it becomes (at some point, one may even get the feeling that relativity is actually a logical necessity!), one can hardly make the same statement

about quantum mechanics. Nevertheless, among all intellectual constructions of the human mind, quantum mechanics may be the most successful of all theories, since, despite all efforts of physicists to find its limits of validity (as they do for all physical theories), and many sorts of speculations, no one for now has been able to obtain clear evidence that these limits even exist. Future will tell us if this is the case; surprises are always possible!

11

Annex: Basic mathematical tools of quantum mechanics

This chapter gives a summary of the mathematical formalism used in quantum mechanics, with a short bibliography put directly at its end. It should be seen as a complement, to be used by readers who wish to know more than what has been recalled about the mathematical tools in the other chapters. Some results are given without explicit proofs; they can be found for instance in Chapters II and IV of [1]. Many quantum mechanics textbooks introduce its general formalism in a more complete way, for instance Chapter VII of [2], Chapter 3 of [3], or Chapter 2 of [4].

We first summarize the general formalism of quantum mechanics for any physical system (§11.1), with the Dirac notation; we then study how this formalism treats the grouping of several physical systems into one single quantum system (§11.2); finally, we study a few simple special cases (§11.3), for instance a single particle in an external potential, with or without spin; the reader who prefers wave functions to more abstract reasonings in spaces of states may begin with this section.

11.1 General physical system

The general formalism of quantum mechanics applies to all physical systems, whether they contain a single particle, many particles of various sorts, one or several fields, etc.

11.1.1 *Quantum space of states*

The physical state of a system at each time is defined in quantum mechanics by a state vector which, in Dirac notation, is written $|\Psi\rangle$ – or $|\Psi(t)\rangle$ if one wishes to make the time dependence explicit. This vector $|\Psi\rangle$ belongs to a complex state vector space $\mathcal{E}$, which may have a more or less complicated structure depending on the system considered. In Dirac notation, the vectors of $\mathcal{E}$ are often called “kets”. We

explain below (§11.1.3) how the physical properties of the system can be calculated from $|\Psi\rangle$. Space $\mathcal{E}$ is called either “space of states”, or “Hilbert space” for historical reasons.

Consider any pair of vectors $|\Psi_1\rangle, |\Psi_2\rangle, \dots$ belonging to $\mathcal{E}$; by definition of a complex vector space, any linear combination with complex coefficients α, β :

$$\alpha |\Psi_1\rangle + \beta |\Psi_2\rangle + \dots \quad (11.1)$$

is another vector $|\Psi\rangle$ belonging to $\mathcal{E}$. The “superposition principle” states that any linear combination of state vectors provides another possible state vector for the considered physical system; this principle has many physical consequences, several of which are discussed in this book.

In this space, one defines the scalar product of state vectors; the scalar product of vector $|\Psi\rangle$ by vector $|\Phi\rangle$ is a complex number that is usually written:

$$\langle \Phi | \Psi \rangle \quad (11.2)$$

This number depends linearly on $|\Psi\rangle$ (it is multiplied by λ if $|\Psi\rangle$ is multiplied by λ) and anti-linearly on $|\Phi\rangle$ (it is multiplied by the complex conjugate μ^* of μ if $|\Phi\rangle$ is multiplied by μ). In Dirac notation, the mathematical object $\langle \Phi |$ is often called “bra”.

A basis $\{|u_i\rangle\}$ of $\mathcal{E}$ is an ensemble of vectors $|u_i\rangle$ such that any vector of $\mathcal{E}$ can be written as a linear combination:

$$|\Psi\rangle = \sum_i x_i |u_i\rangle \quad (11.3)$$

in a unique way; the x_i are called the “components” of $|\Psi\rangle$ in this basis.

If the dimension of state space is finite and equal to P , the number of vectors $|u_i\rangle$ (and therefore the number of terms in this sum) is equal to P . Basis $\{|u_i\rangle\}$ is orthonormal if the scalar products between its vectors obey the relations:

$$\langle u_i | u_j \rangle = \delta_{i,j} \quad (11.4)$$

where $\delta_{i,j}$ is the Kronecker delta. With an orthonormal basis, one can show that the components x_i appearing in (11.3) are given by:

$$x_i = \langle u_i | \Psi \rangle \quad (11.5)$$

so that:

$$|\Psi\rangle = \sum_{i=1}^P \langle u_i | \Psi \rangle |u_i\rangle \quad (11.6)$$

We also have:

$$\langle \Psi | \Psi \rangle = \sum_i^P |x_i|^2 \quad \langle \Phi | \Psi \rangle = \sum_i^P y_i^* x_i \quad (11.7)$$

where the y_i are the components of ket $|\Phi\rangle$.

If the dimension of the state space is infinite, only normalizable kets correspond to a physical state, namely kets for which:

$$\sum_i |x_i|^2 = \text{finite number} \quad (11.8)$$

Some “bases” are not only infinite, but also continuous¹, which means that the sums of (11.3), (11.6), and (11.7) are all replaced by integrals; it is also necessary to replace the Kronecker delta in (11.4) by a Dirac delta distribution.

11.1.2 Operators

One also defines the action of linear operators in the state space. An operator A is linear when the action of A on vector $\alpha|\Psi\rangle + \beta|\Phi\rangle$, where α and β are complex constants, is given by:

$$A[\alpha|\Psi\rangle + \beta|\Phi\rangle] = \alpha A|\Psi\rangle + \beta A|\Phi\rangle \quad (11.9)$$

In any finite basis, operator A may be written as a $P \times P$ matrix with matrix elements $A_{i,j}$; the columns of the matrix contain the components of the transforms of the basis vectors under the action of A . In an orthonormal basis $\{|u_i\rangle\}$, the matrix elements are given by the scalar product of $A|u_j\rangle$ by $|u_i\rangle$:

$$A_{i,j} = \langle u_i | A | u_j \rangle \quad (11.10)$$

One can easily define the product of an operator by any complex constant α , the sum of operators, or more generally the linear combination $\alpha A + \beta B$ of two operators A and B by:

$$[\alpha A + \beta B]|\Psi\rangle = \alpha(A|\Psi\rangle) + \beta(B|\Psi\rangle) \quad (11.11)$$

11.1.2.a Product, commutator, eigenvectors

The product of two operators A and B is defined by its action on any ket $|\Psi\rangle$:

$$AB|\Psi\rangle = A|\Psi'\rangle \quad \text{where:} \quad |\Psi'\rangle = B|\Psi\rangle \quad (11.12)$$

¹ The vectors of continuous basis are not normalizable, and therefore do not really belong to $\mathcal{E}$, which explains our quotation marks; see for instance the discussion in §A-3 of [11-1].

This product AB corresponds to the action of the B operator, followed by that of A . The matrix corresponding to the product operator $C = AB$ is merely given by the product of the matrices associated with A and B , with the usual rule for multiplying matrices (multiplying the lines of the left matrix by the columns of the right matrix). In general, the product operator AB is different from the product operator BA in the reverse order; the operator $[A, B]$ defined by:

$$[A, B] = AB - BA \quad (11.13)$$

is called the “commutator” between operators A and B . If this commutator vanishes, one says that operators A and B commute.

The eigenvectors $|a_k\rangle$ of A associated to the eigenvalues a_k are defined as vectors for which the action of A is simply a multiplication by a_k :

$$A|a_k\rangle = a_k|a_k\rangle \quad (11.14)$$

A method to obtain the eigenvalues and the eigenvectors of A is to proceed to a diagonalization of the matrix of coefficients $A_{i,j}$; nevertheless, for some matrices, the complete diagonalization is not possible. If an eigenvalue corresponds to several eigenvectors that are not proportional, one says that the eigenvalue is “degenerate”.

11.1.2.b Hermitian and unitary operators

- (i) To each operator A one associates an Hermitian conjugate operator $A^\dagger$ with matrix elements obtained by complex conjugation of those of A , followed by a transposition of lines and columns:

$$\langle u_i | A^\dagger | u_j \rangle = \langle u_j | A | u_i \rangle^* \quad (11.15)$$

An operator A is Hermitian if it is equal to its own Hermitian conjugate: $A = A^\dagger$; this means that all its matrix elements obey the relations²:

$$A_{i,j} = [A_{j,i}]^* \quad (11.16)$$

In other words, matrix elements in symmetrical positions with respect to the first diagonal of the matrix are complex conjugate. One can show that any Hermitian operator can always be fully diagonalized, and that all its eigenvalues are real (which does not mean that they are all distinct: some may be degenerate).

² One can show that, if this relation is satisfied in one basis $\{|u_i\rangle\}$, it is also satisfied in any other orthonormal basis.

- (ii) A linear operator U is unitary if the product of its Hermitian conjugate by U is the unity operator:

$$U^\dagger U = 1 \quad (11.17)$$

Consider two kets $|\Phi\rangle$ and $|\Psi\rangle$ in the space of states and the action of U on these kets:

$$|\Phi'\rangle = U|\Phi\rangle \quad \text{and} \quad |\Psi'\rangle = U|\Psi\rangle \quad (11.18)$$

Relation (11.17) shows that:

$$\langle\Phi'|\Psi'\rangle = \langle\Phi|U^\dagger U|\Psi\rangle = \langle\Phi|\Psi\rangle \quad (11.19)$$

This result expresses that the action of a unitary operator does not change the scalar product (as a consequence, it does not change the norm of a ket either). A unitary operator transforms any orthonormal basis into another orthonormal basis of the space of states. Conversely, any orthonormal change of basis in the space of states defines a unitary operator. This property can be used as an alternative definition of unitary operators.

The matrix associated to a unitary operator is called a unitary matrix. Its columns contain the components of the kets $U|u_i\rangle$ in the initial basis $\{|u_i\rangle\}$. Unitarity can easily be checked by selecting any pair of columns i and j , taking the complex conjugate of the first column, and calculating the line-by-line product; the result should be the Kronecker delta $\delta_{i,j}$ for any choice of i and j .

A unitary operator can always be diagonalized in an orthogonal basis, as a Hermitian operator. The eigenvalues are all complex numbers of the form e^{ia} , where a is a real number. A consequence is that U can always be written in the form $U = e^{iA}$, where A is a Hermitian operator; unitary operators are, so to say, complex exponentials of Hermitian operators.

It is easy to show from (11.17) that the product of any number of unitary operators is another unitary operator; unitary operators can be chained without losing unitarity. Nevertheless, two unitary operators U_1 and U_2 do not commute in general.

11.1.2.c Trace

The trace of an operator is a number, the sum of all its diagonal matrix elements:

$$\text{Tr}\{A\} = \sum_i A_{ii} \quad (11.20)$$

One can show that the value of the trace is independent of the basis used to calculate it. Moreover, circular permutations of the operators inside a trace do not change the value of the trace; for any three operators A , B , and C :

$$\text{Tr}\{ABC\} = \text{Tr}\{BCA\} = \text{Tr}\{CAB\} \quad (11.21)$$

11.1.3 Probabilities

Hermitian operators play an important role in quantum mechanics: any physical quantity that can be measured on the system (such as its energy, momentum, angular momentum, etc.) corresponds to a Hermitian operator acting in its state space $\mathcal{E}$.

We have seen that, when A is Hermitian, it can be “diagonalized”. One can then find a basis made of vectors of $\mathcal{E}$ that are all eigenvectors $|a_k\rangle$ of A , associated with real eigenvalues a_k ; in this basis, the operator is represented by a diagonal matrix (having all zero matrix elements except the diagonal elements, which are the a_k).

A fundamental postulate of quantum mechanics is the Born probability rule. It is also discussed in the main text (§1.2.1.c) but, for completeness, we briefly recall it here. The rule says that, when the physical quantity associated to operator A is measured on a system in a normalized state $|\Psi\rangle$:

- (i) the only possible results are the eigenvalues a_k ,
- (ii) the probability $\mathcal{P}(a_k)$ to find a particular result a_k is given by the squared modulus of the scalar product of $|\Psi\rangle$ by the corresponding normalized eigenvector:

$$\mathcal{P}(a_k) = |\langle a_k | \Psi \rangle|^2 \quad (11.22)$$

We have assumed that a single eigenvector corresponds to eigenvalue a_k (non-degenerate eigenvalue); if several distinct eigenvectors correspond to this eigenvalue, one must add the probabilities (11.22) corresponding to all orthonormalized eigenvectors associated with a_k , as in equation (1.4).

Assume we have an ensemble of systems that have been all prepared in the same state $|\Psi\rangle$. If one performs a large number of measurements of the physical quantity associated with operator A , since the Born rule is probabilistic, one generally obtains different values at each measurement. Nevertheless, one can calculate an average value $\langle A \rangle$ and obtain from this rule:

$$\langle A \rangle = \sum_k a_k \mathcal{P}(a_k) = \langle \Psi | A | \Psi \rangle \quad (11.23)$$

Consider the operator:

$$P_\Phi = |\Phi\rangle \langle \Phi| \quad (11.24)$$

where $|\Phi\rangle$ is normalized ($\langle\Phi|\Phi\rangle = 1$). Its action on an arbitrary ket $|\Psi\rangle$ is:

$$P_\Phi |\Psi\rangle = |\Phi\rangle \langle\Phi|\Psi\rangle = c |\Phi\rangle \quad (11.25)$$

where $c = \langle\Phi|\Psi\rangle$ is a number. Any ket is transformed under the action of P_Φ into a ket that is proportional to $|\Phi\rangle$; this indicates that P_Φ is the projector onto state $|\Phi\rangle$, obeying the characteristic projector relation $P^2 = P$. We note $P(a_k)$ the projector onto eigenvector $|a_k\rangle$:

$$P(a_k) = |a_k\rangle \langle a_k| \quad (11.26)$$

if a_k is non-degenerate; if it is degenerate, $P(a_k)$ is defined as the sum of projectors over an orthonormalized ensemble of eigenvectors associated with this same eigenvalue. Applying the Born probability rule shows that the probability of obtaining result a_k is then given by:

$$\mathcal{P}(a_k) = \langle\Psi| P(a_k) |\Psi\rangle = \langle P(a_k) \rangle \quad (11.27)$$

11.1.4 Time evolution

Until now we have considered the properties of the physical system at one time. When time evolution is taken into account, the state vector becomes a time-dependent ket $|\Psi(t)\rangle$, with an evolution given by the Schrödinger equation:

$$i\hbar \frac{d}{dt} |\Psi(t)\rangle = H(t) |\Psi(t)\rangle \quad (11.28)$$

where $H(t)$ is the Hamiltonian operator at time t (operator associated with the energy of the system) and $\hbar$ the Planck constant h divided by 2π . Since (11.28) is a first-order time differential equation, it provides a continuous time evolution of the state vector from any initial value. The Hamiltonian operator $H(t)$ is Hermitian but may take various forms, sometimes very complicated, so that it is not necessarily possible to solve equation (11.28) exactly; nevertheless, the equation is always valid.

The Schrödinger evolution does not change the norm of the state vector:

$$i\hbar \frac{d}{dt} \langle\Psi(t)|\Psi(t)\rangle = \langle\Psi(t)| [H(t) - H^\dagger(t)] |\Psi(t)\rangle = 0 \quad (11.29)$$

which means that the solution $|\Psi(t)\rangle$ of the equation at time t has the same norm as the initial state $|\Psi(t_0)\rangle$. The same reasoning immediately shows that the scalar product $\langle\Psi_1(t)|\Psi_2(t)\rangle$ of two different solutions of the Schrödinger equation does not change in time, whatever the initial values of these solutions were at time t_0 .

In other words, time propagation from t_0 to t corresponds to a unitary evolution (§11.1.2.b), which defines a unitary operator $U(t, t_0)$:

$$|\Psi(t)\rangle = U(t, t_0) |\Psi(t_0)\rangle \quad (11.30)$$

The evolution operator $U(t, t_0)$ between time t_0 and t obeys the chain relation corresponding to an evolution between times t and t'' , followed by evolution between times t'' and t' :

$$U(t', t) = U(t', t'') U(t'', t) \quad (11.31)$$

11.1.5 Density operator

Expressions such as (11.22), (11.23), and (11.27) are not linear with respect to the state vector $|\Psi\rangle$. One can nevertheless obtain linear expressions as functions of an operator ρ called “density operator”, which replaces $|\Psi\rangle$.

11.1.5.a Definition

The density operator associated with a normalized state vector $|\Psi\rangle$ is the projector over this state, defined by:

$$\rho = |\Psi\rangle \langle\Psi| \quad (11.32)$$

The trace of this operator is one³:

$$Tr\{\rho\} = 1 \quad (11.33)$$

It is a Hermitian operator:

$$\rho = \rho^\dagger \quad (11.34)$$

One can then replace expressions (11.22) and (11.23) by:

$$\mathcal{P}(a_k) = Tr\{P(a_k) \rho\} \quad (11.35)$$

and:

$$\langle A \rangle = Tr\{A\rho\} \quad (11.36)$$

A useful property of these expressions is their linearity, which is convenient to combine the notions of classical and quantum probabilities. Assume that the state

³ This is a consequence of the normalization of $|\Psi\rangle$ since the calculation of the trace in any orthonormal basis $\{|u_i\rangle\}$ gives $Tr\{|\Psi\rangle \langle\Psi|\} = \sum_i \langle u_i | \Psi \rangle \langle\Psi | u_i \rangle$, which is equal to $\langle\Psi | \Psi \rangle = 1$.

$|\Psi\rangle$ of a physical system is not known, but that the system is in normalized state $|\Psi_1\rangle$ with probability p_1 , in normalized state $|\Psi_2\rangle$ with probability p_2 , ... in normalized state $|\Psi_n\rangle$ with probability p_n . We define the density operator ρ as the sum of the density operators of these states with a weight equal to their probabilities:

$$\rho = \sum_n p_n |\Psi_n\rangle \langle \Psi_n| \quad (11.37)$$

where:

$$0 \leq p_n \leq 1 \quad \text{and} \quad \sum_n p_n = 1 \quad (11.38)$$

Equations (11.33) to (11.36) then remain valid. For instance, the trace (11.33), obtained in any orthonormal basis $\{|u_i\rangle\}$, can be derived from the normalization of the $|\Psi_n\rangle$ and p_n :

$$\text{Tr}\{\rho\} = \sum_i \sum_n p_n |\langle \Psi_n | u_i \rangle|^2 = \sum_n p_n = 1 \quad (11.39)$$

11.1.5.b Pure states and statistical mixtures

The density operator is a Hermitian, positive (or, more precisely, non-negative) operator; every diagonal element of ρ obeys:

$$0 \leq \langle \Phi | \rho | \Phi \rangle \leq 1 \quad (11.40)$$

for any normalized ket $|\Phi\rangle$. To check this property, it is sufficient to replace ρ by its definition (11.37), to use the Schwarz inequality, and finally to take (11.38) into account. The density operator can always be diagonalized and, if we write $|\theta_m\rangle$ its eigenvectors associated with eigenvalues q_m , we can write ρ in the form:

$$\rho = \sum_m q_m |\theta_m\rangle \langle \theta_m| \quad (11.41)$$

Since $\langle \theta_m | \rho | \theta_m \rangle = q_m$ is a real number between 0 and 1 – relation (11.40) – we have:

$$0 \leq q_m \leq 1 \quad (11.42)$$

Every eigenvalue q_m of ρ is therefore bounded between 0 and 1, and their sum gives the trace of ρ , equal to 1:

$$\sum_m q_m = \sum_m \langle \theta_m | \rho | \theta_m \rangle = \sum_{m,n} p_n |\langle \Psi_n | \theta_m \rangle|^2 = \sum_n p_n = 1 \quad (11.43)$$

We can therefore interpret the eigenvalues of ρ as the occupation probabilities of states $|\theta_m\rangle$; we find a form that is similar to (11.37), but now with projectors onto eigenstates that are necessarily orthogonal. If we square (11.41), we obtain:

$$\rho^2 = \sum_m (q_m)^2 |\theta_m\rangle \langle \theta_m| \quad (11.44)$$

One then distinguishes two cases for the density operator:

Pure state

If only one of its eigenvalues q_m is equal to 1 (all the others being equal to zero), a single term plays a role in the m summation of (11.41); ρ is then the projector onto a single quantum states, its first eigenstate, which we can rename $|\Psi\rangle$. We are then in the case where the system is described by a “pure state” $|\Psi\rangle$, as in (11.32); this description corresponds to the maximal amount of information that can be given on a physical system within quantum mechanics. With (11.44), and since $1^2 = 1$ and $0^2 = 0$, we check that $\rho^2 = \rho$; in terms of density operators, we can characterize a pure state, either by this equality between ρ and ρ^2 , or by relation:

$$\text{Tr} \left\{ \rho^2 \right\} = \text{Tr} \{ \rho \} = 1 \quad (11.45)$$

Statistical mixture

If several eigenvalues q_m are non-zero, it is no longer possible to assign a single state $|\Psi\rangle$ to the system; several states are necessary, with weights given by probabilities that are neither 0, nor 1. One then says that the density operator ρ describes a “statistical mixture”. In contrast with the preceding case, the quantum description provided by ρ is not the most accurate possible within quantum mechanics. In fact, the description may even be very inaccurate, if many states $|\theta_m\rangle$ are associated with equal (or similar) probabilities q_m . Because, for any number q between 0 and 1, we have $q^2 < q$, we see in (11.44) that the density operator is not equal to its square ($\rho^2 \neq \rho$). Another way to characterize a statistical mixture is the relation:

$$\text{Tr} \left\{ \rho^2 \right\} < \text{Tr} \{ \rho \} = 1 \quad (11.46)$$

11.1.5.c Time evolution

When the $|\Psi_n\rangle$ in (11.37) become functions of time $|\Psi_n(t)\rangle$ evolving according to the Schrödinger equation, the density operator ρ becomes time dependent (the kets change in time but the probabilities p_n , which define the initial random choice of the system, remain constant). By inserting (11.28) as well as the associated time evolution for the bra $\langle \Psi_n(t)|$ into (11.37), one obtains the equation of evolution of

the density operator:

$$i\hbar \frac{d\rho(t)}{dt} = [H(t), \rho(t)] \quad (11.47)$$

where $[H(t), \rho(t)]$ is the commutator of $H(t)$ and $\rho(t)$. This equation is often called the “Von Neumann equation”.

The same reasoning applies if one starts from expansion (11.41) of ρ onto its eigenvectors; one then notices that the eigenvalues of ρ remain constant during time evolution, as well as the trace and the trace of ρ^2 .

11.1.5.d Statistical entropy

The statistical entropy (or Von Neumann entropy) S associated with any density operator ρ is defined by:

$$S = -k_B \text{Tr} \{ \rho \ln \rho \} \quad (11.48)$$

where k_B is the Boltzmann constant of statistical mechanics and thermodynamics, and where $\ln$ denotes the natural logarithm (in quantum information, one often prefers to take $k_B = 1$ and the base 2 logarithm, but this changes nothing essential). Equation (11.41) provides:

$$S = -k_B \sum_m q_m \ln q_m \quad (11.49)$$

For a density operator describing a pure state, all eigenvalues are 0, except one that is equal to 1; one then has $S = 0$. For a statistical mixture, relations (11.42) imply⁴ that $S > 0$; for instance, if the system is with equal probabilities in two orthogonal states, two eigenvalues of ρ are equal to 1/2 and $S = 2\ln 2$ (or $S = 2$ if the base 2 logarithm has been chosen). The value of S characterizes the distance between the quantum description provided by ρ and the optimal description with a pure state, the most accurate possible within quantum mechanics.

We have seen that the probabilities q_m remain constant during the time evolution of ρ according to (11.47); the Hamiltonian evolution of a density operator therefore conserves its entropy: $dS/dt = 0$.

11.1.6 Simple case, spin 1/2

The simplest space of states has two dimensions (except the trivial case with one dimension, which is of little interest since the system can never change state). This

⁴ Function $-x \ln x$ is positive when $0 < x < 1$.

case occurs in the study of the properties of a spin 1/2 particle with a single orbital wave function (§11.3.2); one then usually denotes $|+\rangle$ and $|-\rangle$ the two eigenstates of the O_z component of the spin, which provide a basis. The most general state then reads:

$$|\Psi\rangle = \alpha|+\rangle + \beta|-\rangle \quad (11.50)$$

where α and β are any two complex numbers, obeying condition $|\alpha|^2 + |\beta|^2 = 1$ if $|\Psi\rangle$ is normalized; this ket is associated with the column vector:

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} \quad (11.51)$$

The three components $S_{x,y,z}$ of the spin along the three directions Ox , Oy , and Oz correspond to three Hermitian operators; their action onto the column vector (11.51) is given by the three “Pauli matrices” σ_x , σ_y , and σ_z defined by⁵:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (11.52)$$

A simple calculation shows that the squares of the Pauli matrices are all the unit matrix:

$$[\sigma_x]^2 = [\sigma_y]^2 = [\sigma_z]^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (11.53)$$

The commutation relations of these Pauli matrices are:

$$[\sigma_x, \sigma_y] \equiv \sigma_x \sigma_y - \sigma_y \sigma_x = 2i\sigma_z \quad (11.54)$$

but the Pauli matrices anti-commute:

$$[\sigma_x, \sigma_y]_+ \equiv \sigma_x \sigma_y + \sigma_y \sigma_x = 0 \quad (11.55)$$

(in both cases, two other relations may be obtained by circular permutations of σ_x , σ_y , and σ_z).

The density operator ρ associated with state $|\Psi\rangle$ has the following matrix:

$$\rho = \begin{pmatrix} \alpha^* \alpha & \beta^* \alpha \\ \alpha^* \beta & \beta^* \beta \end{pmatrix} \quad (11.56)$$

⁵ More precisely, the three components of the spin correspond to the three Pauli matrices multiplied by constant $\hbar/2$; this is why we use the word “associated” and not “equal”.

which therefore corresponds to a pure state (one can easily check that $\rho^2 = \rho$ if $|\Psi\rangle$ is normalized). The most general density operator (pure state or statistical mixture) is described by the matrix:

$$\rho = \frac{1}{2} [1 + a\sigma_x + b\sigma_y + c\sigma_z] = \frac{1}{2} [1 + \mathbf{M} \cdot \boldsymbol{\sigma}] \quad (11.57)$$

where vector $\mathbf{M}$, with components a, b , and c on three axes Ox , Oy , and Oz , is called the “Bloch vector”. If its length is equal to 1, the state is pure as in (11.56); if its length is smaller than 1, one has a statistical mixture ($\rho^2 \neq \rho$). Chapter IV of [11-1] gives more examples of quantum calculations performed in a space of states with dimension 2.

11.2 Grouping several physical systems

Consider two physical systems 1 and 2, the first with space of states $\mathcal{F}$ containing state vectors $|\Phi\rangle$, the second with space of states $\mathcal{G}$ containing state vectors $|\Xi\rangle$. Each of them may be considered as a sub-system of a larger physical system containing both, and having a space of states $\mathcal{E}$.

11.2.1 Tensor product

Assume that the ensemble of kets $\{|u_i\rangle\}$ provides a basis of $\mathcal{F}$; any vector $|\Phi\rangle$ can then be written as a linear combination:

$$|\Phi\rangle = \sum_i x_i |u_i\rangle \quad (11.58)$$

where the x_i are the (complex) components of $|\Phi\rangle$ in this basis. Similarly, the vectors $\{|v_j\rangle\}$ provide a basis of $\mathcal{G}$ and we can write:

$$|\Xi\rangle = \sum_j y_j |v_j\rangle \quad (11.59)$$

The space of states $\mathcal{E}$ of the physical system made by grouping the two preceding sub-systems is spanned by vectors defining the individual state of each sub-system. These vectors are the “tensor products”, which we write:

$$|1 : u_i\rangle \otimes |2 : v_j\rangle \quad (11.60)$$

We will often use a simplified notation for these kets, assuming that the first state always refer to sub-system 1:

$$|1 : u_i\rangle \otimes |2 : v_j\rangle \equiv |1 : u_i; 2 : v_j\rangle \equiv |u_i, v_j\rangle \quad (11.61)$$

If K is the dimension of $\mathcal{F}$ (number of distinct values for index i) and Q that of $\mathcal{G}$ (number of distinct values of j), the number of distinct product vectors is $R = KQ$, which determines the dimension of $\mathcal{E}$. Any state $|\Psi\rangle$ belonging to $\mathcal{E}$ can then be written:

$$|\Psi\rangle = \sum_{i=1}^K \sum_{j=1}^Q z_{i,j} |u_i, v_j\rangle \quad (11.62)$$

where the complex numbers $z_{i,j}$ are its KQ components. Space $\mathcal{E}$ is called the tensor product of $\mathcal{F}$ and $\mathcal{G}$:

$$\mathcal{E} = \mathcal{F} \otimes \mathcal{G} \quad (11.63)$$

In the simple case where all components $z_{i,j}$ of $|\Psi\rangle$ in (11.62) can be written as products:

$$z_{i,j} = x_i y_j \quad (11.64)$$

the state vector $|\Psi\rangle$ is a tensor product:

$$|\Psi\rangle = |\Phi\rangle \otimes |\Xi\rangle \quad (11.65)$$

This corresponds to two quantum systems without any correlation⁶. But this is only a special case: in general, one has to use relation (11.62), which does not factorize. We discuss the relation between this general expression and the notion of quantum entanglement in Chapter 6.

We have explicitly studied two physical systems grouped into a single system, as for instance a system of two particles (cf. §11.3.3). Needless to say, this operation can be generalized to any number of particles: the space of state $\mathcal{E}$ of a system made of N particles with individual spaces of states $\mathcal{F}_i$ ($i = 1, 2, \dots, N$) is the tensor product of all these spaces:

$$\mathcal{E} = \mathcal{F}_1 \otimes \mathcal{F}_2 \otimes \dots \otimes \mathcal{F}_N \quad (11.66)$$

and (11.62) may be generalized in the same way.

Remark: to simplify the discussion, we have assumed that the individual spaces of states $\mathcal{F}$ and $\mathcal{G}$ have finite dimensions P and Q , but the reasoning can be generalized to cases where one of these dimensions is infinite, or both. For instance, in the case of a particle with spin 1/2 (§11.3.2.b), the orbital space of states (associated to the position of the particle) has an infinite dimension, while that of spin has dimension 2; the result is then that the dimension of the whole space of states (tensor product) is also infinite.

⁶ Similarly, with classical probabilities, a distribution of two variables that is a product corresponds to non-correlated variables.

11.2.2 Ensemble of spins 1/2

The simplest case occurs when both $\mathcal{F}$ and $\mathcal{G}$ are state spaces of spins 1/2, with dimension 2 each. Space of states $\mathcal{E}$ then has dimension 4, with a basis provided by vectors:

$$|1 : +; 2 : +\rangle, \quad |1 : +; 2 : -\rangle, \quad |1 : -; 2 : +\rangle, \quad \text{and} \quad |1 : -; 2 : -\rangle \quad (11.67)$$

where, for instance, $|1 : +; 2 : -\rangle$ denotes the state where the Oz component of the first spin is equal to $+\hbar/2$, and that of the second spin equal to $-\hbar/2$. We will write these four vectors in a simpler way, without making the numbering of the spins explicit, as:

$$|+, +\rangle \quad |+, -\rangle \quad |-, +\rangle \quad |-, -\rangle \quad (11.68)$$

The most general state vector belonging to $\mathcal{E}$ is any linear combination of these four vectors.

A particular state in this space, used in many examples, is the “singlet” state:

$$|\Psi\rangle = \frac{1}{\sqrt{2}} \left[|+, -\rangle - |-, +\rangle \right] \quad (11.69)$$

One special property of this state is that it is rotation invariant: it keeps exactly the same properties if, instead of a reference axis Oz to characterize both spins, we choose any other arbitrary direction Ou . Moreover, the two spins are anti-correlated: if the component of the first spin along any direction is found positive, the component of the second spin on a parallel axis is found negative and opposite. This property is essential for the discussion of §4.1.1.

For N spins 1/2, one proceeds in the same way. Their space of states is the tensor product of all individual spin states, with dimension 2^N and spanned by the 2^N vectors:

$$|\pm, \pm, \pm, \dots, \pm\rangle \quad (11.70)$$

A very special state that generalizes (11.69) is the state often called “GHZ state”, or state “by all or nothing”:

$$|\Psi\rangle = \left[\alpha |+, +, +, \dots, +\rangle + \beta |-, -, -, \dots, -\rangle \right] \quad (11.71)$$

(where α and β are two complex numbers with the sum of their squared modulus equal to 1). This state is remarkable because it contains only two components where all spins change from one individual state to an orthogonal state. It exhibits marked quantum properties, which were discussed in §§5.1 and 5.2. It should not

be confused with a product state, where all spins are uncorrelated, and which can be written:

$$|\Psi\rangle = [\alpha|+\rangle + \beta|-\rangle] [\alpha|+\rangle + \beta|-\rangle] \dots [\alpha|+\rangle + \beta|-\rangle] \quad (11.72)$$

This state exhibits properties that are more similar to those of a classical state.

11.2.3 Partial traces

The density operator may be used to introduce the convenient notion of “partial traces”. Consider a system made of two sub-systems, 1 and 2. If they are uncorrelated and each in a pure state $|\Phi(1)\rangle$ and $|\Xi(2)\rangle$, the state of the whole system is given by:

$$|\Psi(1, 2)\rangle = |\Phi(1)\rangle \otimes |\Xi(2)\rangle \quad (11.73)$$

The density operator of this system is then simply the product of the projectors $\rho_1(1)$ and $\rho_1(2)$ over states $|\Phi(1)\rangle$ and $|\Xi(2)\rangle$:

$$\rho_{12}(1, 2) = |\Psi(1, 2)\rangle \langle\Psi(1, 2)| = |\Phi(1)\rangle \langle\Phi(1)| \otimes |\Xi(2)\rangle \langle\Xi(2)| = \rho_1(1) \rho_1(2) \quad (11.74)$$

All three systems, the whole system and the sub-systems, are in pure states.

But, if $|\Psi(1, 2)\rangle$ is not a product (if the two sub-systems are entangled), the situation is more complex. One can start from the density operator of the whole system:

$$\rho_{12} = |\Psi(1, 2)\rangle \langle\Psi(1, 2)| \quad (11.75)$$

and perform an operation called “partial trace”, which transforms ρ into an operator ρ_1 that acts only in the space of states of the first sub-system:

$$\rho_1 = Tr_2 \{\rho_{12}\} \quad (11.76)$$

The matrix elements of ρ_1 are defined⁷ by:

$$\langle u_i | \rho_1 | u_j \rangle = \sum_k \langle u_i, v_k | \rho_{12} | u_j, v_k \rangle \quad (11.77)$$

From this partial trace one may obtain all probabilities and average values associated with measurements performed only on the first sub-system, since:

$$Tr \{A(1) \rho\} = Tr_1 \{A(1) \rho_1\} \quad (11.78)$$

⁷ The partial density operator ρ_1 is independent of the basis $\{|u_i, v_j\rangle\}$ that is used to define it. One can show from (11.77) that, in any other basis $\{|u'_q, v'_l\rangle\}$ one obtains $\langle u'_q | \rho_1 | u'_r \rangle = \sum_l \langle u'_q, v'_l | \rho_{12} | u'_r, v'_l \rangle$.

where $A(1) \equiv A(1) \otimes I(2)$ is an operator acting only in the space of states of the first sub-system, while the identity operator in the space of states of the sub-system is denoted $I(2)$; the right-hand part of the equation is a trace taken only in the space of states of sub-system 1. Similarly, one defines another partial trace ρ_2 , which relates only to the properties of the other sub-system 2.

A specific property of quantum mechanics is that, even if the whole system is known with the best possible accuracy (it is in a pure state), in general its sub-systems are not in this case: they are described by statistical mixtures, and therefore with an accuracy that is not maximal. A classical example is given by two spins in a singlet state (11.69), for which the whole system is indeed in a pure state; if one calculates the partial trace over any of the two spins, one obtains the matrix associated with individual spins in the form:

$$\rho_1(1) = \rho_2(2) = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix} \quad (11.79)$$

Under these conditions, the state of each spin is completely unknown: it has the same probability to be in state $|+\rangle$ or $|-\rangle$ or, in fact, in any linear combination of these states⁸. Therefore, even if the whole system is known with the best possible accuracy in quantum mechanics, in this case no information whatsoever is available concerning the two sub-systems, a situation that has no equivalent in classical physics (for a more detailed discussion of this unusual situation, in particular by Schrödinger, see §6.1).

11.3 Particles in a potential

We now apply the general formalism to a few simple cases.

11.3.1 Single particle

In classical mechanics, the position of a point particle is defined by the three components of its position $\mathbf{r}$ at time t . The momentum $\mathbf{p}$ of the particle is a vector with three components given by the conjugate momenta of the components of $\mathbf{r}$:

$$\mathbf{p} = m \frac{d}{dt} \mathbf{r} \quad (11.80)$$

where m is the mass of the particle. The dynamical state of the particle at any time is defined by the simultaneous specification of its position $\mathbf{r}$ and of its momentum

⁸ The matrix is proportional to the unit matrix, which keeps the same form in any basis.

$\mathbf{p}$, which corresponds to six scalar variables if the particle moves in ordinary three-dimensional space.

When a potential $V(\mathbf{r})$ acts on the particle, the evolution of the state of the particle may be obtained from an Hamiltonian $\mathcal{H}$, which is nothing but the sum of its kinetic energy and potential energy $\mathcal{V}$:

$$\mathcal{H}(\mathbf{r}, \mathbf{p}; t) = \frac{\mathbf{p}^2}{2m} + \mathcal{V}(\mathbf{r}; t) \quad (11.81)$$

11.3.1.a Wave function

In quantum mechanics, the state of the particle at time t is no longer defined by its position and momentum, but by a state vector $|\Psi(t)\rangle$ belonging to a space of states $\mathcal{E}_{\mathbf{r}}$; in this space, a continuous “basis” (cf. note 1) is given by the ensemble of kets $|\mathbf{r}\rangle$ where the particle is perfectly localized at point $\mathbf{r}$. According to (11.5), the components of $|\Psi\rangle$ in this basis are given by a $\mathbf{r}$ -dependent function:

$$\Psi(\mathbf{r}, t) = \langle \mathbf{r} | \Psi(t) \rangle \quad (11.82)$$

with, according to (11.6):

$$|\Psi(t)\rangle = \int d^3r \Psi(\mathbf{r}, t) |\mathbf{r}\rangle \quad (11.83)$$

The complex function $\Psi(\mathbf{r}, t)$ is called the “wave function” of the particle. Since this function may extend over a whole domain of space, the position of the particle is not perfectly defined. Quantum mechanics only specifies that the probability to find at time t the particle in any volume D is given by:

$$\mathcal{P} = \int_D d^3r |\Psi(\mathbf{r}, t)|^2 \quad (11.84)$$

In other words, $|\Psi(\mathbf{r})|^2$ gives the density of probability $n(\mathbf{r})$ associated with the position variable of the particle:

$$n(\mathbf{r}, t) = |\Psi(\mathbf{r}, t)|^2 \quad (11.85)$$

The normalization condition of the total probability implies that the integral should be equal to 1 when D extends to all space:

$$\int d^3r |\Psi(\mathbf{r}, t)|^2 = \int d^3r n(\mathbf{r}, t) = 1 \quad (11.86)$$

Of course, this condition is crucial to interpret $n(\mathbf{r})$ as a density of probability. If a wave function does not provide a unit value for the integral of its square modulus,

one says that it is not normalized; but it is then sufficient to divide it by the square root of this integral to normalize it. This operation is possible only if the integral in all space is finite; one can normalize only functions that obey this condition, an essential condition for a function to be acceptable as a wave function.

The probability to find the momentum within any domain is given by a formula that is similar to (11.84), but with the probability density $\bar{n}(\mathbf{p})$ replaced by $|\bar{\Psi}(\mathbf{p})|^2$, where $\bar{\Psi}(\mathbf{p})$ is the Fourier transform of $\Psi(\mathbf{r})$.

11.3.1.b Schrödinger equation, probability current

The time evolution of the wave function $\Psi(\mathbf{r}, t)$ is given by the Schrödinger equation:

$$i\hbar \frac{\partial}{\partial t} \Psi(\mathbf{r}, t) = -\frac{\hbar^2}{2m} \Delta \Psi(\mathbf{r}, t) + V(\mathbf{r}; t) \Psi(\mathbf{r}, t) \quad (11.87)$$

From this wave function, one can define a probability current:

$$\mathbf{J}(\mathbf{r}, t) = \frac{\hbar}{2im} \left\{ \Psi^*(\mathbf{r}, t) \nabla \Psi(\mathbf{r}, t) - \Psi(\mathbf{r}, t) \nabla \Psi^*(\mathbf{r}, t) \right\} \quad (11.88)$$

and, by using (11.87), derive the equation for local conservation of the probability:

$$\frac{\partial}{\partial t} n(\mathbf{r}, t) + \nabla \cdot \mathbf{J}(\mathbf{r}, t) = 0 \quad (11.89)$$

Integrated over all space, this relation shows that the norm of any wave function remains constant as a function of time; if it is initially normalized as in (11.86), the Schrödinger equation conserves this normalization during the time evolution.

11.3.2 Spin, Stern–Gerlach experiment

The formalism we have described applies only to spinless particles, also called spin zero particles. We now study how the formalism can be adapted to particles with non-zero spin, which will allow us to better understand the origin of the two-dimensional space of states that was introduced in §11.1.6.

11.3.2.a Introduction of spin

At the time quantum mechanics was invented, the experimental study of atomic spectral lines made physicists realize that a particle such as an electron must possess an additional degree of freedom, in addition to those associated with its position in space – this idea was proposed by Uhlenbeck and Goudsmit in 1925 [11-5]. The additional degree of freedom is called “spin”; it corresponds to an internal rotation

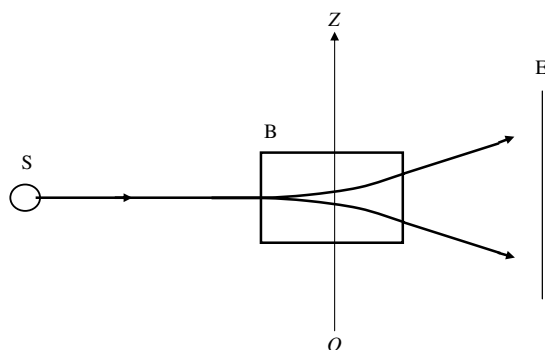


Figure 11.1. Schematic representation of a Stern–Gerlach experiment. A source S emits a beam of particles (Silver atoms), which propagate to region B where they undergo the influence of a strong magnetic gradient along direction Oz created by a magnet (not shown in the figure). The force exerted by this gradient on the magnetic moment of the particle, which is parallel to its spin, then depends on the direction of the spin; the trajectory is therefore bent in a way that depends on the Oz component of the spin.

In classical mechanics, one would expect a continuum of possible deviations for the particles (since this component can take continuous values within a whole interval). But two groups of completely different impacts are observed on screen E, corresponding to two possible deviations. This indicates that the Oz component of the spin can only take two discrete values – one says that it is quantized. In classical physics, no vector can have components on any axis that are equal to constant values. The experiment provides evidence, at the same time, for the existence of spin, and for its completely non-classical properties.

of the particle (rotation around itself). The existence of spin can not be explained with a classical image: a classical object can turn around itself only if it extends over some domain of space, while in quantum mechanics even a point particle (the electron for instance) can have a spin.

Shortly before (1922), Stern and Gerlach [6] had performed an experiment that provided a direct evidence of an internal rotation of quantum particles such as atoms; the particles they had used were Silver atoms, which have a spin arising in particular from the electrons they contain. The experiment is shown very schematically in Figure 11.1. A beam of particles (atomic beam) originates from source S and propagates to region B, where a magnet creates a magnetic field with a strong gradient along direction Oz . The particles carry a magnetic moment that is proportional to their spin, and is therefore colinear with it. The local magnetic gradient creates a force acting on this magnetic moment, so that the trajectory of the particles is bent in a direction that depends on the Oz component of their spin. The impact of the particles is eventually registered on a screen E.

Within classical mechanics one would expect that, initially, the spins should have completely random orientations, uniformly spread in all directions; this component

should then vary continuously between two extreme values. In other words, one should observe a continuum of possible values for the deviations of the particles, resulting in a broad spot on the screen. But the experimental surprise was to observe, instead, two well-separated spots: only two kinds of deviations were obtained, one along Oz , the other in the opposite direction. This result is interpreted by assuming that the component of each spin along Oz can only take two discrete values (which turn out to be equal to $\pm\hbar/2$): this is spin quantization, directly observed experimentally with this device.

In this experiment, direction Oz does not correspond to any particular direction for source S . This means that the spin component can take only one out of two opposite values, whatever measurement direction is chosen. Of course, in classical physics, such a situation is totally impossible: no vector has its component on any direction in space with constant modulus. This provides one more illustration of the completely quantum nature of spin, without any classical equivalent. For a more detailed discussion of the Stern–Gerlach experiment, and in particular of measurements of spin components along various directions, see for instance §A in Chapter IV of [1].

In fact, quantum mechanics does not necessarily imply that the number of discrete components of a spin should be two: in general, their number is given by $2s + 1$, where s is any half integer number⁹. The case we have described corresponds to $s = 1/2$, and this explains why one speaks of “spin 1/2 particle”; this applies to the electron for instance, as well as to other particles such as the proton, the neutron, etc. (but not the photon).

11.3.2.b Space of states

For a spin 1/2 particle, the formalism of non-relativistic quantum mechanics introduces two wave functions instead of one: each spin component has a wave function. One first chooses a fixed reference axis, the “quantization axis” Oz , and one defines two wave functions $\Psi_{\pm}(\mathbf{r})$; index $\pm$ specifies the sign of the Oz component of the spin. For a particle with any spin s , one would introduce $2s + 1$ wave functions, but for the sake of simplicity here we limit ourselves to the case $s = 1/2$.

Introducing two functions $\Psi_{\pm}(\mathbf{r})$, instead of one, doubles the number of components of the state vector $|\Psi\rangle$, as well as the number of vectors in a basis of the space of states. We therefore now replace (11.83) by:

$$|\Psi(t)\rangle = \int d^3r \left[\Psi_+(\mathbf{r}; t) |\mathbf{r}, +\rangle + \Psi_-(\mathbf{r}; t) |\mathbf{r}, -\rangle \right] \quad (11.90)$$

⁹ A half integer number is by definition a number that is, or becomes, integer when multiplied by 2.

where $|\mathbf{r}, +\rangle$ denotes the ket where the particle is localized at point $\mathbf{r}$ with a positive component of its spin along Oz , while $|\mathbf{r}, -\rangle$ is the corresponding state with a negative component.

One sometimes call “spinor” the ensemble of the two components of $|\Psi\rangle$, which are grouped in a column matrix as the two components of a vector in a two-dimensional space:

$$\begin{pmatrix} \Psi_+(\mathbf{r}) \\ \Psi_-(\mathbf{r}) \end{pmatrix} \quad (11.91)$$

Applying the three Pauli matrices (11.52) to the column vector (11.91) defines (after multiplication by a constant factor $\hbar/2$) the action of the operators $S_{x,y,z}$ associated with the three components of the spin. We now see the emergence of a structure of a space of states that combines the properties of spin (space of states with dimension 2, studied in §11.1.6) and those of the orbital variable $\mathbf{r}$ (space of states with infinite dimension), and provides a good illustration of the notion of tensor product (§11.2.1).

11.3.3 Several particles

Assume now that the system under study is made of several particles. In classical mechanics, the evolution of a system containing N particles with masses $m_1, m_2, \dots, m_N$ and positions $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N$ involves an Hamiltonian:

$$\mathcal{H}(\mathbf{r}, \mathbf{p}; t) = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m_i} + \mathcal{V}(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N; t) \quad (11.92)$$

where $\mathcal{V}(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N; t)$ is the sum of the external potential acting on the particles and of their mutual interaction potential.

In quantum mechanics, for an ensemble of N spinless particles, the single particle wave function $\Psi(\mathbf{r}, t)$ is replaced by a wave function $\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N; t)$ depending of the positions $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N$ of all particles:

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N; t) = \langle \mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N | \Psi(t) \rangle$$

The function:

$$n(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N; t) = |\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N; t)|^2 \quad (11.93)$$

generalizes (11.85) and provides the probability to find the first particle at point $\mathbf{r}_1$, the second at point $\mathbf{r}_2$, ... the last at point $\mathbf{r}_N$. The wave function Ψ is physically

acceptable only if the integral of n in all configuration space (with $3N$ dimensions) is finite, which makes the normalization of Ψ possible.

The introduction of the notion of tensor space is natural in terms of wave functions. For two particles, we know that a function $\Psi(\mathbf{r}_1, \mathbf{r}_2)$ can be expanded as a sum of products of $\mathbf{r}_1$ functions by $\mathbf{r}_2$ functions:

$$\Psi(\mathbf{r}_1, \mathbf{r}_2) = \sum_{\mu} \sum_{\nu} \phi_{\mu}(\mathbf{r}_1) \phi_{\nu}(\mathbf{r}_2) \quad (11.94)$$

(where the sums may possibly be infinite, or even continuous so that they become integrals over μ and ν): the space of two-particle wave functions $\Psi(\mathbf{r}_1, \mathbf{r}_2)$ is the tensor product of the spaces of single-particle wave functions. This notion can be generalized to N particles, for which the wave functions can be written:

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = \sum_{\mu} \sum_{\nu} \dots \sum_{\xi} \phi_{\mu}(\mathbf{r}_1) \phi_{\nu}(\mathbf{r}_2) \dots \phi_{\xi}(\mathbf{r}_N) \quad (11.95)$$

and belong to the tensor product space of N spaces of single particle wave functions.

The Schrödinger equation is a generalization of (11.87), which reads:

$$i\hbar \frac{\partial}{\partial t} \Psi(\mathbf{r}_1, \dots, \mathbf{r}_N; t) = -\frac{\hbar^2}{2m} \Delta \Psi(\mathbf{r}_1, \dots, \mathbf{r}_N; t) + V(\mathbf{r}_1, \dots, \mathbf{r}_N; t) \times \Psi(\mathbf{r}_1, \dots, \mathbf{r}_N; t) \quad (11.96)$$

where, as in classical mechanics, the potential V may include a part due to an external potential acting separately on each particle as well as a mutual interaction part (generally the sum of binary interactions). One can define a probability current $\mathbf{J}$ in a space with $3N$ dimensions (configuration space); this generalizes relations (11.88) and (11.89) and provides a multidimensional conservation relation.

If the N particles are spin 1/2 particles, the wave functions splits into 2^N components, labeled with N equal to $\pm$:

$$\Psi_{\pm, \pm, \dots, \pm}(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N; t) \quad (11.97)$$

Except for this change, the general idea remains the same: the space of states of the whole system of particles is the tensor product of the single-particle spaces of states.

Remark: one can also quantize other physical systems than particles, for instance fields. For an introduction, see for instance [7] and [8].

Bibliography

[1] C. Cohen-Tannoudji, B. Diu, and F. Laloë, *Quantum Mechanics*, Wiley (1977).

[2] A. Messiah, *Mécanique Quantique*, Dunod (1962); *Quantum Mechanics*, Wiley (1961), Dover (1999).

[3] A. Peres, *Quantum Theory, Concepts and Methods*, Kluwer Academic Press (1995).

[4] M. Le Bellac, *Quantum Physics*, Cambridge University Press (2011).

[5] G. E. Uhlenbeck and S. Goudsmit, *Naturwissenschaften* **47**, 953 (1925); *Nature* **117**, 264 (1926). For an interesting historical description of the way spin was discovered, see: <http://www.ilorentz.org/history/spin/goudsmit.html>.

[6] W. Gerlach and O. Stern, “Der experimentelle Nachweis der Richtungsquantelung im Magnetfeld”, *Zeit. Phys.* **9**, 349–355 (1922).

[7] L. Schiff, *Quantum Mechanics*, McGraw Hill (1955).

[8] C. Cohen-Tannoudji, J. Dupont-Roc, and G. Grynberg, *Atom–Photon Interactions*, Wiley (1989).

Appendix A

Mental content of the state vector

(i) A relatively recent (1999) article by Englert, Scully, and Walther [85] provides an interesting illustration of the debate about the content of the state vector, in particular because of the wording chosen by the authors. Speaking of standard theory, they write: “(One) is led astray by regarding state reductions as physical processes, rather than accepting that they are nothing but mental processes”. They then advocate a “minimalistic interpretation of state vectors” and actually even give a general warning that it is dangerous to go beyond it (“Van Kampen’s caveat¹”), but do not expand very much on these dangers. It is then interesting to extend this line of thought further: if the state vector can evolve under the effect of a “purely mental process”, it then necessarily acquires elements that are mental (subjective), instead of being related to external reality only (objective). The Schrödinger evolution, on the other hand, is determined by external macroscopic parameters, and therefore has a reality content that is similar with these parameters. Should we consider that the state vector is hybrid, and combines elements describing external reality, at least partially, and others that are purely mental, all contained in a single mathematical object²? In this view, the process of state vector reduction would correspond to times when the state vector suddenly acquires more mental elements (since the reduction process is considered as purely mental); the Schrödinger evolution would correspond to periods of time where the mental content remains constant.

Such a hybrid view concerning the reality content of the state vector is nevertheless not compatible with the standard view where it defines a physical preparation procedure, which can be a measurement process involving state reduction (selection of atoms in one beam at the output of a Stern–Gerlach filter for instance).

¹ “Whoever endows the state vector with more meaning than is needed for computing observable phenomena is responsible for the consequences” [84].

² This in itself would not necessarily be a problem. A classical statistical distribution function for an ensemble of time-dependent systems has similar properties: it does combine real evolution with elements related to imperfect initial knowledge on the systems.

If the effect of the measurement is a purely mental projection of the state vector, the resulting state vector should also be purely mental. With this definition, a state vector should then have no real content at all.

But we have already noted (§10.3.2) that most physicists would probably agree that, at least sometimes, the wave function does contain elements of reality, and illustrated this with a few examples (BCS state of electrons, etc.). One usually considers that Schrödinger evolution contains at the same time a physical evolution for those properties that quantum mechanics does attribute to the system (observables that have $|\Psi\rangle$ as an eigenvector), but also an evolution of probabilities (for all other observables) that represent only our knowledge of the system, and therefore can indeed be seen as mental processes. In fact, the notion of a “purely mental process” is not often mentioned by supporters of the Copenhagen interpretation (maybe with the exception of Wigner, see §10.1.1.b, but the Wigner interpretation is not really standard). Peres for instance, in the quotation of [44] given in §1.2.3.b, while also orthodox, never invokes human minds, but just preparations and tests on physical systems.

(ii) Another interesting illustration is provided by a note published by Peres and Fuchs [86] entitled “Quantum theory needs no interpretation” – especially since this note has stimulated many reactions from authors expressing various points of view. It goes even further than [85] since these authors seem to take explicitly a point of view where the wave function is not absolute, but observer dependent³. After stating that: “Quantum theory does *not* describe physical reality. What it does is provide an algorithm for computing *probabilities* for the macroscopic events that are the consequences of our experimental interventions”, they add that “a wave function is only a mathematical expression for evaluating probabilities and depends on the knowledge of whoever is doing the computing”. The wave function then becomes really similar to a classical probability distribution which, obviously, depends on the knowledge of the experimenter; several different distributions can be associated with the same physical system if there are several observers. As mentioned above (1.2.3.a), associating several different wave functions with one single system is not part of what is usually called the orthodox interpretation (except, of course, for a trivial phase factor); on the other hand, this idea is part of the relational interpretation (§10.3.1).

³ As in §1.2.3.a, we assume that all observers use the same Galilean reference frame. Otherwise, they need different state vectors for describing the system, for trivial reasons (as in classical mechanics).

Appendix B

Bell inequalities in non-deterministic local theories

In the derivations of the various forms of Bell inequalities resulting from local realism (§4), we have assumed that the results of experiments are known functions $A(a, \lambda)$ and $B(b, \lambda)$, which depend on the settings a and b and of the additional variable λ . This is a natural continuation of the EPR theorem, which concludes on the existence of these functions from the assumption of local realism, in other words on determinism. Here we show that the Bell inequalities are actually more general and that, provided locality is preserved, they can still be established within a non-deterministic context.

The basic idea behind this generalization is straightforward: probabilities can always be considered as resulting from a deterministic process controlled by one more additional random variable, which we call μ here. Adding this variable does not change anything to the reasoning that leads to the inequalities; actually, λ can be seen as one single variable with several dimensions, one of its components being μ .

Consider a given value of λ . Within a non-deterministic theory, mathematically we have to replace $A(a, \lambda)$ by two probabilities $\mathcal{P}_+^A(a, \lambda)$ and $\mathcal{P}_-^A(a, \lambda)$, and $B(b, \lambda)$ by two probabilities $\mathcal{P}_+^B(b, \lambda)$ and $\mathcal{P}_-^B(b, \lambda)$. For any a and λ we have:

$$\mathcal{P}_+^A(a, \lambda) + \mathcal{P}_-^A(a, \lambda) = 1 \quad (\text{B.1})$$

with a similar condition for the probabilities $\mathcal{P}_\pm^B$. We can then introduce one additional variable μ having uniform distribution within interval $[0, 1]$, and define a function $\bar{A}(a, \lambda, \mu)$ by¹:

$$\bar{A}(a, \lambda, \mu) = \begin{cases} +1 & \text{if } 0 \leq \mu \leq \mathcal{P}_+^A(a, \lambda) \\ -1 & \text{if } \mathcal{P}_+^A(a, \lambda) < \mu \leq 1 \end{cases} \quad (\text{B.2})$$

¹ We use the notation $\bar{A}$ to make a distinction from A , which was defined in the main text as the result of the experiment. Here, $\bar{A}$ is just a mathematical variable introduced to express probabilities in a convenient way.

We then have:

$$\mathcal{P}_+^A(a, \lambda) - \mathcal{P}_-^A(a, \lambda) = \int_0^1 d\mu \bar{A}(a, \lambda, \mu) \quad (\text{B.3})$$

and, of course, a similar equation for the difference $\mathcal{P}_+^B(b, \lambda) - \mathcal{P}_-^B(b, \lambda)$.

Suppose now that we wish to calculate the average of the product of the two results observed with settings a and b . For each value of λ , we have to sum the probabilities associated with the $+, +$ and $-, -$ events and to subtract the probabilities of the $+, -$ and $-, +$ events, which amounts to introducing the term:

$$\begin{aligned} & \mathcal{P}_+^A(a, \lambda) \mathcal{P}_+^B(b, \lambda) + \mathcal{P}_-^A(a, \lambda) \mathcal{P}_-^B(b, \lambda) - \mathcal{P}_+^A(a, \lambda) \mathcal{P}_-^B(b, \lambda) - \mathcal{P}_-^A(a, \lambda) \mathcal{P}_+^B(b, \lambda) \\ &= \left[\mathcal{P}_+^A(a, \lambda) - \mathcal{P}_-^A(a, \lambda) \right] \times \left[\mathcal{P}_+^B(b, \lambda) - \mathcal{P}_-^B(b, \lambda) \right] \end{aligned} \quad (\text{B.4})$$

The average value of this expression is obtained by a sum over λ which, according to (B.3), gives the integral:

$$\int d\lambda \int_0^1 d\mu \bar{A}(a, \lambda, \mu) \int_0^1 d\mu' \bar{B}(b, \lambda, \mu') \quad (\text{B.5})$$

At this point, we see that we obtain the same expression as with a deterministic theory, with an integral of the product of two functions that are equal to ± 1 . The only difference is the presence of two additional integration variables μ and μ' : the local stochastic theory is therefore equivalent to a deterministic theory with more additional variables. The rest of the reasoning leading to Bell inequalities remains unchanged.

The conclusion is that these inequalities are also valid for non-deterministic theories, provided the dependences of the probabilities are local – if we had assumed that $\mathcal{P}_+^A(a, \lambda)$ also depends on b , the proof of the inequalities would no longer have been possible.

Appendix C

An attempt for constructing a “separable” quantum theory (non-deterministic but local)

We give an example of such a non-deterministic local theory that looks similar to quantum mechanics, and actually even makes use of its formalism, while it is in fact significantly different. This theory includes the non-determinism of quantum mechanics, but gives to the state vector a role that is more local than in standard quantum mechanics. For this, we consider a physicist who has well understood the basic rules of quantum mechanics concerning non-determinism, but who remains sceptical about non-locality (or non-separability; for a detailed discussion of the meaning of these terms, see §3.3.3.b as well as, for instance, [24, 47]). This physicist thinks that, if measurements are performed in remote regions of space, it is more natural to apply the rules of quantum mechanics separately in these two regions. In order to calculate the probability of any measurement result, he/she will then apply the rules of quantum mechanics, in a way that is locally perfectly correct, but that also assumes that it is possible to reason separately in the two regions of space. If for instance the two measurements take place in two different galaxies, our sceptical physicist is prepared to apply quantum mechanics at the scale of a galaxy, but not at an intergalactic scale!

How can one then treat the measurement process that takes place in the first galaxy? It is very natural to assume that the spin it contains is described by a state vector (or by a density operator, it makes no difference here) that may be used to apply the orthodox formula for obtaining the probabilities of each possible result. Of course it would not be a good idea to assume that each spin is described by a single, fixed, density operator: obviously, the model would then exclude correlations between the results of measurements performed in the two galaxies. A better idea is to assume that the density operators in question are random mathematical objects $\rho_1(\lambda)$ and $\rho_2(\lambda)$ that fluctuate because they are functions of a random variable λ , corresponding for instance to fluctuating conditions of emission of the particles. The method is then clear: for any possible condition of the emission, one performs an orthodox quantum calculation in each region of space, and then takes an average

value over the conditions in question. After all, this is nothing but the universal method for calculating correlations in all the rest of physics! This approach takes into account the indeterministic character of quantum mechanics¹, but introduces a notion of space separability that is directly in the line of the EPR reasoning. Our physicist may for instance assume that the two measurement events are separated by a space-like interval in the sense of relativity, so that no causal relation can relate them in any circumstance; this seems to fully justify an independent calculation of both phenomena.

If we note $|+(a)\rangle$ the eigenstate of the measurement corresponding to result $+1$ and $\rho_1(\lambda)$ the density operator of s , the probability of obtaining result $+1$ if the first measurement is made along direction a is then:

$$\mathcal{P}_+^A(a, \lambda) = \langle +(a) | \rho_1(\lambda) | +(a) \rangle \quad (\text{C.1})$$

In the same way, we write the probability for the result -1 in the form:

$$\mathcal{P}_-^B(a, \lambda) = \langle -(a) | \rho_1(\lambda) | -(a) \rangle \quad (\text{C.2})$$

If, instead of direction a , another different direction a' is chosen, the calculations remain the same and lead to two functions $\mathcal{P}_\pm^A(a', \lambda)$. As for measurements performed in the second region of space, they provide two functions $\mathcal{P}_\pm^B(b, \lambda)$ and $\mathcal{P}_\pm^B(b, \lambda)$.

We now calculate the average of the product of the two results, which is nothing but the average over λ of the expression already written (B.4). If now we define $\widehat{A}(\lambda)$ and $\widehat{B}(\lambda)$ by:

$$\begin{aligned} \widehat{A}(\lambda) &= \mathcal{P}_+^A(a, \lambda) - \mathcal{P}_-^A(a, \lambda) \\ \widehat{B}(\lambda) &= \mathcal{P}_+^B(b, \lambda) - \mathcal{P}_-^B(b, \lambda) \end{aligned} \quad (\text{C.3})$$

we can write the average of the product of results as:

$$\int d\lambda \, n(\lambda) \, \widehat{A}(\lambda) \, \widehat{B}(\lambda) \quad (\text{C.4})$$

where $n(\lambda)$ is the distribution density of variable λ .

The difference with the usual result is that, here, $\widehat{A}(\lambda)$ and $\widehat{B}(\lambda)$ are not defined as functions that are always equal to ± 1 . To complete the proof, we have the choice between two different methods:

¹ In §7.2.2.b, we remarked that this is an appropriate method, within standard quantum mechanics, to study situations where Alice and Bob observe correlations when Eve interferes and measures polarizations before the particles reach the two partners.

(i) either we proceed as in Appendix B and introduce an additional variable μ to express the probability differences as in (B.3), in terms of new quantities $\bar{A}(a, \lambda, \mu)$ and $\bar{B}(b, \lambda, \mu')$ that are always equal to ± 1 . This leads to the same mathematical expression and proof of the BCHSH inequalities as above.

(ii) or we introduce the couples of orientations (a, b) , (a, b') , (a', b) , (a', b') to calculate the average over λ of the expression:

$$\hat{A}(\lambda)\hat{B}(\lambda) - \hat{A}(\lambda)\hat{B}'(\lambda) + \hat{A}'(\lambda)\hat{B}(\lambda) + \hat{A}'(\lambda)\hat{B}'(\lambda) \quad (\text{C.5})$$

The $\hat{A}$ and $\hat{B}$, which are now defined as probability differences, are no longer necessarily ± 1 ; but it is nonetheless easy to see that they have values between $+1$ and -1 , whatever the value of λ is. This implies² that expression (C.5) is necessarily between ± 2 , which brings us back to the calculation of Section 4.1.2.

Once more, we find that the Bell theorem holds in a large variety of situations! One may wonder what exactly went wrong in the approach of our sceptical physicist, and why his results, because they satisfy the Bell inequalities, are necessarily incompatible with standard mechanics (not necessarily always, but at least sometimes). After all, his reasoning was based on the use of the usual formalism of quantum mechanics. In fact, what caused the error was the insistence of treating the quantum measurements as separable events, while orthodox quantum mechanics requires us to consider the whole two-spin system as a single, non-separable, system; in this system, no attempt should be made to distinguish sub-systems. The correct reasoning uses only state vectors/density operators that describe this whole system with one single mathematical object. This example illustrates how it is really separability and/or locality that are at stake in a violation of the Bell inequalities, not determinism.

² To see why, let us for a moment consider λ , $\hat{A}$ and $\hat{A}'$ as fixed, keeping only $\hat{B}$ and $\hat{B}'$ as variables; in the space of these variables, expression (C.5) corresponds to a plane surface which, at the four corners of the square $\hat{B} = \pm 1$, $\hat{B}' = \pm 1$, takes values $\pm 2\hat{A}$ or $\pm 2\hat{A}'$, which are between ± 2 ; at the center of the square, the plane goes through the origin. By linear interpolation, it is clear that, within the inside of the square, the function given by (C.5) also remains bounded between ± 2 ; finally, its average value has the same property.

Appendix D

Maximal probability for a state

In this appendix, we give more details on the calculations of §5.4; the two-particle state corresponding to the measurement considered in (i) is the tensor product of ket (5.33) by its correspondent for the second spin:

$$\cos^2 \theta |+, +\rangle + \sin \theta \cos \theta [|+, -\rangle + |-, +\rangle] + \sin^2 \theta |-, -\rangle \quad (\text{D.1})$$

which has the following scalar product with ket (5.38):

$$\cos^2 \theta \sin \theta - 2 \sin \theta \cos^2 \theta = -\sin \theta \cos^2 \theta \quad (\text{D.2})$$

The requested probability is obtained by dividing the square of this expression by the square of the norm of state vector (5.38):

$$\mathcal{P} = \frac{\sin^2 \theta \cos^4 \theta}{2 \cos^2 \theta + \sin^2 \theta} = \frac{\sin^2 \theta (1 - \sin^2 \theta)^2}{2 - \sin^2 \theta} \quad (\text{D.3})$$

A plot of this function shows that it has a maximum of about 0.09.

Appendix E

The influence of pair selection

In the proof of the Bell theorem, we have assumed that all pairs of particles emitted by the source are actually detected, whatever choice is made for the measurement settings a and b ; within local realism, the Bell inequalities are then obeyed, which means that it is impossible to reproduce the a and b dependence (4.9) of the correlation function predicted by quantum mechanics (since it allows violations of the Bell inequalities). In this appendix, we examine what happens when the detection process introduces a selection in the ensemble of emitted pairs; in a first step (§E.1), we assume that this selection is independent of a and b , and in a second step (§E.2), we generalize to include a possible dependence. In the latter case, we will see that it then becomes possible to reproduce any variation of the correlation rate as a function of a and b , including the prediction (4.9) in $\cos(a - b)$ of quantum mechanics, while remaining within local realism; this is the origin of the “loophole” discussed in §4.5.1.a.

E.1 Setting independent selection

Our first model is as follows:

(i) We assume that some process selects particles in a way that is independent of the choice of the experimental settings a and b . We can for instance assume that the particles are emitted randomly by the source in many directions, while only the particles that are emitted into a small solid angle can reach the analyzers and detectors. For each particle, we characterize this emission direction with a random variable ω . The first particle is detected only if its random variable ω_1 falls inside a range $D_\lambda(\Omega_1)$ corresponding to a cone originating at the source and going to the periphery of the input diaphragm of the detector:

$$\omega_1 \in D_1(\Omega_1) \tag{E.1}$$

(otherwise it is lost). Similarly, the second particle is detected only if its random variable ω_2 falls inside domain $D_\lambda(\Omega_2)$, and is lost otherwise. This sort of angular selection occurs in practice in all experiments where the small size of detectors puts a strong angular limit; the two domains are then small ω domains with sizes $d\Omega_{1,2}$ centered around values $\omega_{\text{det } 1,2}$.

(ii) The particles that have survived the preceding selection arrive on the analyzers and detectors, and provide results that depend on a random variable λ_1 for the first, λ_2 for the second. As in §4.1.2, these results depend also of the local choice of the measurement setting; they are noted $A(a, \lambda_1) = \pm 1$ for the first particle, $B(b, \lambda_2)$ for the second (we assume that two channel detectors are used, as in an EPRB experiment). For the particles that have not been selected in process (i), we choose by convention to attribute to them result 0.

The ensemble of variables $\lambda_1, \lambda_2, \omega_1, \omega_2$, can be grouped formally into a single parameter Λ with several components (a vector in a space with many dimensions). Every pair is characterized by a given value of Λ , which determines if the pair will be detected or not by the detection apparatuses and the results that they will provide. Assume now that the source emits the pairs of particles randomly; the values of Λ are then associated to a probability distribution $\rho(\Lambda)$, with the normalization condition:

$$\int d\Lambda \rho(\Lambda) = \int d\lambda_1 \int d\lambda_2 \int d\omega_1 \int d\omega_2 \rho(\lambda_1, \lambda_2, \omega_1, \omega_2) = 1 \quad (\text{E.2})$$

The average value of the product of results is then:

$$\langle AB \rangle = \int d\lambda_1 \int d\lambda_2 \int_{D_1(\Omega_1)} d\omega_1 \int_{D_2(\Omega_2)} d\omega_2 \rho(\Lambda) A(a, \lambda_1) B(b, \lambda_2) \quad (\text{E.3})$$

which, if the sizes of the domains $d\Omega_{1,2}$ centered around values $\omega_{\text{det } 1,2}$ are small, is also given by:

$$\langle AB \rangle \simeq d\Omega_1 d\Omega_2 \int d\lambda_1 \int d\lambda_2 \rho(\lambda_1, \lambda_2, \omega_{\text{det } 1}, \omega_{\text{det } 2}) A(a, \lambda_1) B(b, \lambda_2) \quad (\text{E.4})$$

The two formulas are exactly of the type that can be handled by the Bell theorem; they allow to derive inequalities such as, for instance, the BCHSH inequality (4.8). We note in passing that, if the selection is very efficient, most particles give result zero and do not contribute to the average value; as a consequence, the value of expression (E.4) is necessarily very small (it contains the very small product $d\Omega_1 d\Omega_2$); it is then not surprising that the sum of four average values can never reach values ± 2 . In quantum mechanics, the probability of detecting pairs is also proportional to $d\Omega_1 d\Omega_2$, and therefore low; there is no hope either to obtain violations of the Bell inequalities beyond ± 2 , and the situation is not extremely interesting.

Nevertheless, one can also take a different point of view: instead of normalizing the average values over the ensemble of emitted pairs, we can normalize over the smaller ensemble of detected pairs. The normalization condition (E.2) then becomes:

$$\int d\lambda_1 \int d\lambda_2 \int_{D_1(\Omega_1)} d\omega_1 \int_{D_2(\Omega_2)} d\omega_2 \rho(\lambda_1, \lambda_2, \omega_1, \omega_2) = 1 \quad (\text{E.5})$$

or:

$$d\Omega_1 d\Omega_2 \int d\lambda_1 \int d\lambda_2 \rho(\lambda_1, \lambda_2, \omega_{\text{det } 1}, \omega_{\text{det } 2}) \simeq 1 \quad (\text{E.6})$$

This change of normalization introduces a factor $1/d\Omega_1 d\Omega_2$ into ρ , which cancels that of the average value (E.4). As a consequence, the strong reduction of the average values does not occur anymore, and we merely come back to a case that is exactly equivalent to what happens in the absence of selection. The BCHSH sum of four average values can then reach the values ± 2 . In quantum mechanics, since the small factor has now disappeared, one recovers the $\cos(a - b)$ dependence and the violation of the Bell inequality.

E.2 Setting dependent selection

We now replace assumption (i) above by a more general assumption, which no longer excludes a setting dependence in the selection process:

(iii) An a and b dependent selection process occurs; for instance, after being emitted, when the particles fly towards the detectors, they are part of a physical process (absorption for instance) that destroys some fraction of them in a way that depends on a for the first particle, b for the second. The no-absorption condition for the first particle then becomes:

$$\omega_1 \in D_1(a) \quad (\text{E.7})$$

and a similar condition applies to the second particle. If a and b are angles, we can write the no-absorption condition in the form:

$$a - \Delta a < \omega_1 < a + \Delta a \quad \text{and} \quad b - \Delta b < \omega_2 < a + \Delta b \quad (\text{E.8})$$

where Δa and Δb are some fixed small angles. We can then replace the distribution $\rho(\lambda_1, \lambda_2, \omega_1, \omega_2)$ by $\rho(\lambda_1, \lambda_2, a, b)$, and relation (E.4) by:

$$\langle AB \rangle \simeq \Delta a \Delta b \int d\lambda_1 \int d\lambda_2 \rho(\lambda_1, \lambda_2, a, b) A(a, \lambda_1) B(b, \lambda_2) \quad (\text{E.9})$$

At this stage, it becomes clear that the model gives much more flexibility than above to reproduce arbitrary a and b dependences; this is because we are indeed free to choose the positive distribution $\rho(\Lambda)$ as we wish, in particular its ω_1 and ω_2 dependences which, in turn, introduces an a and b dependence of ρ into (E.9).

To see mathematically why, we can simplify the model even more, while preserving the possibility of reproducing an arbitrary dependence on the settings. Assume for instance that the domains $D_1^\pm$ of variable λ_1 where $A(a, \lambda_1) = \pm 1$ are independent of a , and, similarly, for the other particle. Then:

$$\langle AB \rangle \simeq \Delta a \Delta b [I_{+,+} + I_{-,-} - I_{+,-} - I_{-,+}] \quad (\text{E.10})$$

with:

$$I_{\pm,\pm} = \int_{D_1^\pm} d\lambda_1 \int_{D_2^\pm} d\lambda_2 \rho(\lambda_1, \lambda_2, a, b) \quad (\text{E.11})$$

The four integrals $I_{\pm,\pm}$ correspond to the probabilities of the four possible measurement results $(\pm, \pm)$, which in quantum mechanics have the expressions given in (4.3) and (4.4). By arbitrarily choosing the ω_1 and ω_2 dependence of ρ , and therefore the a and b dependence of $\rho(\lambda_1, \lambda_2, a, b)$, we can obtain any dependence of the four probabilities on the experimental settings, while remaining within local realism. For instance, we can choose a distribution ρ that is constant (independent of λ_1 and λ_2) within the four domains $D_{1,2}^\pm$, and select for ρ a value that is proportional to the target probability, to obtain the desired dependence on a and b . In this way, we reproduce the predictions of any arbitrary theory¹, provided of course it gives positive probabilities for all values of a and b .

Physically, what we have done is merely to assume that, for each value of the settings a and b , a narrow selection occurs for the detected particles; each time the settings are changed, a different class of pairs is detected; it is then sufficient to attribute arbitrary properties to these classes in order to reproduce all possible variations as functions of a and b , including those of quantum mechanics. Nevertheless, we should remark that what is reproduced is only these a and b variations, and not the absolute value of coincidence rates. There is a compromise between accuracy and detection efficiency: the smaller the values of Δa and Δb , the better (E.9) can approximate any variation, but the price to pay is that more and more pairs go undetected.

¹ A few examples of setting dependent pair selection processes that lead to artificial violations of the BCHSH inequalities are given in [153]; in its §3-3 it shows that the selection can lead to a violation of the Cirelson bound and, in §3-4, to the maximal mathematically possible violation, even with a pair of uncorrelated spins.

The conclusion is that the validity of the Bell inequalities is strongly related to assuming that all pairs of a well-defined and a, b independent sample are detected. If, each time one changes the measurement settings one changes the category of detected pairs, no Bell-type limit may exist for the dependence of the correlation rates with respect to the settings a and b .

Appendix F

Impossibility of superluminal communication

F.1 Introduction

In EPR schemes, applying the reduction postulate projects the second particle instantaneously onto an eigenstate corresponding to the same quantization axis as the first measurement. If it were possible to determine this state completely, superluminal communication would become accessible: from this state, the second experimenter could calculate the direction of the quantization axis to which it corresponds, and rapidly know what direction was chosen by the first experimenter¹, even if the experimenters are in two different and remote galaxies. This, obviously, could be used as a sort of telegraph, completely free of any relativistic minimum delay (proportional to the distance covered) for the transmission of information. Nevertheless, we have seen in §7.2.1 that it is impossible to obtain a complete determination of a quantum state from a single realization of this state. Such a realization allows only one single measurement, which (almost always) perturbs the state, so that a second measurement on the same state is not feasible; there is not, and by far, sufficient information in the first measurement for a full determination of the quantum state – see discussion in §7.2. This telegraph would therefore not function.

If a single particle is not sufficient for Bob to get a message, could he use more particles? Suppose for a moment that a perfect “cloning” of quantum states could be performed – more precisely the reproduction with many particles of the unknown state of a single particle². Applying the cloning process to the second particle of an

¹ What is envisaged here is communication through the choice of the settings of the measurement apparatuses; this makes sense since the settings are chosen at will by the experimenters. On the other hand, the results of the experiments are not controlled, but random, so that they cannot be directly used as signals.

² The “cloning” operation is not to be confused with the preparation of a series of particles into a same quantum state that is chosen by the experimentalist. The latter operation can be realized by sending many spin 1/2 half particles through the same Stern–Gerlach magnet, or many photons through the same polarizing filter. What is theoretically impossible is to perfectly duplicate an initially unknown (and arbitrary) state.

EPR pair, one could then make a large number of perfect copies of its state; one could then perform a series of measurements on each of these copies, and progressively determine the state in question with arbitrary accuracy. In this way, the possibility for superluminal communication would be restored! But we have also seen in §7.1 that quantum mechanics does not allow for such a perfect reproduction of quantum states [266, 267]; for instance, if one considers using stimulated emission in order to clone the state of polarization of one single photon into many copies, the presence of spontaneous emission introduces noise in the process and prevents perfect copying.

Even if it is impossible to clone quantum states and to measure single states accurately, this does not provide an obvious answer to the general question: by using only the information that is available in one single measurement in each region of space, is it possible to make use of the instantaneous reduction of the state vector for superluminal communication? After all, it is possible to repeat the experiment many times with many independent pairs of correlated particles, and to try to extract some information from the statistical properties of the results of all measurements. The EPR correlations are very special and exhibit such completely unexpected properties (e.g. violations of the Bell inequalities)! Why not imagine that, by using or generalizing EPR schemes (more than two systems, delocalized systems, etc.), one could invent schemes where superluminal communication becomes possible? Here we show why such schemes do not exist; we will give the general impossibility proof in the case of two particles (or two regions of space), but the generalization to more systems in several different regions of space is straightforward.

F.2 A first scheme

Suppose that, initially, the two remote observers already possess a collection of pairs of correlated particles, which have propagated from their common source to two remote regions of space before the experiment starts; the first has propagated to region A where Alice does experiments, the second to region B containing Bob's laboratory. Each pair is in an arbitrary state of quantum entanglement; we describe its state before measurement with a density operator $\rho(t_0)$ in a completely general way. When the two particles are very distant, they do not interact, and their Hamiltonian is merely the sum $H_1 + H_2$. Alice then chooses a setting a or, more generally, any local observable $O_A(1)$ to measure. Bob is equally free to choose any local observable $O_B(2)$, and may use as many particles as necessary to measure the frequency of occurrence of each result (i.e. probabilities). The question is whether the second observer can extract some information on the choice of O_A from any statistical property of the observed results.

Equation (10.9) provides the probability of an event where Alice observes result m at time t_1 and Bob result n at time t_2 :

$$\mathcal{P}(m, t_1; n, t_2) = \text{Tr} \{ \hat{P}_{O_B}(n, t_2) \hat{P}_{O_A}(m, t_1) \rho(t_0) \hat{P}_{O_A}(m, t_1) \hat{P}_{O_B}(n, t_2) \} \quad (\text{F.1})$$

where $\hat{P}_{O_A}(m, t_1)$ is the projector onto the eigenstates corresponding to Alice's measurement in the Heisenberg picture, and $\hat{P}_{O_B}(n, t_2)$ the corresponding projector for Bob. But Bob does not have access to the results obtained by Alice, so that the probability of the events that he observes is the sum of (F.1) over m (§4.4.1):

$$\mathcal{P}_B(n, t_2) = \sum_m \mathcal{P}(m, t_1; n, t_2) \quad (\text{F.2})$$

To calculate this sum, we first remark that the two projectors $\hat{P}_{O_A}(m, t_1)$ and $\hat{P}_{O_B}(n, t_2)$ commute: they correspond to operators acting on different particles, and evolve in the Heisenberg picture under the effect of independent Hamiltonians H_1 and H_2 . In the right-hand side of (F.1), consider the operator $\hat{P}_{O_A}(m, t_1)$ appearing just before last in the trace; we can bring it to the last position and then, by circular permutation of the operators inside the trace, put it first. It then sits just before the first $\hat{P}_{O_B}(n, t_2)$, and can be put just after since it commutes with that operator; but, since the square of projector $\hat{P}_{O_A}(m, t_1)$ is equal to the projector itself, eventually the result of our operation is just to suppress one of the operators $\hat{P}_{O_A}(m, t_1)$ in (F.1). We now have to perform the summation over m . Since:

$$\sum_m \hat{P}_{O_A}(m, t_1) = 1 \quad (\text{F.3})$$

we obtain at the end:

$$\mathcal{P}_B(n, t_2) = \text{Tr} \{ \hat{P}_{O_B}(n, t_2) \rho(t_0) \hat{P}_{O_B}(n, t_2) \} \quad (\text{F.4})$$

We then see that, in this probability, any dependence on the choice of operator O_A made by Alice has disappeared. By measuring probabilities locally, Bob obtains no information on the choice of operator made by Alice.

The proof can easily be generalized to a situation where Alice and Bob perform several measurements at different times³. We have therefore shown that, in a very

³ The proof proceeds by the same method as above. One takes the last projector corresponding to Alice's measurements, one pushes it to the end of the product in the trace, then at the beginning, and finally groups it with the other occurrence of the same operator. The summation of this residual projector over the last result obtained by Alice then gives 1. Then one proceeds in the same way for the second projector associated to Alice's measurements, etc. When all the m summations are performed, one obtains a probability that is completely independent of the choice of Alice's operators.

Of course, the proof is meaningful only if the time intervals between Alice's measurement remain shorter than the time of propagation to Bob's site at the speed of light. Otherwise the particles would have the time to propagate from one region to the other, and the comparison with relativity is irrelevant.

general way, the second observer receives exactly the same information, in a way that is completely independent of the decisions made by the first observer; even the fact that Alice makes experiments or no measurement at all is impossible to detect in Bob's laboratory. No faster-than-light-communication is therefore possible by this method.

F.3 Generalization

One could object that it is not necessary that particle 1 is located in region A of space and particle 2 in region B : if each of the particles is at the same time in both regions, can we not imagine cases where the probabilities observed by Bob depend on the choices made by Alice? In fact, this situation is not so different from the previous case since, again, all operators corresponding to the measurements performed in region A commute with all those associated with measurements performed in region B . In field theory, this is a consequence of the commutation of field operators at points of space-time that correspond to space like intervals in relativity. In a more elementary theory, if Alice can detect both particles, we can write the corresponding measurement operator as a sum of operators acting on the two particles:

$$O_{Alice} = P_A(1)O_A(1)P_A(1) + P_A(2)O_A(2)P_A(2) \quad (F.5)$$

where $P_A(1, 2)$ is the projector over all states of particle 1, 2 localized in region A ; the first term corresponds to the case where Alice detects particle 1, the second to the case where she detects particle 2. Similarly, the measurement performed by Bob corresponds to the operator:

$$O_{Bob} = P_B(1)O_B(1)P_B(1) + P_B(2)O_B(2)P_B(2) \quad (F.6)$$

It is then easy to check that O_{Alice} and O_{Bob} commute. The reason is that, in the products between these two operators, the terms in $P_A P_B$ vanish when the two projectors act on the same particle, and the two remaining terms have the form:

$$P_A(1)O_A(1)P_A(1) \times P_B(2)O_B(2)P_B(2) \quad (F.7)$$

added to the similar term where the particle numbers are exchanged; this is because, if each of the two operators performs a measurement on a particle, either particle 1 is in region A and particle 2 in region B , or the reverse, but the two particles are never in the same region. One then immediately sees that it is possible to reverse the order of the two factors in (F.7) without changing the result, so that the two operators O_{Alice} and O_{Bob} commute. This result allows us to use the same proof as above and to obtain the same conclusion: the information available in one region of space is completely independent of the sort of measurements that are performed in the other region. No message can then be sent, and quantum mechanics is not contradictory with relativity!

Appendix G

Quantum measurements at different times

In this appendix, we discuss the probabilities associated with several measurements of the same quantum system at different times. We first give a proof of relation (10.7), which we have used in §10.1.2 without justification, in order to derive the Wigner rule (10.9) from the postulate of state vector reduction (§1.2.2). Conversely, this postulate can also be considered as a consequence of a generalized Born rule, dealing with several measurements performed at different times. In a second part of the appendix, we discuss this derivation, with a reasoning that involves the coupling of the system with the environment provided by the measurement apparatuses.

G.1 Wigner formula

To understand how (10.7) can be obtained, let us first calculate the probability that the first measurement provides result m at time t_1 . The usual Born rule (1.6) indicates that this probability is given by the square norm of $|\Psi_m(t_1)\rangle$ defined in (10.3):

$$\mathcal{P}_1(m, t_1) = \langle \Psi_m(t_1) | \Psi_m(t_1) \rangle \quad (\text{G.1})$$

Equivalently, we can express this probability as a trace:

$$\mathcal{P}_1(m, t_1) = \text{Tr} \{ |\Psi_m(t_1)\rangle \langle \Psi_m(t_1)| \} \quad (\text{G.2})$$

(the definition of the trace in any orthonormal basis immediately shows that these two expressions are equal). We now calculate the conditional probability $\mathcal{Q}_{N/M}$ that, if result m has been obtained, the second measurement will provide result n . After the first measurement, the state vector reduction postulate – relation (1.8)– indicates that the normalized state vector is:

$$|\Psi'_m(t_1)\rangle = \frac{1}{\sqrt{\langle \Psi_m(t_1) | \Psi_m(t_1) \rangle}} |\Psi_m(t_1)\rangle \quad (\text{G.3})$$

which, after evolution between time t_1 and time t_2 , is nothing but the ket $|\Psi_m(t_2)\rangle$ given by (10.4) and (10.5), divided by $\sqrt{\langle\Psi_m(t_1)|\Psi_m(t_1)\rangle}$. The same reasoning as above then gives the conditional probability as:

$$\mathcal{Q}_{N/M}(m, t_1; n, t_2) = \frac{1}{\langle\Psi_m(t_1)|\Psi_m(t_1)\rangle} \langle\Psi_{m,n}(t_2)|\Psi_{m,n}(t_2)\rangle \quad (\text{G.4})$$

where $|\Psi_{m,n}(t_2)\rangle$ has been defined in (10.6). From this, we obtain the probability of obtaining the sequence m and n :

$$\mathcal{P}_1(m, t_1; n, t_2) = \mathcal{P}_1(m, t_1) \times \mathcal{Q}_{N/M}(m, t_1; n, t_2) = \langle\Psi_{m,n}(t_2)|\Psi_{m,n}(t_2)\rangle \quad (\text{G.5})$$

(two factors $\langle\Psi_m(t_1)|\Psi_m(t_1)\rangle$ cancel each other in the numerator and the denominator). Equivalently, one can express the right-hand side of (G.5) as a trace:

$$\mathcal{P}_1(m, t_1; n, t_2) = \text{Tr} \{ |\Psi_{m,n}(t_2)\rangle \langle\Psi_{m,n}(t_2)| \} \quad (\text{G.6})$$

By recurrence, (G.5) leads to (10.7).

To prove (10.9), we introduce the projection operator in the Heisenberg picture:

$$\hat{P}_M(m, t_1) = U^\dagger(t_1, t_0) P_M(m) U(t_1, t_0) \quad (\text{G.7})$$

where $U(t_1, t_0)$ is the unitary evolution operator between times t_0 and t_1 (§11.1.4). From (10.3) and (10.1), we obtain, since $U(t_1, t_0)U^\dagger(t_1, t_0) = 1$:

$$|\Psi_m(t_1)\rangle = P_M(m) U(t_1, t_0) |\Psi(t_0)\rangle = U(t_1, t_0) \hat{P}_M(m, t_1) |\Psi(t_0)\rangle$$

and relation (G.2) gives, since $P_M(m)$ is Hermitian:

$$\begin{aligned} \mathcal{P}_1(m, t_1) &= \text{Tr} \left\{ U(t_1, t_0) \hat{P}_M(m, t_1) |\Psi(t_0)\rangle \langle\Psi(t_0)| \hat{P}_M(m, t_1) U^\dagger(t_1, t_0) \right\} \\ &= \text{Tr} \left\{ \hat{P}_M(m, t_1) |\Psi(t_0)\rangle \langle\Psi(t_0)| \hat{P}_M(m, t_1) \right\} \\ &= \text{Tr} \left\{ \hat{P}_M(m, t_1) \rho(t_0) \hat{P}_M(m, t_1) \right\} \end{aligned} \quad (\text{G.8})$$

(we have used the circular permutation of operators under the trace, which makes the product of two operators disappear).

A similar calculation can also be made from (G.5). The ket $|\Psi_{m,n}(t_2)\rangle$ is given by:

$$\begin{aligned} |\Psi_{m,n}(t_2)\rangle &= P_N(n) U(t_2, t_1) |\Psi_m(t_1)\rangle \\ &= P_N(n) U(t_2, t_1) U(t_1, t_0) U^\dagger(t_1, t_0) P_M(m) U(t_1, t_0) |\Psi(t_0)\rangle \end{aligned} \quad (\text{G.9})$$

where we have inserted the relation $U(t_1, t_0)U^\dagger(t_1, t_0) = 1$ in order to make the Heisenberg projector $\hat{P}_M(m, t_1)$ appear again; we then simplify the product of operators $U(t_2, t_1)U(t_1, t_0)$ into $U(t_2, t_0)$ and use the unitarity of U to write:

$$|\Psi_{m,n}(t_2)\rangle = U(t_2, t_0) \hat{P}_N(n, t_2) \hat{P}_M(m, t_1) |\Psi(t_0)\rangle \quad (\text{G.10})$$

Finally, using (G.6) and after circular permutation of operators under the trace (which make two U operators disappear), we obtain:

$$\mathcal{P}_1(m, t_1; n, t_2) = \text{Tr} \left\{ \hat{P}_N(n, t_2) \hat{P}_M(m, t_1) |\Psi(t_0)\rangle \langle \Psi(t_0)| \hat{P}_M(m, t_1) \hat{P}_N(n, t_2) \right\} \quad (\text{G.11})$$

which leads to (10.9). By the same method, it is straightforward to generalize this formula to more than two measurements, with additional projectors in the reverse order on both sides of $\rho(t_0)$. By linearity, this result also applies to situations where the density operator $\rho(t_0)$ is not a projector (a pure state) as in (10.8), but a statistical mixture.

G.2 Generalized Born rule

We now take a different approach by taking into account the entanglement of the measured system S with the measurement apparatuses¹. A measurement associated with operator M is performed at time t_1 , another measurement associated with operator N at time t_2 , etc. (in the calculation, we assume that two measurements are performed, but the generalization to an arbitrary number of measurements is straightforward).

Initially, at time t_0 , the system S is in state $|\Psi^S(t_0)\rangle$; the two measurement apparatuses M and N , which have not yet operated, are in states $|\Psi^M(t_0)\rangle$ and $|\Psi^N(t_0)\rangle$, and the state vector $|\overline{\Psi}(t_0)\rangle$ of the whole system including these apparatuses is the (tensor) product:

$$|\overline{\Psi}(t_0)\rangle = |\Psi^S(t_0)\rangle \otimes |\Psi^M(t_0)\rangle \otimes |\Psi^N(t_0)\rangle \quad (\text{G.12})$$

(all these states are normalized). We assume that the three systems evolve independently (without any interaction), except between times t_1 and t'_1 where S interacts with the first measurement apparatus M , and between times t_2 and t'_2 , where it interacts with the second measurement apparatus N . The two measurement apparatuses are macroscopic; each includes a pointer that indicates, after the measurement, the result that has been obtained. They never interact with each other, but only with system S .

Between time t_0 and time t_1 the state of S evolves from $|\Psi^S(t_0)\rangle$ to $|\Psi^S(t_1)\rangle$ according to the Schrödinger equation (10.1) and, similarly, the state vectors of the measurement apparatuses become $|\Psi^M(t_1)\rangle$ and $|\Psi^N(t_1)\rangle$.

We now consider the effect of the first measurement. As in §§1.2.2.a and 10.1.2, we call $P_M(m)$ the projectors over the possible eigenvectors of operator M , with

¹ This approach is necessary if one chooses the Everett interpretation (§ 10.11).

eigenvalues $m = m_1, m_2, \dots, m_i, \dots$, and expand $|\Psi^S(t_1)\rangle$ over these eigenvectors; with our notation, relations (10.2) and (10.3) become:

$$|\Psi^S(t_1)\rangle = \sum_m |\Psi_m^S(t_1)\rangle \quad (\text{G.13})$$

and:

$$|\Psi_m^S(t_1)\rangle = P_M(m) |\Psi^S(t_1)\rangle \quad (\text{G.14})$$

If it turns out that $|\Psi^S(t_1)\rangle$ is an eigenvector of M with eigenvalue m_i , then only one term $m = m_i$ occurs in the summation; the result of the first measurement is certain. At time t'_1 just after this measurement, the first apparatus will reach a well-defined normalized state $|\Psi_{m_i}^M(t'_1)\rangle$ where its pointer indicates the result. Therefore, in this particular case, just after the first measurement the state of the whole system will be:

$$|\overline{\Psi}(t'_1)\rangle = |\Psi_{m_i}^S(t'_1)\rangle \otimes |\Psi_{m_i}^M(t'_1)\rangle \otimes |\Psi^N(t'_1)\rangle \quad (\text{G.15})$$

In general, for any $|\Psi^S(t_1)\rangle$, the linearity of the Schrödinger equation implies that the state of the whole system just after the first measurement is:

$$|\overline{\Psi}(t'_1)\rangle = \sum_m |\Psi_m^S(t'_1)\rangle \otimes |\Psi_m^M(t'_1)\rangle \otimes |\Psi^N(t'_1)\rangle \quad (\text{G.16})$$

The first measurement apparatus has then reached a state that depends on the eigenvalue m , but not the second, which has not yet interacted with S . All the evolutions of the individual state vectors correspond to unitary time evolutions, which conserve the norm; nevertheless, expansion (G.13) contains kets $|\Psi_m^S(t_1)\rangle$ having a norm that is in general smaller than that of $|\Psi^S(t_1)\rangle$.

For the second measurement, we repeat the same calculation. Between time t'_1 and t_2 , for each value of m in (G.16) every term in the product evolves independently (unitary evolution), and $|\overline{\Psi}(t_2)\rangle$ is obtained by replacing t'_1 by t_2 in (G.16). The analogous expressions of (G.13) and (G.14) are now:

$$|\Psi_m^S(t_2)\rangle = \sum_n |\Psi_{m,n}^S(t_2)\rangle \quad (\text{G.17})$$

with:

$$|\Psi_{m,n}^S(t_2)\rangle = P_N(n) |\Psi_m^S(t_2)\rangle \quad (\text{G.18})$$

The second measurement correlates the state of the second apparatus to that of system S . Finally, the state of the whole system at time t'_2 after the second measurements is:

$$|\bar{\Psi}(t'_2)\rangle = \sum_m \sum_n |\Psi_{m,n}^S(t'_2)\rangle \otimes |\Psi_m^M(t'_2)\rangle \otimes |\Psi_n^N(t'_2)\rangle \quad (\text{G.19})$$

This state vector contains a coherent superposition of various components associated with all possible pairs (m, n) of results of measurements; in each of these components, the state of the measurement apparatuses have registered the results.

At this stage, we can introduce a “generalized Born rule” and postulate that the probability of obtaining results (m, n) is given by the square of the norm of each component. We have already remarked that, as long as the three sub-systems do not interact, the norm of each factor in one component remains constant. Actually, the only changes of norm occur during the interactions (the measurements) and are contained in expansions (G.13) and (G.17). Since the states of the measurement apparatuses in (G.19) are normalized, the probability $\mathcal{P}_{m,n}$ of any result (m, n) is then:

$$\mathcal{P}_{m,n} = \langle \Psi_{m,n}^S(t'_2) | \Psi_{m,n}^S(t'_2) \rangle \quad (\text{G.20})$$

which is nothing but the norm of the ket:

$$\begin{aligned} & P_N(n)U(t_2, t'_1) \left| \Psi_m^S(t'_1) \right\rangle \\ &= P_N(n)U(t_2, t'_1)P_M(m)U(t_1, t_0) \left| \Psi^S(t_0) \right\rangle \end{aligned} \quad (\text{G.21})$$

The state vectors associated to the measurement apparatuses have completely disappeared from this result. It has a simple interpretation in terms of system S only, if one reads the second line from the left to the right: system S was initially in state $|\Psi^S(t_0)\rangle$, evolved freely from time t_0 to time t_1 , then underwent a projection over the eigenstates corresponding to the result of measurement m , then evolved freely again from time t'_1 to time t_2 , to be projected a second time over the eigenstates corresponding to the result of measurement n . Moreover, (G.20) is equivalent to relation (G.5), which was obtained by applying the postulate of state vector reduction. From the probabilities of combined (m, n) events, one can derive the conditional probabilities of other events, using the usual rules of statistics. The final conclusion is that both methods provide exactly the same results for all probabilities: one may either involve the measurement apparatuses and use a generalized Born rule, or consider the measured system S only and apply the state vector reduction postulate (§1.2.2.a). It is a matter of taste to choose one method as a postulate and derive the other as a consequence, or the reverse.

Appendix H

Manipulating and preparing additional variables

Introducing the hydrodynamic equations (Madelung [416]) associated with the evolution of the wave function, and using them to guide the evolution of the additional variables (positions of particles), is a natural idea. In the dynamics of fluids, hydrodynamic equations can be obtained by taking averages of microscopic quantities over positions and velocities of point-like particles; for instance, the Navier–Stokes macroscopic equations can be derived from the Boltzmann transport equation by appropriate microscopic averages (Chapman–Enskog method); conversely, the hydrodynamic variables will influence the motion of individual particles. Moreover, there is some analogy between the guiding term and the force term in a Landau type kinetic equation, where each particle is subject to an average force proportional to the gradient of the density of the others. Nevertheless, here we are dealing with a single particle, so that the guiding term cannot be associated with interactions between particles. Moreover, we also know from the beginning that rather unusual properties must be contained in the guiding equations, at least if we wish to exactly reproduce the predictions of usual quantum mechanics: the Bell theorem states that the additional variables have to evolve non-locally in ordinary three-dimensional space (they evolve locally only in the configuration space of the system, exactly as for the state vector). In other words, in real space the additional variables must be able to influence each other at an arbitrary distance. Indeed, in the Bohmian equation of motion of the additional variables, the velocity of a particle contains an explicit dependence on its own position, as expected, but also a dependence on the positions of all the other particles (assuming that the particles are entangled). This is not a problem in itself: as mentioned in the main text, one can consider that making non-locality completely explicit in the equations is actually an advantage of Bohmian mechanics.

But one also has to be careful when this non-local term is included in the equations of motion: since relativity is based on the idea that it is totally impossible to send a message at a velocity exceeding the velocity of light, one must avoid

features in the theory that would create conflicts with this principle. We must distinguish two cases, depending whether we consider influences on the additional variables that are direct (one modifies them “by hand”, in a completely arbitrary way, as for instance the position of a billiard ball), or indirect (applying external fields changes the Hamiltonian of the system, therefore modifies the evolution of the wave function so that, in turn, the evolution of the additional variables is affected). In the latter case, one can check that the non-local Bohmian term creates no problem: it cannot be used to transmit instantaneous information through the additional variables. This is a general result, which holds simply because the statistical predictions of Bohmian theory are equivalent to usual quantum mechanics, which itself does not allow superluminal communication (§4.4 and Appendix F). But assume for instance that we could manipulate directly the additional variable attached to a particle belonging to an EPR correlated pair, in a completely arbitrary way (even at a microscopic scale) and without changing the wave function. Then, the “quantum velocity term” acting on the additional variables of the other particle would instantaneously be affected, and so would be its subsequent position in space. Since that particle may be in principle at an arbitrary distance, one could use this property to send messages at a velocity exceeding the velocity of light. The conclusion is that such manipulations should be considered as impossible: the only possible source of evolution of the additional variables has to be the wave-function-dependent term, without any possibility for direct human action.

If the additional variables cannot be directly manipulated at a microscopic scale, can we then assume that it is possible somehow to filter them in a range of values, as one does for the state vector when the Oz component is filtered in a Stern–Gerlach apparatus? Suppose for instance that we could, for a particle in an eigenstate of the Oz component of its spin, select the values of the additional variable that will correspond to a result $+1$ in a future measurement of the Ox component; were such a selection possible with the help of any physical device, the theory with additional variables would obviously no longer be completely equivalent to standard quantum mechanics, introducing determinism where this theory forbids it¹. Moreover, Valentini has shown [421] that, if the initial distribution of the Bohmian positions differs from the usual “quantum equilibrium” distribution (§10.6.1.a), faster than light communication becomes possible. If one could somehow prepare by hand a distribution of the Bohmian positions that differs from $|\Psi(\mathbf{Q}_1, \mathbf{Q}_2, \dots)|^2$, a narrower distribution for instance, then contradictions with relativity would

¹ Within orthodox theory, if a spin $1/2$ particle is initially selected into the $+1$ spin state by an Oz oriented Stern–Gerlach apparatus, it becomes completely impossible to make any prediction on the deviation observed later in an Ox oriented Stern–Gerlach apparatus.

occur. This is the reason why one generally considers that such preparations are impossible².

To summarize, if one provides the additional variable theories with features that make them equivalent to orthodox theory, it is necessary to assume that the additional variables can neither be manipulated directly nor filtered, as opposed to the state vector. The additional variables describe an objective reality, but at a different level from the reality of the field of the wave function, since only the latter can be influenced directly by human decisions. Additional variables are indeed easily visible (the results of the experiments) but not controllable, while wave functions have the complementary properties. At the end, we have two levels of reality, one for experimentally controllable classical fields (the wave functions), and one for positions which can be observed but not controlled.

Moreover, when empty waves appear during measurement processes, classical wave functions in turn split into two sub-levels of reality: part of them continues to play an effective role as a pilot of the position of particles, while another part is made of empty waves that play no role at all in determining the result of future measurements. This second part becomes, so to say, virtual. If for instance one considers the wave function of the universe, as in the Everett theory it is split into a fantastic number of orthogonal components (although, here, these components have a different status, with no special relation to memory registers of observers). Among all these components, only one plays a role to pilot the position of the point representing the universe, in a configuration space with enormous dimension. All the others are too far from the real position in this space, and will remain in limbo forever.

In conclusion, the description of physical reality within the de Broglie–Bohm theory is certainly not as direct as in classical theory of particles and fields. It is not entirely free of conceptual difficulties, which have some similarity with those encountered in the standard interpretation of quantum mechanics.

² We should mention that, historically, Bohm and Bud have developed a theory where this possibility is not excluded [7] – see §10.8.1.a. Conceptual revolutions are of course always possible, but for the moment it may seem safer to provide the additional variable theories with features that make them equivalent to orthodox theory.

Appendix I

Correlations in Bohmian theory

I.1 Time correlation functions

In standard quantum mechanics, the calculation of any two-time correlation function has to include the evolution of the system between the two times considered; this evolution is contained in the unitary evolution operator $U(t', t)$, as for instance in relation (10.9). In Bohmian theory, it is important to take into account the effect of the first measurement, which correlates the system under study to a measurement apparatus and creates “empty waves” (§10.6.1.c). Otherwise one obtains contradictions with the standard predictions.

For instance the author of [443] considers a one-dimensional harmonic oscillator that is initially in a stationary state, and studies the correlation function of the position at times t and time t' , in the particular case where $t' - t$ is equal to half the period $2\pi/\omega$ of the oscillator. In standard quantum mechanics, it is easy to show that the corresponding position operators $X(t)$ and $X(t')$ are then opposite, so that the correlation function $\langle X(t)X(t') \rangle$ is equal to $-\langle [X(t)]^2 \rangle$, therefore negative. In Bohmian mechanics, the particle is initially static since the wave function is real. If one ignores the effect of the first measurement, the position of the particle will remain at the same place, which corresponds to a correlation function equal to $\langle [X(t)]^2 \rangle$, therefore positive; one reaches an apparent complete contradiction. But, if one takes into account the effect of the first measurement on the particle, one finds that, just after this measurement, each position of the oscillator becomes correlated with a different Bohmian position of the pointer. Since the wave function is no longer a product, the motions of both systems become correlated: for each position of the pointer, the measured particle takes a different velocity. In practice, just after the measurement, the position distribution of the oscillator becomes a narrow function that starts to oscillate in the potential well, giving an oscillating character to two-time position distribution. The sign difference

with the standard correlation function then disappears and a perfect agreement is obtained.

A similar case is studied in [442], with two independent one dimension harmonic oscillators, initially in the state:

$$|\Psi\rangle = \frac{2}{\sqrt{2}} \left[|1, 0\rangle + |0, 1\rangle \right] \quad (\text{I.1})$$

where $|n, p\rangle$ is the state where the first oscillator has quantum number n and the second quantum number p ; for the sake of simplicity, we assume that the frequencies of the two oscillators are the same, equal to $\omega/2\pi$. Since the stationary wave functions of the harmonic oscillator can be chosen real, the associated wave functions are also real, meaning that none of the Bohmian positions of the two particles moves at all.

The position operators of the two oscillators commute and can be measured, either simultaneously, or with any delay between the measurements (one can assume for instance that the two oscillators are centered at different points of space, so that differentiating between the two positions is easy). Since the two operators commute, the standard quantum calculation of the correlation coefficient of the two positions is simple; one finds that it contains a component oscillating at frequency $\omega/2\pi$. On the other hand, we have seen that the Bohmian positions are static, so that the average of their product over all possible trajectories gives a constant result. It therefore seems again that one reaches a contradiction between the predictions of standard mechanics and those of Bohmian mechanics.

But actually this is not the way correlation functions should be calculated within Bohmian mechanics: again, the effect of the first measurement must be properly taken into account, even if the two measured observables commute, and even if they correspond to independent systems. The first measurement correlates the Bohmian position of the particle with that of a pointer, resulting in “empty waves” and modified dynamics of the system, which then affects the correlation function. The case studied in [442] is actually particularly interesting, since the effect of this first measurement is to switch off (instead of on) non-local effects. The following scenario takes place:

(i) Initially, since the wave function is not a product, the motions of the two particles are subject to complicated non-local effects. For a real wave function such as (I.1), the non-local effects combined with the local effects of the potentials exactly cancel, which results in a very simple motion: the position of none of them moves.

(ii) When particle 1 is measured, the system becomes a three (or more) body problem including the positions of the two particles and that of the pointer. Each Bohmian trajectory selects a pointer position (for instance that corresponding to a

positive measurement in the detection volume), and for each such trajectory the two particles move guided by a wave function that is now a product. The correlations have vanished so that the non-local effects have disappeared; both particles then follow a local Schrödinger evolution.

(iii) After the measurement, both particles oscillate in their potential wells, in a way that exactly reproduces the standard quantum result for the correlation function.

I.2 Two-particle correlations

I.2.1 EPRB experiment

Consider an EPRB experiment (§3.3.1) with two remote particles, the first with wave function $\varphi(\mathbf{r}_1)$ and the second with wave function $\chi(\mathbf{r}_2)$. We assume that each of these particles has a spin $1/2$ and that their spins are entangled in a singlet spin state. In Bohmian theory, it is convenient to write explicit wave functions, so that here we will use a mixed notation with wave functions for orbital variables and spins treated as kets in the Dirac notation. We can then write the wave function/state vector of the two-particle system as:

$$|\Psi\rangle = \varphi(\mathbf{r}_1) \chi(\mathbf{r}_2) \frac{1}{\sqrt{2}} \left[|1: +, 2: -\rangle - |1: -, 2: +\rangle \right] \quad (\text{I.2})$$

Assume now that the first particle enters a magnetic field gradient such as that of a Stern–Gerlach magnet oriented along direction a , which splits the wave function $\varphi(\mathbf{r}_1)$ into two spatially separated components $\varphi'_+(\mathbf{r}_1)$ and $\varphi'_-(\mathbf{r}_1)$. Then $|\Psi\rangle$ becomes:

$$|\Psi'\rangle = \frac{1}{\sqrt{2}} \left[\varphi'_+(\mathbf{r}_1) |1: +\rangle_a |2: -\rangle_a - \varphi'_-(\mathbf{r}_1) |1: -\rangle_a |2: +\rangle_a \right] \chi(\mathbf{r}_2) \quad (\text{I.3})$$

where the indices a in the spin states refer to a quantization direction that has been chosen parallel to a . In such a situation, the Bohmian position of the first particle must be either in wave packet $\varphi'_+(\mathbf{r}_1)$, or in wave packet $\varphi'_-(\mathbf{r}_1)$. The other wave is necessarily an “empty wave”, which does not play any role and may be ignored. One of the components of (I.3) then disappears, so that the spin state of the second particle has been projected onto the opposite state of that of the first particle. We see that a spatial separation of the wave packet of one of the particles is already sufficient to project both spin states, even before the second particle has entered any magnetic gradient. One therefore obtains in this way a very efficient mechanism for reproducing the standard state vector projection.

1.2.2 Two-photon interferences

An experimental refutation of Bohmian mechanics was published in [444], involving correlations between two photons, each going through different slits of a screen; the refutation was based on a previous theoretical claim of a discrepancy between Bohmian theory and standard quantum mechanics [445]. But, here again, what is disproved is not the full Bohmian theory, but a modified version. The full Bohmian theory is built to reproduce exactly the same results as standard quantum mechanics, provided the same wave function is used in both cases (symmetrical with respect to exchange between the two bosons in this case); such discrepancies in predictions concerning position measurements cannot occur. In the present case, the extra assumption is that the trajectories of the two bosons are always symmetric with respect to the symmetry plane of the experiment. But, in Bohmian theory, the positions of the particles fill the whole volume available in six dimensional configuration space, and within this volume non-symmetric pairs of positions are perfectly accessible.

Appendix J

Models for spontaneous reduction of the state vector

In this appendix, we introduce a few simple models involving modified Schrödinger dynamics with stochasticity, in order to illustrate how such models may lead to an evolution that reproduces the reduction of the state vector during a measurement (emergence of a single eigenvalue during a single realization, with a random value). For the sake of simplicity, we ignore the usual Hamiltonian evolution during the time of measurement, assuming for instance that this time is too short for this evolution to be significant; otherwise, it would be necessary to use the interaction representation with respect to the Hamiltonian, which does not change the calculations much, except that this introduces a time dependence of the operators.

J.1 Single operator

We consider the measurement of some quantum observable associated with an Hermitian operator A ; we look for an equation of evolution containing a state vector reduction process associated with this particular measurement. Since the final eigenvector must vary randomly from one realization to the next, the evolution equation necessarily contains a random component. In our case, it will take the form of a random function of time (as opposed to the GRW theory where the stochasticity is introduced by the discontinuous “hitting processes”, see §10.8.1.b).

J.1.1 Equation of evolution

We assume that the state vector $|\Psi(t)\rangle$ evolves according to:

$$\frac{d}{dt} |\Psi(t)\rangle = -[w(t) - A]^2 |\Psi(t)\rangle \quad (\text{J.1})$$

where $w(t)$ is a real random function of time. In order to simplify the model as much as possible, we may discretize time into small finite intervals Δt , during

which we assume that $w(t)$ remains constant; moreover, we may also assume that the possible values of $w(t)$ belong to a finite discrete ensemble $w_1, w_2, \dots, w_N$. One may then choose a rule to specify the time progression of $w(t)$, and possibly an interpolation rule to make the functions continuous. Another choice would be to assume that $w(t)$ corresponds to a white noise (Wiener process), with no time memory. For the moment, we do not specify the properties of this random function any further.

Equation (J.1) clearly does not conserve the norm of $|\Psi(t)\rangle$, but it is possible to define a normalized state vector $|\Phi(t)\rangle$ by:

$$|\Phi(t)\rangle = \frac{|\Psi(t)\rangle}{\langle\Psi(t)|\Psi(t)\rangle^{1/2}} \quad (\text{J.2})$$

We then have:

$$\begin{aligned} \frac{d}{dt} |\Phi(t)\rangle &= -[w(t) - A]^2 |\Phi(t)\rangle \\ &\quad - \frac{1}{2 \langle\Psi(t)|\Psi(t)\rangle^{3/2}} (-2) \langle\Psi(t)|[w(t) - A]^2 |\Psi(t)\rangle \times |\Psi(t)\rangle \end{aligned} \quad (\text{J.3})$$

or:

$$\frac{d}{dt} |\Phi(t)\rangle = \left\{ -[w(t) - A]^2 + \langle\Psi(t)|[w(t) - A]^2 |\Psi(t)\rangle \right\} |\Phi(t)\rangle \quad (\text{J.4})$$

With this non-linear equation, the norm of $|\Phi(t)\rangle$ does not vary in time, whatever choice is made for the random function $w(t)$.

J.1.2 Solution of the equation

We denote $|a_n\rangle$ the eigenvectors of A with eigenvalues¹ a_n ; we may then expand $|\Psi(t)\rangle$ as:

$$|\Psi(t)\rangle = \sum_n x_n(t) |a_n\rangle \quad (\text{J.5})$$

We then have:

$$\frac{d}{dt} x_n(t) = -[w(t) - A]^2 x_n(t) \quad (\text{J.6})$$

so that:

$$x_n(t) = e^{-\int_0^t dt' [w(t') - a_n]^2} x_n^0 \quad (\text{J.7})$$

¹ In case of degeneracy, several consecutive values of a_n are equal, but they correspond to different (orthogonal) eigenstates.

where x_n^0 is the initial value of $x_n(t)$ at time $t = 0$. Equivalently, we may also write:

$$|\Psi(t)\rangle = e^{-\int_0^t dt' [w(t') - A]^2} |\Psi(0)\rangle \quad (\text{J.8})$$

According to (J.7), all components of $|\Psi(t)\rangle$ constantly decrease in time, except if $w(t) = a_n$; in this particular case, the components associated with one eigenvalue of A remain constant, as long as the equality is obeyed.

J.1.3 CSL probability rule

We now assume that the probability of each realization of $w(t)$ (the number of these realizations is finite if we take the simple assumptions mentioned above) is proportional to the square of the norm of the value of $|\Psi(t)\rangle$ obtained from (J.1):

$$\mathcal{P}(w_1, w_2, \dots, w_N) = c_N \langle \Psi(t) | \Psi(t) \rangle^{1/2} \quad (\text{J.9})$$

where $t = N\Delta t$, and where $w_1, w_2, \dots, w_N$ are the selected values for $w(t)$; the normalization coefficient c_N is obtained by writing that the sum of probabilities associated with all realizations is 1. Relation (J.9) may be called the CSL probability rule. From this condition, the Bayes theorem provides the probability that, if $w(t)$ has a given value at some time, the function will jump to any other value in the next discrete time interval.

Among all possible realizations of the random functions $w(t)$, the CSL probability rule strongly favors a small sub-ensemble, that containing functions conserving a large norm for $|\Psi(t)\rangle$, in other words functions that remain constantly equal (or almost equal) to one of the eigenvalues a_n . All other possibilities, even if there are very many of them, are assumed to be very unlikely. For each realization, the mechanism obtained in this way breaks the symmetry between all eigenvalues: the same random function cannot remain very close all the time to more than one eigenvalue a_n ; it has to make a choice among all of them. The net result is that, after some time, one always ends up with a state vector that differs very little from an eigenvector of A . Nevertheless, depending on the random function $w(t)$, which may be different for each realization of the experiment, a different eigenvalue is obtained each time. This is precisely the behavior that is needed to reproduce state vector reduction.

Remarks:

(i) this model does not correspond to a hidden variable theory *stricto sensu*, but is relatively similar; what is added to standard mechanics is a random function playing a role in the dynamics of the state vector, not a hidden variable giving directly the

measured observable, the position for instance. One may call it a “hidden function model”.

(ii) the CSL probability rule gives the probability of a particular realization of a random function, but not of a particular value of $|\Psi(t)\rangle$, since several different random functions may lead to the same state vector.

J.2 Several operators

To measure the position of a particle, one may imagine that many detectors are located at different places in space, each signalling the possible presence of the particle within its spatial range. This amounts to the simultaneous measurement of a large number of observables, all commuting with each other. We therefore generalize the preceding model to a case where several commuting operators A_k play a role in the dynamics of a state vector.

J.2.1 Equation of evolution

We now postulate the equation of evolution:

$$\frac{d}{dt} |\Psi(t)\rangle = - \sum_k [w_k(t) - A_k]^2 |\Psi(t)\rangle \quad (\text{J.10})$$

which contains a series of random real functions $w_k(t)$ and a series of commuting operators A_k . For instance, these operators are assumed to be diagonal in the position representation, and their effect is to multiply the wave function by a given function $\varphi_k(\mathbf{r})$. These functions are, for instance, Gaussian functions with width $\alpha^{-1/2}$ centered on a regular lattice in all space, labeled by an index k . The A_k then all commute, but their product for two different values of k is not necessarily 0 since the $\varphi_k(\mathbf{r})$ have mutual overlap.

We now introduce another, much finer, lattice, made of cubic “cells” having dimensions significantly smaller than $\alpha^{-1/2}$, which are labeled with index q . Within each of these cells, each function $\varphi_k(\mathbf{r})$ remains practically constant, so that in this volume the action of A_k may be approximated by a multiplication by a constant, the value φ_k^q of $\varphi_k(\mathbf{r})$ at the center of the cell. We then expand the state vector onto its components in the cells²:

$$|\Psi(0)\rangle = \sum_q |\bar{\Psi}_q^0\rangle \quad (\text{J.11})$$

² By definition, the component $|\bar{\Psi}_q^0\rangle$ has the same wave function as $|\Psi(0)\rangle$ within cell q , but is 0 outside of this cell.

We then have:

$$A_k \left| \overline{\Psi}_q^0 \right\rangle \simeq \varphi_k^q \left| \overline{\Psi}_q^0 \right\rangle \quad (\text{J.12})$$

which means that $\left| \overline{\Psi}_q^0 \right\rangle$ is almost an eigenvector of A_k . Under these conditions:

$$\left| \Psi(t) \right\rangle \simeq \sum_q \left[e^{-\int_0^t dt' \sum_k [w_k(t') - \varphi_k^q]^2} \right] \left| \overline{\Psi}_q^0 \right\rangle \quad (\text{J.13})$$

We can also write the evolution equation of the normalized state vector $\left| \Phi(t) \right\rangle$ defined in (J.2); we then obtain:

$$\frac{d}{dt} \left| \Phi(t) \right\rangle = \left\{ - \sum_k [w_k(t) - A]^2 + \langle \Phi(t) | \sum_k [w_k(t) - A]^2 | \Phi(t) \rangle \right\} \left| \Phi(t) \right\rangle \quad (\text{J.14})$$

J.2.2 Spontaneous localization of the state vector

The situation is then rather similar to that occurring for one operator A . Equation (J.13) shows that the component associated with a particular cell q can keep a large norm only if each random function $w_k(t)$ remains very close to value φ_k^q during all time interval $[0, t]$. This means that the function $w_k(t)$ corresponding to the point k that is closest to cell q must take a significant value; as for the other random functions $w_{k'}(t)$ with $k' \neq k$, they must be much smaller, since $\varphi_{k'}^q$ takes an exponentially smaller and smaller value when point k' is far from the considered cell. As a consequence, the random functions $w_k(t)$ may select a particular cell q , or even an ensemble of close cells if they are small and if the time is not too large, but certainly not remote cells at the same time. Finally, if one postulates the CSL probability rule as above, one favors a very particular ensemble of functions, those for which the wave function becomes localized in close cells, or even a single cell after some time. This realizes the equivalent of a state vector reduction in an arbitrary and random small region of space.

Appendix K

Consistent families of histories

This appendix provides a discussion of the consistency condition (10.39) and of the construction of consistent families of histories. First, we should mention that other conditions have been proposed and used in the literature; in the initial article on histories [17], a weaker condition involving only the cancellation of the real part of (10.39) was introduced. For simplicity, here we limit ourselves to the stronger condition (10.39), which is a sufficient but not necessary condition to the weaker form; it turns out that, as noted in [519], it seems more useful in this context to introduce selectivity than generality in the definition of consistent histories.

At first sight, a natural question that comes to mind is whether or not it is easy, or even possible at all, to fulfil exactly the large number of conditions contained in (10.39). Actually, it has been proposed by Gell-Mann and Hartle to give a fundamental role to families that satisfy consistency conditions in only an approximate way [510], but here we leave aside this possibility and consider only exact consistency conditions. Let us assume for instance that the system under study is a particle propagating in free space; the various projectors may then define ranges of positions for the particle, playing a role similar to diaphragms or spatial filters in optics that confine an optical beam in the transverse direction. Then the consistency condition will appear as similar to a non-interference condition for the Huyghens wavelets that are radiated by the inner surface of each diaphragm. But we know that diffraction is unavoidable in the propagation of light; even if it can be a very small effect when the wavelength is sufficiently short and the diaphragms sufficiently broad, it is never strictly 0. Can we then satisfy the non-interference conditions exactly? The answer is not obvious. It turns out to be yes, but it is necessary to exploit the enormous flexibility that we have in the choice of subspaces and projectors in a large space of states, and not to limit ourselves to projectors over well-defined positions only. To understand why, we now

briefly sketch one possible systematic method to construct consistent families of histories.

The simplest method is to guide the construction on the structure of (10.39), and to introduce the eigenstates $|\varphi_n^0\rangle$ of the density operator $\rho(t_0)$ (an Hermitian operator can always be diagonalized); let us then define the operators $\hat{P}_{1,j_1}(t_1)$ as:

$$\hat{P}_{1,n}(t_1) = |\varphi_n^0\rangle\langle\varphi_n^0| \quad (\text{K.1})$$

which is equivalent to assuming that their Schrödinger counterparts $P_{1,j}$ are the projectors over the states that have evolved from the $|\varphi_n^0\rangle$ from time t_0 to time t_1 . Because $\rho(t_0)$ is of course diagonal in its own basis, this choice already ensures the presence of a factor $\delta_{j_1,j_1'}$ in the right-hand side of (10.39). Now, we can also assume that the P_{2,j_2} are defined as the projectors over the states that have evolved from the $|\varphi_n^0\rangle$ from time t_0 to time t_2 , so that a relation similar to (K.1) is obtained again; this will ensure, not only the presence of factors $\delta_{j_2,j_2'}$ in the right hand side of (10.39), but actually also the appearance of a delta function δ_{j_1,j_2} . The procedure can be repeated as many times as needed, and in this way a consistent family is built.

It is nevertheless a very special family, for several reasons. The first is that each projector corresponds to a subspace of dimension 1 only, which corresponds to histories that are “maximally accurate”; the second is that most histories of the family have zero probability: in fact, only those with $j_1 = j_2 = j_3 = \dots$ are possible, which means that the only randomness occurs at the initial time t_1 , and that all subspaces at later times are then perfectly determined. The description that we obtain is, in a sense, trivial: initially, the system is in one of the eigenstates that are contained in $\rho(t_0)$, and then evolves deterministically from this initial state.

But it is possible to make the family less singular by grouping together, for each time t_i , several projectors into one single projector; different associations of projectors may be used at different times. In this way, the description of the evolution of the state within this family becomes less accurate, but also less trivial since projectors at different times are no longer associated pair by pair. On the other hand, it is possible to see that this grouping of projectors has not destroyed the consistent character of the family; other methods for constructing consistent families are also possible.

References

- [1] N. Bohr, *Atomic Physics and Human Knowledge*, Wiley (1958), and Dover (2011), see in particular chapter “Discussions with Einstein on epistemological problems in atomic physics” – or, in French, with bibliography and glossary by C. Chevalley: *Physique atomique et connaissance humaine*, folio essais, Gallimard (1991); *Essays 1933 to 1957 on Atomic Physics and Human Knowledge*, Ox Bow Press (1987); *Essays 1958–62 on Atomic Physics and Human Knowledge*, Wiley (1963) and Ox Bow Press (1987); *Atomic Physics and the Description of Nature*, Cambridge University Press (1934 and 1961).
- [2] *Albert Einstein: Philosopher-Scientist*, P.A. Schilpp editor, Open Court and Cambridge University Press (1949).
- [3] G. Bacchiagaluppi and A. Valentini, *Quantum Theory at the Crossroads: Reconsidering the 1927 Solvay Conference*, Cambridge University Press (2009).
- [4] J. von Neumann, *Mathematische Grundlagen der Quantenmechanik*, Springer, Berlin (1932); *Mathematical Foundations of Quantum Mechanics*, Princeton University Press (1955).
- [5] J.S. Bell, “On the problem of hidden variables in quantum mechanics”, *Rev. Mod. Phys.* **38**, 447–452 (1966); reprinted in *Quantum Theory of Measurement*, J.A. Wheeler and W.H. Zurek editors, Princeton University Press (1983), 396–402 and in Chapter 1 of [6].
- [6] J.S. Bell, *Speakable and Unspeakable in Quantum Mechanics*, Cambridge University Press (1987); second augmented edition (2004), which contains the complete set of J. Bell’s articles on quantum mechanics.
- [7] D. Bohm and J. Bub, “A proposed solution of the measurement problem in quantum mechanics by a hidden variable theory”, *Rev. Mod. Phys.* **38**, 453–469 (1966).
- [8] D. Bohm and J. Bub, “A refutation of the proof by Jauch and Piron that hidden variables can be excluded in quantum mechanics”, *Rev. Mod. Phys.* **38**, 470–475 (1966).
- [9] N.D. Mermin, “Hidden variables and the two theorems of John Bell”, *Rev. Mod. Phys.* **65**, 803–815 (1993); in particular see §III.
- [10] A. Shimony, “Role of the observer in quantum theory”, *Am. J. Phys.* **31**, 755–773 (1963).
- [11] D. Bohm, “A suggested interpretation of the quantum theory in terms of ‘hidden’ variables”, *Phys. Rev.* **85**, 166–179 and 180–193 (1952); by the same author, see also *Quantum Theory*, Constable (1954), although this book does not discuss theories with additional variables.

- [12] N. Wiener and A. Siegel, “A new form for the statistical postulate of quantum mechanics”, *Phys. Rev.* **91**, 1551–1560 (1953); A. Siegel and N. Wiener, “Theory of measurement in differential space quantum theory”, *Phys. Rev.* **101**, 429–432 (1956).
- [13] P. Pearle, “Reduction of the state vector by a nonlinear Schrödinger equation”, *Phys. Rev. D* **13**, 857–868 (1976).
- [14] P. Pearle, “Toward explaining why events occur”, *Int. J. Theor. Phys.* **18**, 489–518 (1979).
- [15] G.C. Ghirardi, A. Rimini, and T. Weber, “Unified dynamics for microscopic and macroscopic systems”, *Phys. Rev. D* **34**, 470–491 (1986); “Disentanglement of quantum wave functions”, *Phys. Rev. D* **36**, 3287–3289 (1987).
- [16] B.S. DeWitt, “Quantum mechanics and reality”, *Phys. Today* **23**, 30–35 (September 1970).
- [17] R.B. Griffiths, “Consistent histories and the interpretation of quantum mechanics”, *J. Stat. Phys.* **36**, 219–272 (1984).
- [18] S. Goldstein, “Quantum theory without observers”, *Phys. Today* **51**, 42–46 (March 1998) and 38–41 (April 1998).
- [19] “Quantum mechanics debate”, *Phys. Today* **24**, 36–44 (April 1971); “Still more quantum mechanics”, *Phys. Today* **24**, 11–15 (Oct. 1971).
- [20] B.S. DeWitt and R.N. Graham, “Resource letter IQM-1 on the interpretation of quantum mechanics”, *Am. J. Phys.* **39**, 724–738 (1971).
- [21] M. Jammer, *The Conceptual Development of Quantum Mechanics*, McGraw Hill (1966), second edition (1989).
- [22] J. Mehra and H. Rechenberg, *The Historical Development of Quantum Theory*, Springer (1982).
- [23] O. Darrigol, *From c-Numbers to q-Numbers: The Classical Analogy in the History of Quantum Theory*, University of California Press (1992).
- [24] B. d’Espagnat, *Conceptual Foundations of Quantum Mechanics*, Benjamin, New York (1971).
- [25] B. d’Espagnat, *Veiled Reality: An Analysis of Present Day Quantum Mechanics Concepts*, Addison Wesley (1995); *Le réel voilé, analyse des concepts quantiques*, Fayard, Paris (1994); *Une incertaine réalité, la connaissance et la durée*, Gauthier-Villars, Paris (1985); *A la recherche du réel*, Gauthier Villars Bordas, Paris (1979).
- [26] M. Planck, “Über eine Verbesserung der Wienerschen Spektralgleichung”, *Verhandlungen der Deutschen Physikalischen Gesellschaft* **2**, 202–204 (1900). *Physikalische Abhandlungen und Vorträge*, vol. 1, 493–600, Friedrich Vieweg und Sohn (1958).
- [27] G.N. Lewis, “The conservation of photons”, *Nature* **118**, 874–875 (1926).
- [28] A. Pais, “Einstein and the quantum theory”, *Rev. Mod. Phys.* **51**, 863–914 (1979).
- [29] E.H. Lieb and R. Seiringer, *The Stability of Matter in Quantum Mechanics*, Cambridge University Press (2010).
- [30] L. de Broglie, “Recherches sur la théorie des quanta”, thesis Paris (1924).
- [31] C.J. Davisson and L.H. Germer, “Reflection of electrons by a crystal of nickel”, *Nature* **119**, 558–560 (1927).
- [32] O. Darrigol, “Strangeness and soundness in Louis de Broglie’s early works”, *Physica* **30**, 303–372 (1993).
- [33] E. Schrödinger, “Quantisierung als Eigenwert Problem”, *Annalen der Physik*, 1st communication: **79**, 361–376 (1926); 2nd communication: **79**, 489–527 (1926); 3rd communication: **80**, 437–490 (1926); 4th communication: **81**, 109–139 (1926).
- [34] M. Born, “Quantenmechanik der Stossvorgänge”, *Zeitschrift für Physik* **38**, 803–827 (1926); “Zur Wellenmechanik der Stossvorgänge”, *Göttingen Nachrichten* 146–160 (1926).

- [35] E.A. Cornell and C.E. Wieman, “The Bose–Einstein condensate”, *Scientific American* **278**, 26–31 (March 1998).
- [36] W. Heisenberg, *The Physical Principles of the Quantum Theory*, University of Chicago Press (1930).
- [37] P. Jordan, “Bemerkungen zur Theorie der Atomstruktur”, *Zeitschrift für Physik* **33**, 563–570 (1925); “Über eine neue Begründung der Quantenmechanik I und II”, *Zeitschrift für Physik* **40**, 809–838 (1926) and **44**, 1–25 (1927); “Austauschprobleme und zweite Quantelung”, *Zeitschrift für Physik* **91**, 284–288 (1934).
- [38] B. Schroer, “Pascual Jordan, his contributions to quantum mechanics, and his legacy in contemporary local quantum physics”, arXiv: hep-th/0303241v2 (2003).
- [39] P.A.M. Dirac, *The Principles of Quantum Mechanics*, Oxford University Press (1930, 1958).
- [40] D. Howard, “Who invented the Copenhagen interpretation? A study in mythology”, *Philos. Sci.* **71**, 669–682 (2004).
- [41] N. Bohr, “Can quantum-mechanical description of physical reality be considered complete?”, *Phys. Rev.* **48**, 696–702 (1935).
- [42] H.P. Stapp, “S-matrix interpretation of quantum theory”, *Phys. Rev.* **D 3**, 1303–1320 (1971).
- [43] H.P. Stapp, “The Copenhagen interpretation”, *Am. J. Phys.* **40**, 1098–116 (1972).
- [44] A. Peres, “What is a state vector?”, *Am. J. Phys.* **52**, 644–650 (1984).
- [45] J.B. Hartle, “Quantum mechanics of individual systems”, *Am. J. Phys.* **36**, 704–712 (1968).
- [46] N. Bohr, “On the notions of causality and complementarity”, *Dialectica* **2**, 312–319 (1948).
- [47] J.S. Bell, “Quantum mechanics for cosmologists”, in *Quantum Gravity*, C. Isham, R. Penrose, and D. Sciama editors, **2**, 611–637, Clarendon Press (1981); pages 117–138 of [6].
- [48] D. Mermin, “Is the moon there when nobody looks? Reality and the quantum theory”, *Phys. Today* **38**, 38–47 (April 1985).
- [49] F. London and E. Bauer, “La théorie de l’observation en mécanique quantique”, no. 775 of *Actualités scientifiques et industrielles, exposés de physique générale*; Hermann, Paris (1939); translated into English as “The theory of observation in quantum mechanics” in *Quantum Theory of Measurement*, J.A. Wheeler and W.H. Zurek editors, Princeton University Press, pp. 217–259 (1983); see in particular §11, but also 13 and 14.
- [50] M. Jammer, *The Philosophy of Quantum Mechanics*, Wiley (1974).
- [51] E. Schrödinger, “Die gegenwärtige Situation in der Quantenmechanik”, *Naturwissenschaften* **23**, 807–812, 823–828, 844–849 (1935).
- [52] J.D. Trimmer, “The present situation in quantum mechanics: a translation of Schrödinger’s cat paradox paper”, *Proc. Amer. Phil. Soc.* **124**, 323–338 (1980). Also available in pp. 152–167, *Quantum Theory of Measurement*, J.A. Wheeler and W.H. Zurek editors, Princeton University Press (1983).
- [53] A. Einstein, letter to Schrödinger dated 8 August 1935; available for instance (translated into French) page 238 of [75].
- [54] E.P. Wigner, “The problem of measurement”, *Am. J. Phys.* **31**, 6–15 (1963); reprinted in *Symmetries and Reflections*, Indiana University Press, pp. 153–170; or again in *Quantum Theory of Measurement*, J.A. Wheeler and W.H. Zurek editors, Princeton University Press (1983), pp. 324–341.

- [55] E.P. Wigner, “Remarks on the mind-body question” in *The Scientist Speculates*, I.J. Good editor, Heinemann, London (1961), pp. 284–302; reprinted in E.P. Wigner, *Symmetries and Reflections*, Indiana University Press (1967), pp. 171–184.
- [56] A.C. Elitzur and L. Vaidman, “Quantum mechanical interaction-free measurements”, *Found. Phys.* **23**, 987–997 (1993).
- [57] P. Kwiat, H. Weinfurter, T. Herzog, A. Zeilinger, and M.A. Kasevich, “Interaction-free measurement”, *Phys. Rev. Lett.* **74**, 4763–4766 (1995).
- [58] L. Hardy, “On the existence of empty waves in quantum theory”, *Phys. Lett. A* **167**, 11–16 (1992).
- [59] L. Hardy, “Quantum mechanics, local realistic theories, and Lorentz invariant realistic theories”, *Phys. Rev. Lett.* **68**, 2981–2984 (1992).
- [60] Tae-Gon Noh, “Counterfactual quantum cryptography”, *Phys. Rev. Lett.* **103**, 230501 (2009).
- [61] A. Petersen, “The philosophy of Niels Bohr”, in *Bulletin of the Atomic Scientists* **XIX**, 8–14 (September 1963).
- [62] C. Chevalley, “Niels Bohr’s words and the Atlantis of Kantianism”, in *Niels Bohr and Contemporary Philosophy*, J. Faye and H. Folse editors, Dordrecht Kluwer (1994), pp. 33–57.
- [63] N. Bohr, “The unity of human knowledge” (October 1960); *Atomic Physics and Human Knowledge*, Wiley (1958 and 1963).
- [64] C. Norris, “*Quantum Theory and the Flight from Realism: Philosophical Responses to Quantum Mechanics*, Routledge (2000), p. 233.
- [65] N. Bohr, “Quantum physics and philosophy: causality and complementarity”, in *Philosophy in the Mid-Century: A Survey*, R. Klibansky editor, La Nuova Italia Editrice, Firenze (1958). See also “The quantum of action and the description of nature”, in *Atomic Theory and the Description of Nature*, Cambridge University Press (1934), pp. 92–101.
- [66] P. Bokulich and A. Bokulich, “Niels Bohr’s generalization of classical mechanics”, *Found. Phys.* **35**, 347–371 (2005).
- [67] N. Bohr, “Atomic theory and mechanics”, *Nature* **116**, 845–852 (1925).
- [68] N. Bohr, *Collected Works*, edited by F. Aaserud, Elsevier (2008); see also *Collected Works, Complementarity beyond Physics (1928–1962)*.
- [69] M. Born, “Physical aspects of quantum mechanics”, *Nature* **119**, 354–357 (1927).
- [70] W. Heisenberg, *Physics and Philosophy*, Harper & Brothers (1958); Harper Perennial Modern Classics (2007).
- [71] J.S. Bell, “Bertlmann’s socks and the nature of reality”, *J. Physique colloques* **C2**, 41–62 (1981). This article is reprinted in pages 139–158 of [6].
- [72] L.D. Landau and E.M. Lifshitz, *Quantum Mechanics, Non-Relativistic Theory*, Pergamon Press (1958), Butterworth-Heinemann Ltd (1996).
- [73] E. Schrödinger, *What is Life? Mind and Matter*, Cambridge University Press (1944 and 1967), p. 137.
- [74] A. Einstein, letter to Schrödinger dated 31 May 1928; available for instance (translated into French) page 213 of [75].
- [75] F. Balibar, O. Darrigol, and B. Jech, *Albert Einstein, oeuvres choisies I, quanta*, Editions du Seuil et Editions du CNRS (1989).
- [76] A. Einstein, “Physik und Realität”, *Journal of the Franklin Institute* **221**, 313–347 (1936).
- [77] L. de Broglie, “La physique quantique restera-t-elle indéterministe?”, *Revue des sciences et de leurs applications* **5**, 289–311 (1952). French Academy of Sciences, 25 April 1953 session, <http://www.sofrphilos.fr/telecharger.php?id=74>

- [78] J.S. Bell, “Against measurement”, in *Sixty Two Years of Uncertainty: Historical, Philosophical and Physical Enquiries into the Foundations of Quantum Mechanics*, Erice meeting in August 1989, A.I. Miller editor (Plenum Press); reprinted in pp. 213–231 of the 2004 edition of [6].
- [79] L. Rosenfeld, “The measuring process in quantum mechanics”, *Suppl. Prog. Theor. Phys.*, extra number 222 “Commemoration of the thirtieth anniversary of the meson theory by Dr. H. Yukawa” (1965).
- [80] K. Gottfried, *Quantum Mechanics*, Benjamin (1966); second edition, K. Gottfried and Yan Tun-Mow, Springer (2003).
- [81] A.J. Leggett, “Testing the limits of quantum mechanics: motivation, state of play, prospects”, *J. Phys. Condens. Matter* **14**, R415–R451 (2002).
- [82] A.J. Leggett, “Macroscopic quantum systems and the quantum theory of measurement”, Supplement of the *Progr. Theor. Phys.* no 69, 80–100 (1980).
- [83] A.J. Leggett, *The Problems of Physics*, Oxford University Press (1987).
- [84] N.G. van Kampen, “Ten theorems about quantum mechanical measurements”, *Physica A* **153**, 97–113 (1988).
- [85] B.G. Englert, M.O. Scully, and H. Walther, “Quantum erasure in double-slit interferometers with which-way detectors”, *Am. J. Phys.* **67**, 325–329 (1999); see the first few lines of §IV.
- [86] C.A. Fuchs and A. Peres, “Quantum theory needs no ‘interpretation’”, *Phys. Today* **53**, March 2000, 70–71; see also various reactions to this text in the letters of the September 2000 issue.
- [87] C.F. von Weizsäcker, *Voraussetzungen des naturwissenschaftlichen Denkens*, Hanser Verlag (1971) and Herder (1972).
- [88] A. Einstein, B. Podolsky, and N. Rosen, “Can quantum-mechanical description of physical reality be considered complete?”, *Phys. Rev.* **47**, 777–780 (1935); or in *Quantum Theory of Measurement*, J.A. Wheeler and W.H. Zurek editors, Princeton University Press (1983), pp. 138–141.
- [89] M. Born editor, *The Einstein–Born Letters (1916–1955)*, MacMillan, London (1971).
- [90] A. Einstein, letter to Schrödinger dated 19 June 1935; available for instance (translated into French) page 234 of [75].
- [91] A. Einstein, “Quantenmechanik und Wirklichkeit”, *Dialectica* **2**, 320–324 (1948).
- [92] A. Einstein, “Autobiographical notes” pages 5–94 (especially p. 85) and “Reply to criticism” pages 663–688 (especially pp. 681–3) in *Albert Einstein: Philosopher-Scientist*, edited by P.A. Schilpp, Open Court and Cambridge University Press (1949).
- [93] T. Sauer, “An Einstein manuscript on the EPR paradox for spin observables”, *Studies in History and Philosophy of Modern Physics* **38**, 879–887 (2007).
- [94] A. Peres, “Einstein, Podolsky, Rosen, and Shannon”, *Found. Phys.* **35**, 511–514 (2005).
- [95] D. Home and F. Selleri, “Bell’s theorem and the EPR paradox”, *Rivista del Nuov. Cim.* **14**, 1–95 (1991).
- [96] D. Bohm, *Quantum Theory*, Prentice Hall (1951).
- [97] N. Bohr, “Quantum mechanics and physical reality”, *Nature* **136**, 65 (1935).
- [98] P. Pearle, “Alternative to the orthodox interpretation of quantum theory”, *Am. J. Phys.* **35**, 742–753 (1967).
- [99] J.F. Clauser and A. Shimony, “Bell’s theorem : experimental tests and implications”, *Rep. Progr. Phys.* **41**, 1881–1927 (1978).
- [100] F. Laloë, “The hidden phase of Fock states; quantum non-local effects”, *Europ. Phys. J.* **33**, 87–97 (2005); “Bose–Einstein condensates and quantum non-locality”, in *Beyond the Quantum*, T.M. Nieuwenhuizen *et al.* editors, World Scientific (2007).

- [101] P.W. Anderson, in *The Lesson of Quantum Theory*, editors J. de Boer, E. Dahl, and O. Ulfbeck, Elsevier, New York (1986).
- [102] W.J. Mullin and F. Laloë, “Quantum non-local effects with Bose–Einstein condensates”, *Phys. Rev. Lett.* **99**, 150401 (2007); “EPR argument and Bell inequalities for Bose–Einstein spin condensates”, *Phys. Rev. A* **77**, 022108 (2008).
- [103] J.S. Bell, “On the Einstein–Podolsky–Rosen paradox”, *Physics* **I**, 195–200 (1964); reprinted in Chapter 2 of [6].
- [104] F. Laloë, “Les surprenantes prédictions de la mécanique quantique”, *La Recherche* no. 182, 1358–1367 (November 1986).
- [105] F. Laloë, “Cadre général de la mécanique quantique, les objections de Einstein, Podolsky et Rosen”, *J. Physique Colloques* **C-2**, 1–40 (1981). See also the other articles following in this issue, especially that by J. Bell on Bertlmann socks, which is a classic!
- [106] P. Eberhard, “Bell’s theorem without hidden variables”, *Nuov. Cim.* **38 B**, 75–79 (1977); “Bell’s theorem and the different concepts of locality”, *Nuov. Cim.* **46 B**, 392–419 (1978).
- [107] J.F. Clauser, M.A. Horne, A. Shimony, and R.A. Holt, “Proposed experiment to test local hidden-variables theories”, *Phys. Rev. Lett.* **23**, 880–884 (1969).
- [108] A. Peres, “Unperformed experiments have no results”, *Am. J. Phys.* **46**, 745–747 (1978).
- [109] J.A. Wheeler, “Niels Bohr in today’s words” in *Quantum Theory and Measurement*, J.A. Wheeler and W.H. Zurek editors, Princeton University Press (1983), pp. 182–213.
- [110] E.P. Wigner, “On hidden variables and quantum mechanical probabilities”, *Am. J. Phys.* **38**, 1005–1009 (1970).
- [111] K. Hess and W. Philipp, “The Bell theorem as a special case of a theorem of Bass”, *Found. Phys.* **35**, 1749–1767 (2005).
- [112] J. Bass, “Sur la compatibilité des fonctions de répartition”, *C.R. Académie des Sciences* **240**, 839–841 (1955).
- [113] C.A. Kocher and E.D. Commins, “Polarization correlation of photons emitted in an atomic cascade”, *Phys. Rev. Lett.* **18**, 575–577 (1967).
- [114] S.J. Freedman and J.F. Clauser, “Experimental test of local hidden variable theories”, *Phys. Rev. Lett.* **28**, 938–941 (1972); S.J. Freedman, thesis, University of California, Berkeley.
- [115] J.F. Clauser, “Experimental investigations of a polarization correlation anomaly”, *Phys. Rev. Lett.* **36**, 1223 (1976).
- [116] E.S. Fry and R.C. Thompson, “Experimental test of local hidden variable theories”, *Phys. Rev. Lett.* **37**, 465–468 (1976).
- [117] M. Laméhi-Rachti and W. Mittig, “Quantum mechanics and hidden variables: a test of Bell’s inequality by the measurement of the spin correlation in low energy proton-proton scattering”, *Phys. Rev. D* **14**, 2543–2555 (1976).
- [118] A. Aspect, P. Grangier, and G. Roger, “Experimental tests of realistic local theories via Bell’s theorem”, *Phys. Rev. Lett.* **47**, 460–463 (1981).
- [119] A. Aspect, P. Grangier and G. Roger, “Experimental realization of Einstein–Podolsky–Bohm Gedanken experiment: a new violation of Bell’s inequalities”, *Phys. Rev. Lett.* **49**, 91–94 (1982).
- [120] A. Aspect, J. Dalibard, and G. Roger, “Experimental tests of Bell’s inequalities using time varying analyzers”, *Phys. Rev. Lett.* **49**, 1804–1807 (1982).
- [121] W. Perrie, A.J. Duncan, H.J. Beyer, and H. Kleinpoppen, “Polarization correlations of the two photons emitted by metastable atomic deuterium: a test of Bell’s inequality”, *Phys. Rev. Lett.* **54**, 1790–1793 (1985).

- [122] T.E. Kiess, Y.E. Shih, A.V. Sergienko, and C.O. Alley, “Einstein–Podolsky–Rosen–Bohm experiments using pairs of light quanta produced by type-II parametric down conversion”, *Phys. Rev. Lett.* **71**, 3893–3897 (1993).
- [123] W. Tittel, J. Brendel, H. Zbinden, and N. Gisin, “Violations of Bell inequalities by photons more than 10 km apart”, *Phys. Rev. Lett.* **81**, 3563–3566 (1998).
- [124] T. Scheidl, R. Ursin, J. Kofler, S. Ramelow, X.S. Ma, T. Herbst, L. Ratschbacher, A. Fedrizzi, N.K. Langford, T. Jennewein, and A. Zeilinger, “Violations of local realism with freedom of choice”, *Proc. Nat. Acad. Sciences* **107**, 19708–19713 (November 16, 2010).
- [125] J.C. Howell, A. Lamas-Linares, and D. Bouwmeester, “Experimental violation of a spin-1 Bell inequality using maximally entangled four photon states”, *Phys. Rev. Lett.* **88**, 030401 (2002).
- [126] N.D. Mermin, “Bringing home the atomic world: quantum mysteries for anybody”, *Am. J. Phys.* **49**, 940–943 (1981).
- [127] A. Fine, “Hidden variables, joint probability, and the Bell inequalities”, *Phys. Rev. Lett.* **48**, 291–295 (1982).
- [128] J.S. Bell, “Introduction to the hidden variable question”, contribution to *Foundations of Quantum Mechanics*, Proceedings of the International School of Physics Enrico Fermi, course II, Academic, New York (1971), p.171; reprinted in pages 29–39 of [6].
- [129] J.D. Franson, “Bell inequality for position and time”, *Phys. Rev. Lett.* **62**, 2205–2208 (1989).
- [130] J. Brendel, E. Mohler, and W. Martienssen, “Experimental test of Bell’s inequality for energy and time”, *Eur. Phys. Lett.* **20**, 575–580 (1992).
- [131] V. Capasso, D. Fortunato, and F. Selleri, “Sensitive observables of quantum mechanics”, *Int. J. Theor. Phys.* **7**, 319–326 (1973).
- [132] N. Gisin, “Bell’s inequality holds for all non-product states”, *Phys. Lett.* **A154**, 201–202 (1991).
- [133] N. Gisin and A. Peres, “Maximal violation of Bell’s inequality for arbitrarily large spin”, *Phys. Lett. A* **162**, 15–17 (1992).
- [134] S. Popescu and D. Rohrlich, “Generic quantum non-locality”, *Phys. Lett.* **A166**, 293–297 (1992).
- [135] S.L. Braunstein, A. Mann, and M. Revzen, “Maximal violation of Bell inequalities for mixed states”, *Phys. Rev. Lett.* **68**, 3259–3261 (1992).
- [136] R.F. Werner, “Quantum states with Einstein–Podolsky–Rosen correlations admitting a hidden variable model”, *Phys. Rev. A* **40**, 4277–4281 (1989).
- [137] S. Popescu, “Bell’s inequalities and density matrices: revealing ‘hidden’ nonlocality”, *Phys. Rev. Lett.* **74**, 2619–2622 (1995).
- [138] A. Peres, “Collective tests for quantum nonlocality”, *Phys. Rev. A* **54**, 2685–2689 (1996).
- [139] B. Yurke and D. Stoler, “Bell’s-inequality experiments using independent-particle sources”, *Phys. Rev. A* **46**, 2229–2234 (1992).
- [140] F. Laloë and W.J. Mullin, “Interferometry with independent Bose–Einstein condensates: parity as an EPR/Bell quantum variable”, *Eur. Phys. J. B* **70**, 377–396 (2009).
- [141] A.J. Leggett and A. Garg, “Quantum mechanics versus macroscopic realism: is the flux there when nobody looks?”, *Phys. Rev. Lett.* **54**, 857–860 (1985).
- [142] S.M. Tan, D.F. Walls, and M.J. Collett, “Nonlocality of a single photon”, *Phys. Rev. Lett.* **66**, 252–255 (1991).
- [143] L. Hardy, “Nonlocality of a single photon revisited”, *Phys. Rev. Lett.* **73**, 2279–2283 (1994).

- [144] L. Heaney, A. Cabello, M.F. Santos, and V. Vedral, “Extreme non-locality with one photon”, arXiv: 0911.0770v3 [quant-ph] (2009); *New J. Phys.* **13**, 053054 (2011).
- [145] B.F. Toner and D. Bacon, “Communication cost of simulating Bell correlations”, *Phys. Rev. Lett.* **91**, 187904 (2003).
- [146] B.S. Cirel’son, “Quantum generalizations of Bell’s inequality”’ *Lett. Math. Phys.* **4**, 93–100 (1980).
- [147] L.J. Landau, “On the violations of Bell’s inequality in quantum theory”, *Phys. Lett. A* **120**, 54–56 (1987).
- [148] J.P. Jarrett, “On the physical significance of the locality conditions in the Bell arguments”, *Noûs* **18**, 569–589 (1984).
- [149] L.E. Ballentine and J.P. Jarrett, “Bell’s theorem: does quantum mechanics contradict relativity?”, *Am. J. Phys.* **55**, 696–701 (1987).
- [150] A. Shimony, “Events and processes in the quantum world”, in *Quantum Concepts in Space and Time*, edited by R. Penrose and C.J. Isham, Oxford University Press (1986), pp. 182–203.
- [151] S. Popescu and D. Rohrlich, “Quantum nonlocality as an axiom”, *Found. Phys.* **24**, 379–385 (1994).
- [152] J. Barrett, N. Linden, S. Massar, S. Pironio, S. Popescu, and D. Roberts, “Non-local correlations as an information-theoretic resource”, *Phys. Rev. A* **71**, 022101 (2005).
- [153] N. Gisin, “Hidden quantum nonlocality revealed by local filters”, *Phys. Lett. A* **210**, 151–156 (1996); see in particular §3.
- [154] D.S. Tasca, S.P. Walborn, F. Toscano, and P.H. Souto Ribeiro, “Observation of tunable Popescu–Rohrlich correlations through postselection of a Gaussian state”, *Phys. Rev. A* **80**, 030101 (2009).
- [155] I. Gerhardt, Q. Liu, A. Lamas-Linares, J. Skaar, V. Scarani, V. Makarov, and C. Kurtsiefer, “Experimentally faking the violation of Bell’s inequalities”, *Phys. Rev. Lett.* **107**, 170404 (2011).
- [156] L. Masanes, A. Acin, and N. Gisin, “General properties of nonsignalling theories”, *Phys. Rev. A* **73**, 012112 (2006).
- [157] G. Brassard, H. Buhrman, N. Linden, A.A. Méthot, A. Tapp, and F. Unger, “Limit on nonlocality in any world in which communication complexity is not trivial”, *Phys. Rev. Lett.* **96**, 250401 (2006).
- [158] M. Pawłowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, and D. Rohrlich, “Information causality as a physical principle”, *Nature* **461**, 1101–1104 (2009).
- [159] H. Barnum, S. Beigi, S. Boixo, M.B. Elliott, and S. Wehner, “Local quantum measurements and no-signaling imply quantum correlations”, *Phys. Rev. Lett.* **104**, 140401 (2010).
- [160] M.L. Almeida, J.-D. Bancal, N. Brunner, A. Acin, N. Gisin, and S. Pironio, “Guess your neighbor’s input: a multipartite nonlocal game with no quantum advantage”, *Phys. Rev. Lett.* **104**, 230404 (2010).
- [161] P. Pearle, “Hidden-variable example based upon data rejection”, *Phys. Rev. D* **2**, 1418–1425 (1970).
- [162] J.F. Clauser and M.A. Horne, “Experimental consequences of objective local theories”, *Phys. Rev. D* **10**, 526–535 (1974).
- [163] J.S. Bell, Oral presentation to the EGAS conference in Paris, July 1979 (published in an abridged version in the next reference).
- [164] J.S. Bell, “Atomic cascade photons and quantum mechanical non-locality”, *Comments on Atomic and Molecular Physics* **9**, 121 (1980); CERN preprint TH.2053 and TH 2252; Chapter 13 of [6].

- [165] P.H. Eberhard, “Background level and counter efficiencies required for a loophole-free Einstein–Podolsky–Rosen experiment”, *Phys. Rev. A* **47**, R747–R750 (1993).
- [166] P.G. Kwiat, P.H. Eberhard, A.M. Steinberg, and R.Y. Chiao, “Proposal for a loophole-free Bell inequality experiment”, *Phys. Rev. A* **49**, 3209–3220 (1994).
- [167] E.S. Fry, T. Walther, and S. Li, “Proposal for a loophole-free test of the Bell inequalities”, *Phys. Rev. A* **52**, 4381–4395 (1995).
- [168] R. Garcia-Patron, J. Fiurasek, N.J. Cerf, J. Wenger, R. Tualle-Brouri, and P. Grangier, “Proposal for a loophole-free Bell test using homodyne detection”, *Phys. Rev. Lett.* **93**, 130409 (2004).
- [169] J. Wenger, M. Hafezi, F. Grosshans, R. Tualle-Brouri, and P. Grangier, “Maximal violation of Bell inequalities using continuous-variable measurements”, *Phys. Rev. A* **67**, 012105 (2003).
- [170] M.A. Rowe, D. Kielpinski, V. Meyer, C.A. Sackett, W.M. Itano, C. Monroe, and D.J. Wineland, “Experimental violation of a Bell’s inequality with efficient detection”, *Nature* **409**, 791–794 (2001).
- [171] C. Simon and W.T.M. Irvine, “Robust long-distance entanglement and a loophole-free Bell test with ions and photons”, *Phys. Rev. Lett.* **91**, 110405 (2003).
- [172] D.N. Matsukevich, T. Chanelière, S.D. Jenkins, S.Y. Lan, T.A.B. Kennedy, and A. Kuzmich, “Entanglement of remote atomic qubits”, *Phys. Rev. Lett.* **96**, 030405 (2006).
- [173] D.N. Matsukevich, P. Maunz, D.L. Moehring, S. Olmschenk, and C. Monroe, “Bell inequality violation with two remote atomic qubits”, *Phys. Rev. Lett.* **100**, 150404 (2008).
- [174] M. Ansmann, H. Wang, R.C. Bialczak, M. Hofheinz, E. Lucero, M. Neeley, A.D. O’Connell, D. Sank, M. Weides, J. Wenner, A.N. Cleland, and J.M. Martinis, “Violation of Bell’s inequality in Josephson phase qubits”, *Nature* **461**, 504–506 (2009).
- [175] G. Weihs, T. Jennewein, C. Simon, H. Weinfurter, and A. Zeilinger, “Violation of Bell’s inequality under strict Einstein locality conditions”, *Phys. Rev. Lett.* **81**, 5039–5043 (1998).
- [176] C.H. Brans, “Bell’s theorem does not eliminate fully causal hidden variables”, *Int. J. Theor. Phys.* **27**, 219–226 (1988).
- [177] J.S. Bell, “Free variables and local causality”, Chapter 12 of [6].
- [178] M.J.W. Hall, “Local deterministic model of singlet state correlations based on relaxing measurement independence”, *Phys. Rev. Lett.* **105**, 250404 (2010).
- [179] J.S. Bell, “La nouvelle cuisine”, §24 of the second edition of [6].
- [180] H.P. Stapp, “Whiteheadian approach to quantum theory and the generalized Bell’s theorem”, *Found. Phys.* **9**, 1–25 (1979); “Bell’s theorem and the foundations of quantum physics”, *Am. J. Phys.* **53**, 306–317 (1985).
- [181] H.P. Stapp, “Nonlocal character of quantum theory”, *Am. J. Phys.* **65**, 300–304 (1997).
- [182] D. Mermin, “Nonlocal character of quantum theory?”, *Am. J. Phys.* **66**, 920–924 (1998).
- [183] M. Redhead, *Incompleteness, Nonlocality and Realism: A Prolegomenon to the Philosophy of Quantum Mechanics*, Chapter 4, Clarendon Press (1988).
- [184] H.P. Stapp, “Meaning of counterfactual statements in quantum physics”, *Am. J. Phys.* **66**, 924–926 (1998).
- [185] B. d’Espagnat, “Nonseparability and the tentative descriptions of reality”, *Phys. Rep.* **110**, 201–264 (1984).
- [186] B. d’Espagnat, *Reality and the Physicist*, Cambridge University Press (1989).
- [187] R.B. Griffiths, “Consistent quantum counterfactuals”, *Phys. Rev. A* **60**, R5–R8 (1999).

- [188] D.M. Greenberger, M.A. Horne, and A. Zeilinger, in *Bell's Theorem, Quantum Theory, and Conceptions of the Universe*, M. Kafatos editor, Kluwer (1989), pp. 69–72; this reference is not always easy to find, but one can also read the following article, published one year later.
- [189] D.M. Greenberger, M.A. Horne, A. Shimony, and A. Zeilinger, “Bell’s theorem without inequalities”, *Am. J. Phys.* **58**, 1131–1143 (1990).
- [190] N.D. Mermin, “Quantum mysteries revisited”, *Am. J. Phys.* **58**, 731–733 (1990).
- [191] D. Bouwmeester, J.W. Pan, M. Daniell, H. Weinfurter, and A. Zeilinger, “Observation of three-photon Greenberger–Horne–Zeilinger entanglement”, *Phys. Rev. Lett.* **82**, 1345–1349 (1999).
- [192] J.W. Pan, D. Bouwmeester, M. Daniell, H. Weinfurter, and A. Zeilinger, “Experimental test of quantum nonlocality in three-photon Greenberger–Horne–Zeilinger entanglement”, *Nature* **403**, 515–519 (2000).
- [193] Z. Zhao, T. Yang, Y.-A. Chen, A.-N. Zhang, M. Zukowski, and J.W. Pan, “Experimental violation of local realism by four-photon Greenberger–Horne–Zeilinger entanglement”, *Phys. Rev. Lett.* **91**, 180401 (2003).
- [194] R. Laflamme, E. Knill, W.H. Zurek, P. Catasi, and S.V.S. Mariappan, “NMR GHZ”, arXiv: quant-physics/9709025 (1997) and *Phil. Trans. Roy. Soc. Lond.* **A 356**, 1941–1948 (1998).
- [195] S. Lloyd, “Microscopic analogs of the Greenberger–Horne–Zeilinger experiment”, *Phys. Rev.* **A 57**, R1473–1476 (1998).
- [196] U. Sakaguchi, H. Ozawa, C. Amano, and T. Fokumi, “Microscopic analogs of the Greenberger–Horne–Zeilinger experiment on an NMR quantum computer”, *Phys. Rev.* **60**, 1906–1911 (1999).
- [197] N.D. Mermin, “Extreme quantum entanglement in a superposition of macroscopically distinct states”, *Phys. Rev. Lett.* **65**, 1838–1841 (1990).
- [198] J. Acacio de Barros, and P. Suppes, “Inequalities for dealing with detector efficiencies in Greenberger–Horne–Zeilinger experiments”, *Phys. Rev. Lett.* **84**, 793–797 (2000).
- [199] B. Yurke and D. Stoler, “Einstein–Podolsky–Rosen effects from independent particle sources”, *Phys. Rev. Lett.* **68**, 1251–1254 (1992).
- [200] S. Massar and S. Pironio, “Greenberger–Horne–Zeilinger paradox for continuous variables”, *Phys. Rev.* **A 64**, 062108 (2001).
- [201] H.J. Bernstein, D.M. Greenberger, M.A. Horne, and A. Zeilinger, “Bell theorem without inequalities for two spinless particles”, *Phys. Rev.* **A 47**, 78–84 (1993).
- [202] F. Laloë, “Correlating more than two particles in quantum mechanics”, *Current Science* **68**, 1026–1035 (1995); <http://hal.archives-ouvertes.fr/hal-00001443>.
- [203] D.J. Wineland, J.J. Bollinger, W.M. Itano, F.L. Moore, and D.J. Heinzen, “Spin squeezing and reduced quantum noise in spectroscopy”, *Phys. Rev.* **A 46**, R6797–6800 (1992).
- [204] J.J. Bollinger, W.M. Itano, D.J. Wineland, and D.J. Heinzen, “Optimal frequency measurements with maximally correlated states”, *Phys. Rev.* **A 54**, R4649–4652 (1996).
- [205] J.A. Dunningham, K. Burnett, and S.M. Barnett, “Interferometry below the standard limit with Bose–Einstein condensates”, *Phys. Rev. Lett.* **89**, 150401 (2002).
- [206] A.N. Boto, P. Kok, D.S. Abrams, S.L. Braunstein, C.P. Williams, and J.P. Dowling, “Quantum interferometric optical lithography: exploiting entanglement to beat the diffraction limit”, *Phys. Rev. Lett.* **85**, 2733–2736 (2000).
- [207] G. Björk, L.L. Sanchez-Soto, and J. Söderholm, “Entangled state lithography: tailoring any pattern with a single state”, *Phys. Rev. Lett.* **86**, 4516–4519 (2001).

- [208] M. d'Angelo, M.V. Chekhova, and Y. Shih, "Two-photon diffraction and quantum lithography", *Phys. Rev. Lett.* **87**, 013602 (2001).
- [209] A. Zeilinger, M.A. Horne, H. Weinfurter, and M. Zukowski, "Three-particle entanglements from two entangled pairs", *Phys. Rev. Lett.* **78**, 3031–3034 (1997).
- [210] K. Mølmer and A. Sørensen, "Multiparticle entanglement of hot trapped ions", *Phys. Rev. Lett.* **82**, 1835–1838 (1999).
- [211] C.A. Sackett, D. Klepinski, B.E. King, C. Langer, V. Meyer, C.J. Myatt, M. Rowe, O.A. Turchette, W.M. Itano, D.J. Wineland, and C. Monroe, "Experimental entanglement of four particles", *Nature* **404**, 256–259 (2000).
- [212] A. Cabello, "Violating Bell's inequalities beyond Cirelson's bound", *Phys. Rev. Lett.* **88**, 060403 (2002).
- [213] S. Marcovitch, B. Reznik, and L. Vaidman, "Quantum mechanical realization of a Popescu–Rohrlich box", *Phys. Rev. A* **75**, 022102 (2007).
- [214] N.D. Mermin, "What's wrong with this temptation?", *Phys. Today* **47**, June 1994, pp. 9–11; "Quantum mysteries refined", *Am. J. Phys.* **62**, 880–887 (1994).
- [215] D. Boschi, S. Branca, F. De Martini, and L. Hardy, "Ladder proof of nonlocality without inequalities: theoretical and experimental results", *Phys. Rev. Lett.* **79**, 2755–2758 (1997).
- [216] S. Goldstein, "Nonlocality without inequalities for almost all entangled states for two particles", *Phys. Rev. Lett.* **72**, 1951–1954 (1994).
- [217] G. Ghirardi and L. Marinatto, "Proofs of nonlocality without inequalities revisited", *Phys. Lett. A* **372**, 1982–1985 (2008).
- [218] S. Kochen and E.P. Specker, "The problem of hidden variables in quantum mechanics", *J. Math. Mech.* **17**, 59–87 (1967).
- [219] F. Belinfante, *Survey of Hidden Variables Theories*, Pergamon Press (1973).
- [220] A. Cabello, J.M. Estebaranz, and G. Garcia-Alcaine, "Bell–Kochen–Specker theorem: a proof with 18 vectors", *Phys. Lett. A* **212**, 183–187 (1996).
- [221] A. Peres, "Incompatible results of quantum measurements", *Phys. Lett. A* **151**, 107–108 (1990).
- [222] N.D. Mermin, "Simple unified form for the major no-hidden-variables theorems", *Phys. Rev. Lett.* **65**, 3373 (1990).
- [223] A. Cabello, "Experimentally testable state-independent quantum contextuality", *Phys. Rev. Lett.* **101**, 210401 (2008).
- [224] A. Cabello and G. Garcia-Alcaine, "Proposed experimental tests of the Bell–Kochen–Specker theorem", *Phys. Rev. Lett.* **80**, 1797–1799 (1998).
- [225] C. Simon, M. Zukowski, H. Weinfurter, and A. Zeilinger, "Feasible Kochen–Specker experiment with single particles", *Phys. Rev. Lett.* **85**, 1783–1786 (2000).
- [226] Y.-F. Huang, C.-F. Li, Y.-S. Zhang, J.-W. Pan, and G.-C. Guo, "Experimental test of the Kochen–Specker theorem with single photons", *Phys. Rev. Lett.* **90**, 250401 (2003).
- [227] A.A. Klyachko, M.A. Can, S. Binicioglu, and A.S. Shumovsky, "Simple tests for hidden variables in spin-1 systems", *Phys. Rev. Lett.* **101**, 020403 (2008).
- [228] R. Lapkiewicz, P. Li, C. Schaeff, N.K. Langford, S. Ramelow, M. Wieskiak, and A. Zeilinger, "Experimental non-classicality of an indivisible quantum system", *Nature* **474**, 490–493 (2011).
- [229] Y. Hasegawa, R. Loidl, G. Badurek, M. Baron, and H. Rauch, "Quantum contextuality in a single-neutron optical experiment", *Phys. Rev. Lett.* **97**, 230401 (2006).
- [230] G. Kirchmair, F. Zähringer, R. Gerritsma, M. Kleinmann, O. Gühne, A. Cabello, R. Blatt, and C.F. Roos, "State-independent experimental test of quantum contextuality", *Nature* **460**, 494–497 (2009).

- [231] O. Moussa, C.A. Ryan, D.G. Gory, and R. Laflamme, “Testing contextuality on quantum ensembles with one clean qubit”, *Phys. Rev. Lett.* **104**, 160501 (2010).
- [232] P. Grangier, “Contextual objectivity: a realistic interpretation of quantum mechanics”, *Eur. J. Phys.* **23**, 331–337 (2002); arXiv: quant-ph/0012122 (2000), 0111154 (2001), 0301001 (2003), and 0407025 (2004).
- [233] E. Schrödinger, “Discussion of probability relations between separated systems”, *Proc. Cambridge Phil. Soc.* **31**, 555 (1935); “Probability relations between separated systems”, *Proc. Cambridge Phil. Soc.* **32**, 446 (1936).
- [234] http://en.wikiquote.org/wiki/Werner_Heisenberg.
- [235] M. Horodecki, P. Horodecki, and R. Horodecki, “Limits for entanglement measures”, *Phys. Rev. Lett.* **84**, 2014–2017 (2000).
- [236] M. B. Plenio and S. Virmani, “An introduction to entanglement measures”, quant-ph/0504163 (2006); *Quant. Info. Comput.* **7**, 1–51 (2007).
- [237] A. Méthot and V. Scarani, “An anomaly of non-locality”, quant-ph/0601210 (2006); *Quant. Info. Comput.* **7**, 157–170 (2007).
- [238] V. Coffman, J. Kundu, and W.K. Wootters, “Distributed entanglement”, *Phys. Rev. A* **61**, 052306 (2000).
- [239] T.J. Osborne and F. Verstraete, “General monogamy inequality for bipartite qubit entanglement”, *Phys. Rev. Lett.* **96**, 220503 (2006).
- [240] B. Toner, “Monogamy of non-local quantum correlations”, *Proc. Roy. Soc. A* **465**, 59–68 (2009).
- [241] B. Toner and F. Verstraete, “Monogamy of Bell correlations and Tsirelson’s bound”, quant-ph/0611001 (2006).
- [242] A. Peres, “Separability criterion for density matrices”, *Phys. Rev. Lett.* **77**, 1413–1415 (1996).
- [243] M. Horodecki, P. Horodecki, and R. Horodecki, “Separability of mixed states: necessary and sufficient conditions”, *Phys. Lett. A* **223**, 1–8 (1996).
- [244] E. Hagley, X. Maître, G. Nogues, C. Wunderlich, M. Brune, J.M. Raimond, and S. Haroche, “Generation of Einstein–Podolsky–Rosen pairs of atoms”, *Phys. Rev. Lett.* **79**, 1–5 (1997).
- [245] Q.A. Turchette, C.S. Wood, B.E. King, C.J. Myatt, D. Leibfried, W.M. Itano, C. Monroe, and D.J. Wineland, “Deterministic entanglement of two trapped ions”, *Phys. Rev. Lett.* **81**, 3631–4 (1998).
- [246] J.I. Cirac and P. Zoller, “Quantum computations with cold trapped ions”, *Phys. Rev. Lett.* **74**, 4091–4094 (1995).
- [247] R. Blatt and D. Wineland, “Entangled states of trapped atomic ions”, *Nature* **453**, 1008–1015 (2008).
- [248] M. Steffen, M.A. Ansmann, R.C. Bialczak, N. Katz, E. Lucero, R. McDermott, M. Neeley, E.M. Weig, A.N. Cleland, and J.M. Martinis, “Measurement of the entanglement of two superconducting qubits via state tomography”, *Science* **313**, 1423–1425 (2006).
- [249] M. Zukowski, A. Zeilinger, M.A. Horne, and A.K. Ekert, “Event-ready-detectors Bell experiment via entanglement swapping”, *Phys. Rev. Lett.* **71**, 4287–4290 (1993).
- [250] J.W. Pan, D. Bouwmeester, H. Weinfurter, and A. Zeilinger, “Experimental entanglement swapping: entangling photons that never interacted”, *Phys. Rev. Lett.* **80**, 3891–3894 (1998).
- [251] D. Leibfried, E. Knill, S. Seidelin, J. Britton, R.B. Blakestad, J. Chiaverini, D.B. Hume, W.M. Itano, J.D. Jost, C. Langer, R. Ozeri, R. Reichle, and D.J. Wineland, “Creation of a six-atom ‘Schrödinger cat’ state”, *Nature* **438**, 639–642 (2005).

- [252] H. Häffner, W. Hänsel, C.F. Roos, J. Benhelm, D. Chek-al-kar, M. Chwalla, T. Körber, U.D. Rapol, M. Riebe, P.O. Schmidt, C. Becher, O. Gühne, W. Dür, and R. Blatt, “Scalable multiparticle entanglement of trapped ions”, *Nature* **438**, 643–646 (2005).
- [253] M. Radmark, M. Zukowski, and M. Bourennane, “Experimental tests of fidelity limits in six-photon interferometry and of rotational invariance properties of the photonic six-qubit entanglement singlet state”, *Phys. Rev. Lett.* **103**, 150501 (2009).
- [254] M. Radmark, M. Wiesniak, M. Zukowski, and M. Bourennane, “Experimental filtering of two-, four-, and six-photon singlets from a single parametric down-conversion source”, *Phys. Rev. A* **80**, 040302(R) (2009).
- [255] J.F. Poyatos, J.I. Cirac, and P. Zoller, “Quantum reservoir engineering with laser cooled trapped ions”, *Phys. Rev. Lett.* **77**, 4728–4731 (1996).
- [256] S. Diehl, A. Micheli, A. Kantian, B. Kraus, H.P. Büchler, and P. Zoller, “Quantum states and phases in driven open quantum systems with cold atoms”, *Nature Physics* **4**, 878–883 (2008).
- [257] B. Kraus, H.P. Büchler, S. Diehl, A. Kantian, A. Micheli, and P. Zoller, “Preparation of entangled states by quantum Markov processes”, *Phys. Rev. A* **78**, 042307 (2008).
- [258] M. Schlosshauer, “Decoherence, the measurement problem, and interpretations of quantum mechanics”, *Rev. Mod. Phys.* **76**, 1267–1305 (2005).
- [259] M. Brune, E. Hagley, J. Dreyer, X. Maître, A. Maali, C. Wunderlich, J.M. Raimond, and S. Haroche, “Observing the progressive decoherence of the ‘meter’ in a quantum measurement”, *Phys. Rev. Lett.* **77**, 4887–4890 (1996).
- [260] C.H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J.A. Smolin, and W.K. Wootters, “Purification of noisy entanglement and faithful teleportation via noisy channels”, *Phys. Rev. Lett.* **76**, 722–725 (1996).
- [261] C.H. Bennett, H. Bernstein, S. Popescu, and B. Schumacher, “Concentrating partial entanglement by local operations”, *Phys. Rev. A* **53**, 2046–2052 (1996).
- [262] C.H. Bennett, D.P. DiVincenzo, J.A. Smolin, and W.K. Wootters, “Mixed-state entanglement and quantum error correction”, *Phys. Rev. A* **54**, 3824–3851 (1996).
- [263] J.W. Pan, C. Simon, C. Brukner, and A. Zeilinger, “Entanglement purification for quantum communication”, *Nature* **410**, 1067–1070 (2001).
- [264] S. Haroche and J.-M. Raimond, *Exploring the Quantum: Atoms, Cavities and Photons*, Oxford University Press (2008).
- [265] C. Cohen-Tannoudji, J. Dupont-Roc, and G. Grynberg, *Atom–Photon Interactions*, Wiley (1992).
- [266] W.K. Wootters and W.H. Zurek, “A single quantum cannot be cloned”, *Nature* **299**, 802–803 (1982).
- [267] D. Dieks, “Communication by EPR devices”, *Phys. Lett. A* **92**, 271–272 (1982).
- [268] A. Peres, “How the no-cloning theorem got its name”, *Fortschritte der Phys.* **51**, 458–461 (2003).
- [269] N. Gisin and S. Massar, “Optimal quantum cloning machines”, *Phys. Rev. Lett.* **79**, 2153–2156 (1997).
- [270] D.T. Smithey, M. Beck, M.G. Raymer, and A. Faridani, “Measurement of the Wigner distribution and the density matrix of a light mode using optical homodyne tomography: application to squeezed states and the vacuum”, *Phys. Rev. Lett.* **70**, 1244–1247 (1993).
- [271] D.T. Smithey, M. Beck, J. Cooper, and M.G. Raymer, “Measurement of number-phase uncertainty relations of optical fields”, *Phys. Rev. A* **48**, 3159–3167 (1993).
- [272] U. Leonhardt, *Measuring the Quantum State of Light*, Cambridge University Press (1997).

- [273] Y. Aharonov and D. Rohrlich, *Quantum Paradoxes; Quantum Theory for the Perplexed*, Wiley-VCH (2005).
- [274] Y. Aharonov, D.Z. Albert, and L. Vaidman, “How the result of a measurement of a component of the spin of a spin-1/2 particle can turn out to be 100”, *Phys. Rev. Lett.* **60**, 1351–1354 (1988).
- [275] Y. Aharonov, S. Popescu, and J. Tollaksen, “A time-symmetric formulation of quantum mechanics”, *Physics Today* (November 2010), 27–32.
- [276] J.S. Lundeen, B. Sutherland, A. Patel, C. Stewart, and C. Bamber, “Direct measurement of the quantum wavefunction”, *Nature* **474**, 188–191 (2011).
- [277] A.E. Allahverdyan, R. Balian, and Th. M. Nieuwenhuizen, “Determining a quantum state by means of a single apparatus”, *Phys. Rev. Lett.* **92**, 120402 (2004).
- [278] C.H. Bennett and G. Brassard, “Quantum cryptography: public key distribution and coin tossing”, in *Proceedings of the IEEE International Conference on Computers Systems and Signal Processing*, Bangalore India (1984), pp. 175–179.
- [279] A.K. Ekert, “Quantum cryptography based on Bell’s theorem”, *Phys. Rev. Lett.* **67**, 661–663 (1991).
- [280] C.H. Bennett, G. Brassard, and N.D. Mermin, “Quantum cryptography without Bell’s theorem”, *Phys. Rev. Lett.* **68**, 557–559 (1992).
- [281] C.H. Bennett, G. Brassard, and A.K. Ekert, “Quantum cryptography”, *Scientific American* **267**, 50–57 (October 1992).
- [282] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, “Quantum cryptography”, *Rev. Mod. Phys.* **74**, 145–195 (2002).
- [283] C.H. Bennett, “Quantum cryptography using any two nonorthogonal states”, *Phys. Rev. Lett.* **68**, 3121–3124 (1992).
- [284] P. Townsend, J.G. Rarity, and P.R. Tapster, “Single photon interference in a 10 km-long optical fiber interferometer”, *Electron. Lett.* **29**, 634–635 (1993).
- [285] C.H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W.L. Wootters, “Teleporting an unknown quantum state via dual classical and Einstein–Podolsky–Rosen channels”, *Phys. Rev. Lett.* **70**, 1895–1898 (1993).
- [286] D. Bouwmeester, J.W. Pan, K. Mattle, M. Eibl, H. Weinfurter, and A. Zeilinger, “Experimental quantum teleportation”, *Nature* **390**, 575–579 (1997).
- [287] A. Peres, *Quantum Theory: Concepts and Methods*, Kluwer (1993); see also [94].
- [288] M. Le Bellac, *Physique Quantique*, 2nd edition, CNRS Editions and EDP Sciences (2007).
- [289] S. Massar and S. Popescu, “Optimal extraction of information from finite quantum ensembles”, *Phys. Rev. Lett.* **74**, 1259–1263 (1995).
- [290] S. Popescu, “Bell’s inequalities versus teleportation: what is nonlocality?”, *Phys. Rev. Lett.* **72**, 797–800 (1994); arXiv: quant-ph/9501020 (1995).
- [291] T. Sudbury, “The fastest way from A to B”, *Nature* **390**, 551–552 (1997).
- [292] G.P. Collins, “Quantum teleportation channels opened in Rome and Innsbruck”, *Phys. Today* **51**, 18–21 (February 1998).
- [293] Y. Xia, J. Song, P.-M. Lu, and H.-S. Song, “Teleportation of an N-photon Greenberger–Horne–Zeilinger (GHZ) polarization-entangled state using linear elements”, *J. Opt. Soc. Am.* **B 27**, A1–A6 (2010).
- [294] X.M. Jin, J.G. Ren, B. Yang, Z.H. Yi, F. Zhou, X.F. Xu, S.K. Wang, D. Yang, Y.F. Hu, S. Jiang, T. Yang, H. Yin, K. Chen, C.Z. Peng, and J.W. Pan, “Experimental free-space quantum teleportation”, *Nature Photonics* **4**, 376–381 (2010).
- [295] C.H. Bennett, “Quantum information and computation”, *Phys. Today* **48**, 24–30 (October 1995).
- [296] D.P. DiVincenzo, “Quantum computation”, *Science* **270**, 255–261 (October 1995).

- [297] C.H. Bennett, and D.P. DiVincenzo, “Quantum information and computation”, *Science* **404**, 247–255 (2000).
- [298] D. Bouwmeester, A.K. Ekert, and A. Zeilinger editors, *The Physics of Quantum Information: Quantum Cryptography, Quantum Teleportation, Quantum Computation*, Springer (2000).
- [299] D. Mermin, *Quantum Computer Science: An Introduction*, Cambridge University Press (2007).
- [300] D. Deutsch, “Quantum theory, the Church–Turing principle and the universal quantum computer”, *Proc. Roy. Soc. A* **400**, 97–117 (1985).
- [301] http://en.wikipedia.org/wiki/History_of_quantum_computing
- [302] M. Le Bellac, *Le monde quantique*, EDP Sciences (2010).
- [303] P. Shor, *Proceedings of the 55th Annual Symposium on the Foundations of Computer Science*, IEEE Computer Society Press, Los Alamitos, California (1994), pp. 124–133.
- [304] D. Mermin, “What has quantum mechanics to do with factoring?”, *Phys. Today* **60**, 8–9 (April 2007); “Some curious facts about quantum factoring”, *Phys. Today* **60**, 10–11 (October 2007).
- [305] L.K. Grover, “A fast quantum mechanical algorithm for database search”, *Proceedings, 28th Annual ACM Symposium on the Theory of Computing* (May 1996), p. 212; “From Schrödinger’s equation to quantum search algorithm”, *Am. J. Phys.* **69**, 769–777 (2001).
- [306] D. Deutsch and R. Jozsa, “Rapid solution of problems by quantum computation”, *Proceedings of the Royal Society of London A* **439**, 553–558 (1992).
- [307] D.S. Abrams and S. Lloyd, “Simulation of many-body Fermi systems on a universal quantum computer”, *Phys. Rev. Lett.* **79**, 2586–2589 (1997).
- [308] A.W. Harrow, A. Hassidim, and S. Lloyd, “Quantum algorithm for linear systems of equations”, *Phys. Rev. Lett.* **103**, 150502 (2009).
- [309] L.M.K. Vandersypen, M. Steffen, G. Breyta, C.S. Yannoni, M.H. Sherwood and I.L. Chuang, “Experimental realization of quantum Shor’s factoring algorithm using nuclear magnetic resonance”, *Nature* **414**, 883–887 (2001).
- [310] P.W. Shor, “Scheme for reducing decoherence in quantum computer memory”, *Phys. Rev. A* **52**, R2493–R2496 (1995).
- [311] A.M. Steane, “Error correcting codes in quantum theory”, *Phys. Rev. Lett.* **77**, 793–796 (1996).
- [312] J. Preskill, “Battling decoherence: the fault-tolerant quantum computer”, *Phys. Today* **52**, 24–30 (June 1999).
- [313] C.H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J.A. Smolin, and W.K. Wootters, “Purification of noisy entanglement and faithful teleportation via noisy channels”, *Phys. Rev. Lett.* **76**, 722–725 (1996).
- [314] C.H. Bennett, D.P. DiVincenzo, J.A. Smolin, and W.A. Wootters, “Mixed-state entanglement and quantum error correction”, *Phys. Rev. A* **54**, 3824–3851 (1996).
- [315] H.J. Briegel, W. Dür, J.I. Cirac, and P. Zoller, “Quantum repeaters: the role of imperfect local operations in quantum communication”, *Phys. Rev. Lett.* **81**, 5932–5935 (1998).
- [316] R.B. Griffiths and Chi-Sheng Niu, “Semiclassical Fourier transform for quantum computation”, *Phys. Rev. Lett.* **76**, 3228–3231 (1996).
- [317] F. Verstraete, M.M. Wolf, and J.I. Cirac, “Quantum computation and quantum-state engineering driven by dissipation”, *Nature Physics* **5**, 633–636 (2009).
- [318] S. Haroche and J.M. Raimond, “Quantum computing: dream or nightmare?”, *Phys. Today* **49**, 51–52 (August 1996).

- [319] P. Grangier, J.A. Levenson, and J.P. Poizat, “Quantum non demolition measurements in optics”, *Nature* **396**, 537–542 (1998).
- [320] H.D. Zeh, “On the interpretation of measurement in quantum theory”, *Found. Phys.* **I**, 69–76 (1970).
- [321] W.H. Zurek, “Pointer basis of quantum apparatus: into what mixture does the wave packet collapse?”, *Phys. Rev.* **D 24**, 1516–1525 (1981); “Environment-induced superselection rules”, *Phys. Rev.* **D 26**, 1862–1880 (1982).
- [322] W.H. Zurek, “Decoherence, einselection, and the quantum origin of the classical”, *Rev. Mod. Phys.* **75**, 715–775 (2003).
- [323] K. Hepp, “Quantum theory of measurement and macroscopic observables”, *Helv. Phys. Acta* **45**, 237–248 (1972).
- [324] J.S. Bell, “On wave packet reduction in the Coleman–Hepp model”, *Helv. Phys. Acta* **48**, 93–98 (1975); reprinted in [6].
- [325] W.H. Zurek, “Preferred states, predictability, classicality and the environment-induced decoherence”, *Progr. Theor. Phys.* **89**, 281–312 (1993); a shorter version is available in “Decoherence and the transition from quantum to classical”, *Phys. Today* **44**, 36–44 (October 1991).
- [326] W.H. Zurek, “Environment-assisted invariance, entanglement, and probabilities in quantum physics”, *Phys. Rev. Lett.* **90**, 120404 (2003).
- [327] F. Hund, “Zur Deutung der Molekelspektren III”, *Zeit. Phys.* **43**, 805–826 (1927).
- [328] J. Trost and K. Hornberger, “Hund’s paradox and the collisional stabilization of chiral molecules”, *Phys. Rev. Lett.* **103**, 023202 (2009).
- [329] Y. Aharonov, J. Anadan, S. Popescu, and L. Vaidman, “Superpositions of time evolutions of a quantum system and a quantum time-translation machine”, *Phys. Rev. Lett.* **64**, 2965–2968 (1990).
- [330] N.W.M. Richtie, J.G. Story, and R.G. Hulet, “Realization of a measurement of a weak value”, *Phys. Rev. Lett.* **66**, 1107–1110 (1991).
- [331] D.R. Solli, C.F. McCormick, R.Y. Chiao, S. Popescu, and J.M. Hickmann, “Fast light, slow light, and phase singularities: a connection to generalized weak values”, *Phys. Rev. Lett.* **92**, 043601 (2004).
- [332] N. Brunner, V. Scarani, M. Wegmüller, M. Legré, and N. Gisin, “Direct measurement of superluminal group velocity and signal velocity in an optical fiber”, *Phys. Rev. Lett.* **93**, 203902 (2004).
- [333] G.J. Pryde, J.L. O’Brien, A.G. White, T.C. Ralph, and H.M. Wiseman, “Measurement of quantum weak values of photon polarization”, *Phys. Rev. Lett.* **94**, 220405 (2005).
- [334] R. Mir, J.S. Lundeen, M.W. Mitchell, A.M. Steinberg, J.L. Garretson, and H.M. Wiseman, “A double slit ‘which way’ experiment on the complementarity-uncertainty debate”, *New. J. Phys.* **9**, 287–297 (2007).
- [335] J.S. Lundeen and A.M. Steinberg, “Experimental joint weak measurement on a photon pair as a probe of Hardy’s paradox”, *Phys. Rev. Lett.* **102**, 020404 (2009).
- [336] K. Yokota, T. Yamamoto, M. Koashi, and N. Imoto, “Direct observation of Hardy’s paradox by joint weak measurement with an entangled photon pair”, *New. J. Phys.* **11**, 033011 (2009).
- [337] P. Ben Dixon, D.J. Starling, A.N. Jordan, and J.C. Howell, “Ultrasensitive beam deflection measurement via interferometric weak value amplification”, *Phys. Rev. Lett.* **102**, 173601 (2009). D.J. Starling, P. Ben Dixon, A.N. Jordan, and J.C. Howell, “Optimizing the signal-to-noise ratio of a beam-deflection measurement with interferometric weak values”, *Phys. Rev.* **A80**, 041803 (2009).
- [338] N. Brunner and C. Simon, “Measuring small longitudinal phase shifts: weak measurements or standard interferometry”, *Phys. Rev. Lett.* **105**, 010405 (2010).

- [339] N.S. Williams and A.N. Jordan, “Weak values and the Leggett–Garg inequality in solid-state qubits”, *Phys. Rev. Lett.* **100**, 026804 (2008).
- [340] D.T. Gillespie, “The mathematics of Brownian motion and Johnson noise”, *Am. J. Phys.* **64**, 225–240 (1995).
- [341] H.P. McKean, *Stochastic Integrals*, AMS Chelsea Publishing, Providence (1969).
- [342] N. Gisin, “A simple nonlinear dissipative quantum evolution equation”, *J. Phys. A* **14**, 2259–2267 (1981).
- [343] N. Gisin, “Irreversible quantum dynamics and the Hilbert space structure of quantum dynamics”, *J. Math. Phys.* **24**, 1779–1782 (1983).
- [344] N. Gisin, “Quantum measurements and stochastic processes”, *Phys. Rev. Lett.* **52**, 1657–1660 (1984).
- [345] T.A. Brun, “A simple model of quantum trajectories”, *Am. J. Phys.* **70**, 719–737 (2002).
- [346] K. Jacobs and D.A. Steck, “A straightforward introduction to continuous quantum measurement”, *Contemp. Phys.* **47**, 279–303 (2007); arXiv: quant-ph/0611067 (2006).
- [347] V.P. Belavkin, “Non-demolition measurement and control in quantum dynamical systems”, *Proc. of CISM Seminars on Information Complexity and Control in Quantum Systems*, A. Blaquière, S. Diner, and G. Lochak editors, Springer Verlag (1987), pp. 311–329.
- [348] N.F. Mott, “The wave mechanics of α -ray tracks”, *Proc. Royal Soc. A* **126**, 79–84 (1929); reprinted in “Quantum theory of measurement”, J.A. Wheeler and W.H. Zurek editors, Princeton University Press (1983), pp. 129–134.
- [349] W. Nagourney, J. Sandberg, and H. Dehmelt, “Shelved optical electron amplifier: observation of quantum jumps”, *Phys. Rev. Lett.* **56**, 2797–2799 (1986); H. Dehmelt, “Experiments with an isolated subatomic particle at rest”, *Rev. Mod. Phys.* **62**, 525–530 (1990).
- [350] T. Sauter, W. Neuhauser, R. Blatt, and P.E. Toschek, “Observation of quantum jumps”, *Phys. Rev. Lett.* **57**, 1696–1698 (1986).
- [351] J.C. Bergquist, R.G. Hulet, W.M. Itano, and D.J. Wineland, “Observation of quantum jumps in a single atom”, *Phys. Rev. Lett.* **57**, 1699–1702 (1986).
- [352] W.M. Itano, J.C. Bergquist, R.G. Hulet, and D.J. Wineland, “Radiative decay rates in Hg^+ from observation of quantum jumps in a single ion”, *Phys. Rev. Lett.* **59**, 2732–2735 (1987).
- [353] E. Schrödinger, “Are there quantum jumps?”, *British J. Phil. Sci.* **3**, 109–123 and 233–242 (1952).
- [354] G. Greenstein and A.G. Zajonc, “Do quantum jumps occur at well-defined moments of time?”, *Am. J. Phys.* **63**, 743–745 (1995).
- [355] C. Cohen-Tannoudji and J. Dalibard, “Single-atom laser spectroscopy looking for dark periods in fluorescence light”, *Europhys. Lett.* **1**, 441–448 (1986).
- [356] M. Porriati and S. Puttermann, “Wave-function collapse due to null measurements: the origin of intermittent atomic fluorescence”, *Phys. Rev. A* **36**, 929–932 (1987).
- [357] S. Peil and G. Gabrielse, “Observing the quantum limit of an electron cyclotron: QND measurements of quantum jumps between Fock states”, *Phys. Rev. Lett.* **83**, 1287–1290 (1999).
- [358] D. Hanneke, S. Fogwell, and G. Gabrielse, “New measurement of the electron magnetic moment and the fine structure constant”, *Phys. Rev. Lett.* **100**, 120801 (2008).
- [359] M. Brune, S. Haroche, V. Lefevre, J.M. Raimond, and N. Zagury, “Quantum non-demolition measurement of small photon numbers by Rydberg-atom phase-sensitive detection”, *Phys. Rev. Lett.* **65**, 976–979 (1990).

- [360] S. Gleyzes, S. Kuhr, C. Guerlin, J. Bernu, S. Deleglise, U.B. Hoff, M. Brune, J.-M. Raimond, and S. Haroche, “Quantum jumps of light recording the birth and death of a photon in a cavity”, *Nature* **446**, 297–300 (2007); C. Guerlin, J. Bernu, S. Deleglise, C. Sayrin, S. Gleyzes, S. Kuhr, M. Brune, J.M. Raimond, and S. Haroche, “Progressive state collapse and quantum non-demolition photon counting”, *Nature* **448**, 889–893 (2007).
- [361] J. Javanainen and S.M. Yoo, “Quantum phase of a Bose–Einstein condensate with arbitrary number of atoms”, *Phys. Rev. Lett.* **76**, 161–164 (1996).
- [362] M.R. Andrews, C.G. Townsend, H.J. Miesner, D.S. Durfee, D.M. Kurn, and W. Ketterle, “Observation of interference between two Bose condensates”, *Science* **275**, 637–641 (1997).
- [363] A.J. Leggett and F. Sols, “On the concept of spontaneously broken gauge symmetry in condensed matter physics”, *Found. Phys.* **21**, 353–364 (1991).
- [364] E.P. Wigner, “Interpretation of quantum mechanics”, lectures given in 1976 at Princeton University, later published in *Quantum Theory of Measurement*, J.A. Wheeler and W.H. Zurek editors, Princeton University Press (1983), pp. 260–314; see also Wigner’s contribution in “Foundations of quantum mechanics”, *Proc. Enrico Fermi Int. Summer School*, B. d’Espagnat editor, Academic Press (1971).
- [365] N.D. Mermin, “What is quantum mechanics trying to tell us?”, *Am. J. Phys.* **66**, 753–767 (1998).
- [366] B. Misra and E.C.G. Sudarshan, “The Zeno’s paradox in quantum theory”, *J. Math. Phys.* (NY) **18**, 756–763 (1977).
- [367] A. Zeilinger, “A foundational principle for quantum mechanics”, *Found. Phys.* **29**, 631–643 (1999).
- [368] C. Brukner and A. Zeilinger, “Operationally invariant information in quantum measurements”, *Phys. Rev. Lett.* **83**, 3354–3357 (1999).
- [369] C.A. Fuchs, “Quantum foundations in the light of quantum information”, arXiv: quant-ph/0106166 (2001).
- [370] C.A. Fuchs, “Quantum mechanics as quantum information (and only a little more)”, arXiv: quant-ph/0205039 (2002).
- [371] I. Pitowsky, “Betting on the outcomes of measurements: a Bayesian theory of quantum probability”, *Studies in History and Philosophy of Modern Physics* **34**, 395–414 (2003).
- [372] G. Auletta, “Quantum information as a general paradigm”, *Found. Phys.* **35**, 787–815 (2005).
- [373] J. Bub, “Quantum probabilities: an information-theoretic interpretation”, in *Probabilities in Physics*, C. Beisbart and S. Hartmann editors, Oxford University Press (2011).
- [374] D. Deutsch and P. Hayden, “Information flow in entangled quantum systems”, *Proc. Royal Soc. A* **456**, 1759–1774 (2000).
- [375] L.E. Ballentine, “The statistical interpretation of quantum mechanics”, *Rev. Mod. Phys.* **42**, 358–381 (1970).
- [376] A.J. Leggett, “Probing quantum mechanics towards the everyday world: where do we stand?”, *Physica Scripta* **T102**, 69–73 (2002).
- [377] Y. Aharonov, J. Anandan, and L. Vaidman, “Meaning of the wave function”, *Phys. Rev. A* **47**, 4616–4626 (1993).
- [378] C. Rovelli, “Relational quantum mechanics”, *Int. J. Theor. Phys.* **35**, 1637–1678 (1996).
- [379] F. Laudisa and C. Rovelli, “Relational quantum mechanics”, *Stanford Encyclopedia of Philosophy* (2008), <http://plato.stanford.edu/entries/qm-relational/>

- [380] M. Smerlak and C. Rovelli, “Relational EPR”, *Found. Phys.* **37**, 427–445 (2007).
- [381] P. Mittelstaedt, *Quantum Logic*, Kluwer Academic Publishers (1978).
- [382] E.G. Beltrametti and G. Cassinelli, *The Logic of Quantum Mechanics*, Cambridge University Press (1984).
- [383] A. Grinbaum, “The significance of information in quantum theory”, Ph.D. thesis, Ecole polytechnique (2004), arXiv: quant-ph/0410071 (2004). “Reconstruction of quantum theory”, *Brit. J. Phil. Sci.* **58**, 387–408 (2007).
- [384] F. Zwicky, “On a new type of reasoning and some of its possible consequences”, *Phys. Rev.* **43**, 1031–1033 (1933).
- [385] G. Birkhoff and J. von Neumann, “The logic of quantum mechanics”, *Ann. Math.* **37**, 823–843 (1936).
- [386] P. Jordan, “Zur Quanten-Logik”, *Archiv der Mathematik* **2**, 166–171 (1949).
- [387] M. Strauss, “Grundlagen der modernen Physik”, in *Mikrokosmos-Makrokosmos: Philosophisch-theoretische Probleme der Naturwissenschaften, Technik und Medizin*, Akademie Verlag, Berlin (1967).
- [388] K.R. Popper, “Birkhoff and von Neumann’s interpretation of quantum mechanics”, *Nature* **219**, 682–685 (1968).
- [389] R. Hughes, “La logique quantique”, *Pour la Science* December 1981, 36–49.
- [390] H. Reichenbach, *Philosophic Foundations of Quantum Mechanics*, University of California Press (1965).
- [391] C.F. von Weizsäcker, *Göttingische Gelehrte Anzeigen* **208**, 117–136 (1954).
- [392] P. Jordan, J. von Neumann, and E. Wigner, “An algebraic generalization of the quantum mechanical formalism”, *Ann. Math.* **35**, 29–64 (1934).
- [393] I.M. Gelfand and M.A. Naimark, “On the embedding of normed rings into the ring of operators in Hilbert space”, *Mat. Sbornik* **12**, 197–213 (1943).
- [394] I.E. Segal, “Irreducible representations of operator algebras”, *Bull. Amer. Math. Soc.* **61**, 69–105 (1947); “Postulates for general quantum mechanics”, *Ann. Math.* **48**, 930–948 (1947).
- [395] R. Haag and D. Kastler, “An algebraic approach to quantum field theory”, *J. Math. Phys.* **7**, 848–861 (1964).
- [396] F.J. Murray and J. von Neumann, “On rings of operators”, *Ann. Math.* **37**, 116–229 (1936).
- [397] A. Connes, “Une classification des facteurs du type III”, *Ann. Sci. Ecole Norm. Sup.* **6**, 133–252 (1973).
- [398] A. Connes, *Noncommutative Geometry*, Academic Press (1994).
- [399] G. Mackey, *Mathematical Foundations of Quantum Mechanics*, Benjamin, New York (1963).
- [400] C. Piron, “Axiomatique quantique”, *Helv. Phys. Acta* **37**, 439–468 (1964).
- [401] J.M. Jauch and C. Piron, “On the structure of quantal proposition systems”, *Helv. Phys. Acta* **42**, 842–848 (1969).
- [402] M.P. Solèr, “Characterization of Hilbert spaces by orthomodular spaces”, *Comm. Algebra* **23**, 219–243 (1995).
- [403] B. Coecke and E.O. Paquette, “Categories for the practicing physicist”, arXiv: 0905.3010v2 [quant-ph]; in *New Structures for Physics*, Springer (2011), pp. 173–286.
- [404] B. Coecke, “Quantum picturalism”, *Contemp. Phys.* **51**, 59–83 (2010).
- [405] H. Barnum and A. Wilce, “Information processing in convex operational theories”, *Electronic Notes in Theor. Computer Sci.* **12**, 3–15 (2011).

- [406] A. Wilce, “Quantum logic and probability theory”, *Stanford Encyclopedia of Philosophy* (2008), <http://plato.stanford.edu/entries/qt-quantlog/>.
- [407] A.M. Gleason, “Measures on the closed subspaces of a Hilbert space”, *J. Math. and Mech.* **6**, 885–893 (1957).
- [408] P. Bush, “Quantum states and generalized observables: a simple proof of Gleason’s theorem”, *Phys. Rev. Lett.* **91**, 120403 (2003).
- [409] J.S. Bell, “Are there quantum jumps?”, in *Schrödinger – Centenary Celebration of a Polymath*, C.W. Kilmister editor, Cambridge University Press (1987), p. 41; see also Chapter 22 of [6].
- [410] J.S. Bell, “Beables for quantum field theory”, CERN-TH.4035/84 (August 2 1984); *Phys. Rep.* **137**, 49–54 (1986); Chapter 19 of [6].
- [411] L. de Broglie, “La mécanique ondulatoire et la structure atomique de la matière et du rayonnement”, *J. Physique et le Radium*, série VI, tome VIII, 225–241 (1927); “Interpretation of quantum mechanics by the double solution theory”, *Ann. Fond. Louis de Broglie* **12**, 1 (1987).
- [412] L. de Broglie, *Tentative d’Interprétation Causale et Non-linéaire de la Mécanique Ondulatoire*, Gauthier-Villars, Paris (1956).
- [413] L. de Broglie, *Les Incertitudes d’Heisenberg et l’Interprétation Probabiliste de la Mécanique Ondulatoire*, Gauthier-Villars and Bordas, Paris (1982).
- [414] J.T. Cushing, *Quantum Mechanics*, University of Chicago Press (1994).
- [415] D. Bohm, “Proof that probability density approaches $|\Psi|^2$ in causal interpretation of quantum theory”, *Phys. Rev.* **89**, 458–466 (1953).
- [416] E. Madelung, “Quantentheorie in hydrodynamische Form”, *Z. Phys.* **40**, 322–326 (1927).
- [417] D. Dürr, S. Goldstein, and N. Zanghì, “Quantum equilibrium and the origin of absolute uncertainty”, *J. Stat. Phys.* **67**, 843–907 (1992).
- [418] D. Bohm and J.P. Vigier, “Model of the causal interpretation of quantum theory in terms of a fluid with irregular fluctuations”, *Phys. Rev.* **96**, 208–216 (1954).
- [419] A. Valentini, “Signal-locality, uncertainty, and the subquantum H -theorem” I, *Phys. Lett. A* **156**, 5–11 (1991); II, *Phys. Lett. A* **158**, 1–8 (1991).
- [420] A. Valentini and H. Westman, “Dynamical origin of quantum probabilities”, *Proc. Roy. Soc. A* **461**, 253–272 (2004).
- [421] A. Valentini, “Signal-locality in hidden-variables theories”, *Phys. Lett. A* **297**, 273–278 (2002); “Beyond the quantum”, *Physics World* (November 2009).
- [422] P. Holland, “Hamiltonian theory of wave and particle in quantum mechanics II: Hamilton–Jacobi theory and particle back-reaction”, *Nuov. Cim. B* **116**, 1143–1172 (2001).
- [423] C. Philippidis, C. Dewdney, and B.J. Hiley, “Quantum interference and the quantum potential”, *Nuov. Cim.* **52 B**, 15–23 (1979).
- [424] P.R. Holland, *The Quantum Theory of Motion*, Cambridge University Press (1993).
- [425] J.S. Bell, “De Broglie–Bohm, delayed choice double slit experiment, and density matrix”, *International Journal of Quantum Chemistry* **18**, supplement symposium 14, 155–159 (1980); Chapter 14 of [6].
- [426] B.J. Hiley, “Welcher Weg experiments from the Bohm perspective”, contribution to the Växjö conference (2005), <http://www.bbk.ac.uk/tpru/BasilHiley/WelcherWegBohmBJH2.pdf>.
- [427] L. Vaidman, “The reality of Bohmian quantum mechanics or can you kill with an empty wave bullet?”, *Found. Phys.* **35**, 299–312 (2005).
- [428] E. Deotto and G.C. Ghirardi, “Bohmian mechanics revisited”, *Found. Phys.* **28**, 1–30 (1998).

- [429] B.G. Englert, M.O. Scully, G. Süssmann, and H. Walther, “Surrealistic Bohm trajectories”, *Z. Naturforschung* **47a**, 1175–1186 (1992).
- [430] C. Dewdney, P.R. Holland, and A. Kyprianidis, “What happens in a spin measurement?”, *Phys. Lett. A* **119**, 259–267 (1986).
- [431] E.P. Wigner, “Rejoinder”, *Am. J. Phys.* **39**, 1097 (1971).
- [432] J. Clauser, “Von Neumann’s informal hidden-variable argument”, *Am. J. Phys.* **39**, 1095 (1971); “Reply to Dr Wigner’s objections”, *Am. J. Phys.* **39**, 1098 (1971).
- [433] D. Dürr, S. Goldstein, R. Tumulka, and N. Zanghì, “Bohmian mechanics and quantum field theory”, *Phys. Rev. Lett.* **93**, 090402 (2004).
- [434] D. Dürr, S. Goldstein, R. Tumulka, and N. Zanghì, “Trajectories and particle creation and annihilation in quantum field theory”, *J. Phys. A Math. Gen.* **36**, 4143–4149 (2003).
- [435] W. Struyve, “Field beables for quantum field theory”, arXiv: 0707.3685v2 [quant-ph] (2007).
- [436] K. Berndl, D. Dürr, S. Goldstein, and N. Zanghì, “Nonlocality, Lorentz invariance, and Bohmian quantum theory”, *Phys. Rev. A* **53**, 2062–2073 (1996).
- [437] H. Nikolić, “Relativistic quantum mechanics and the Bohmian interpretation”, *Found. Phys. Lett.* **18**, 549–561 (2005); “QFT as pilot-wave theory of particle creation and destruction”, *J. Mod. Phys. A* **25**, 1477–1505 (2010); “Bohmian mechanics in relativistic quantum mechanics, quantum field theory and string theory”, *J. Phys. Conference Series* **67**, 012035 (2007).
- [438] J.S. Bell, Chapter 18 of [47] (page 128 of [6]).
- [439] C. Dewdney, L. Hardy, and E.J. Squires, “How late measurements of quantum trajectories can fool a detector”, *Phys. Lett. A* **184**, 6–11 (1993).
- [440] C. Cohen-Tannoudji, B. Diu, and F. Laloë, *Mécanique quantique*, Hermann (1973 and 1977); *Quantum mechanics*, Wiley (1977).
- [441] R.B. Griffiths, “Bohmian mechanics and consistent histories”, *Phys. Lett. A* **261**, 227–234 (1999).
- [442] M. Correggi and G. Morchio, “Quantum mechanics and stochastic mechanics for compatible observables at different times”, *Ann. Physics* **296**, 371–389 (2002).
- [443] A. Neumaier, “Bohmian mechanics contradicts quantum mechanics”, arXiv: quant-ph/0001011 (2000).
- [444] G. Brida, E. Cagliero, G. Falzetta, M. Genovese, M. Gramegna, and C. Novero, “A first experimental test of the de Broglie–Bohm theory against standard quantum mechanics”, arXiv: quant-ph/0206196 (2002).
- [445] P. Ghose, “An experiment to distinguish between de Broglie–Bohm and standard quantum mechanics”, arXiv: quant-ph/0003037 (2003).
- [446] R.P. Feynman and A.R. Hibbs, *Quantum Mechanics and Path Integrals*, McGraw Hill (1965).
- [447] I. Fényes, “Eine wahrscheinlichkeitstheoretische Begründung und Interpretation der Quantenmechanik”, *Zeit. Physik* **132**, 81–106 (1952).
- [448] E. Nelson, “Derivation of the Schrödinger equation from Newtonian mechanics”, *Phys. Rev.* **150**, 1079–1085 (1966).
- [449] R. Werner, “A generalization of stochastic mechanics and its relation to quantum mechanics”, *Phys. Rev. D* **34**, 463–469 (1986).
- [450] T.C. Wallstrom, “Inequivalence between the Schrödinger equation and the Madelung hydrodynamic equations”, *Phys. Rev. A* **49**, 1613–1617 (1994).
- [451] P. Damgaard and H. Hüffel editors, *Stochastic Quantization*, World Scientific (1988).
- [452] M. Masujima, *Path Integral Quantization and Stochastic Quantization*, Springer Verlag (2000 and 2009).

- [453] G. Parisi and Y.-S. Wu, "Perturbation theory without gauge fixing", *Sci. Sin.* **24**, 483–496 (1981).
- [454] E. Gozzi, "Functional-integral approach to Parisi-Wu stochastic quantization: scalar theory", *Phys. Rev. D* **28**, 1922–1930 (1983).
- [455] M. Dickson and D. Dieks, "Modal interpretation of quantum mechanics", *Stanford Encyclopedia of Philosophy* (2007), <http://plato.stanford.edu/entries/qm-modal/>
- [456] B.C. van Fraassen, "A formal approach to the philosophy of science", in *Paradigms and Paradoxes: The Philosophical Challenge of the Quantum Domain*, R. Colodny editor, University of Pittsburgh Press (1972), pp. 303–366; "The Einstein–Podolsky–Rosen paradox", *Synthese* **29**, 291–309 (1974); *Quantum Mechanics: An Empiricist View*, Clarendon Press, Oxford, (1991).
- [457] S. Kochen, "A new interpretation of quantum mechanics", in *Symposium on the Foundations of Modern Physics*, P. Mittelstaedt and P. Lahti editors, World Scientific (1985), pp. 151–169.
- [458] D. Dieks, "The formalism of quantum theory: an objective description of reality?", *Annalen der Physik* **500**, 174–190 (1988); "Quantum mechanics without the projection postulate and its realistic interpretation", *Found. Phys.* **19**, 1397–1423 (1989); "Resolution of the measurement problem through decoherence of the quantum state" *Phys. Lett. A* **142**, 439–446 (1989); "Modal interpretation of quantum mechanics, measurements, and macroscopic behaviour", *Phys. Rev. A* **49**, 2290–2300 (1994).
- [459] R. Healey, *The Philosophy of Quantum Mechanics: An Interactive Interpretation*, Cambridge University Press (1989); "Measurement and quantum indeterminateness", *Found. Phys. Lett.* **6**, 307–316 (1993).
- [460] G. Bacciagaluppi, "Topics in the modal interpretation of quantum mechanics", dissertation, Cambridge University (1996); "Delocalized properties in the modal interpretation of a continuous model of decoherence", *Found. Phys.* **30**, 1431–1444 (2000).
- [461] M. Dickson, "Wavefunction tails in the modal interpretation", *Proceedings of the Philosophy of Science Association 1994*, D. Hull, M. Forbes, and R. Burian editors, Vol. 1, 366–376 (1994).
- [462] J. Berkovitz and M. Hemmo, "Modal interpretations of quantum mechanics and relativity: a reconsideration", *Found. Phys.* **35**, 373–397 (2005).
- [463] R. Healey, "Modal interpretation, decoherence, and the quantum measurement problem", in *Quantum Measurement: Beyond Paradox*, R. Healey and G. Hellmann editors, *Minnesota Studies in the Philosophy of Science* **17**(1998), pp. 52–86.
- [464] W. Myrvold, "Modal interpretation and relativity", *Found. Phys.* **32**, 1173–1184 (2002).
- [465] R. Clifton, "The modal interpretation of algebraic quantum field theory", *Phys. Lett. A* **271**, 167–177 (2000).
- [466] L. Diosi, "Quantum stochastic processes as models for state vector reduction", *J. Phys. A* **21**, 2885–2898 (1988).
- [467] R. Haag, "Fundamental irreversibility and the concept of events", *Comm. Math. Phys.* **132**, 245–251 (1990); "An evolutionary picture for quantum physics", *Comm. Math. Phys.* **180**, 733–743 (1996).
- [468] A. Jadczyk, "On quantum jumps, events, and spontaneous localization models", *Found. Phys.* **25**, 743–762 (1995).
- [469] P. Pearle, "How stands collapse I", *J. Phys. A: Math. Theor.* **40**, 3189–3204 (2007).
- [470] P. Pearle, "On the time it takes a state vector to reduce", *J. Stat. Phys.* **41**, 719–727 (1985).
- [471] A. Barchielli, L. Lanz, and G.M. Prosperi, "A model for the macroscopic description and continual observations in quantum mechanics", *Nuov. Cim.* **42 B**, 79–121 (1982).

- [472] A. Barchielli, “Continual measurements for quantum open systems”, *Nuov. Cim.* **74 B**, 113–138 (1983); “Measurement theory and stochastic differential equations in quantum mechanics”, *Phys. Rev. A* **34**, 1642–1648 (1986).
- [473] F. Benatti, G.C. Ghirardi, A. Rimini, and T. Weber, “Quantum mechanics with spontaneous localization and the quantum theory of measurement”, *Nuov. Cim.* **100 B**, 27–41 (1987).
- [474] F. Benatti, G.C. Ghirardi, A. Rimini, and T. Weber, “Operations involving momentum variables in non-hamiltonian evolution equations”, *Nuov. Cim.* **101 B**, 333–355 (1988).
- [475] P. Blanchard, A. Jadczyk, and A. Ruschhaupt, “How events come into being: EEQT, particle tracks, quantum chaos and tunneling time”, in *Mysteries, Puzzles and Paradoxes in Quantum Mechanics*, R. Bonifacio editor, American Institute of Physics, AIP Conference Proceedings, no. 461 (1999); *J. Mod. Optics* **47**, 2247–2263 (2000).
- [476] P. Pearle, “Combining stochastic dynamical state-vector reduction with spontaneous localization”, *Phys. Rev. A* **39**, 2277–2289 (1989).
- [477] G.C. Ghirardi, P. Pearle, and A. Rimini, “Markov processes in Hilbert space and continuous spontaneous localization of systems of identical particles”, *Phys. Rev. A* **42**, 78–89 (1990).
- [478] P. Pearle, “Cosmogenesis and collapse”, arXiv: 1003.5582v2 [gr-qc] (2010); *Found. Phys.* **42**, 4–18 (2012).
- [479] *Experimental Metaphysics: Quantum Mechanical Studies for Abner Shimony*, Festschrift volumes 1 and 2, R.S. Cohen, M.A. Horne, and J.J. Stachel editors, *Boston Studies in the Philosophy of Science*, vol. 193 and 194, Kluwer Academic Publishers (1997); P. Pearle, vol. 1, p. 143; G. Ghirardi and T. Weber, vol. 2, p. 89.
- [480] A. Shimony, “Desiderata for a modified quantum dynamics”, pp. 49–59 in “PSA 1990 vol. 2, *Proceedings of the 1990 Biennial Meeting of the Philosophy of Science Association*, A. Fine, M. Forbes, and L. Wessel editors, Philosophy of Science Association, East Lansing, MI (USA).
- [481] P. Pearle, “How stands collapse II”, in *Quantum Reality, Relativistic Causality and Closing the Epistemic Circle: Essays in Honour of Abner Shimony*, W. Mryvold and J. Christian editors, Springer (2009), pp. 257–292.
- [482] L.F. Santos and C.O. Escobar, “A proposed solution to the tail problem of dynamical reduction models”, *Phys. Lett. A* **278**, 315–318 (2001).
- [483] L. Diosi, “Continuous quantum measurement and Itô formalism”, *Phys. Lett.* **129 A**, 419–423 (1988).
- [484] L. Diosi, “Models for universal reduction of macroscopic quantum fluctuations”, *Phys. Rev. A* **40**, 1165–1174 (1989).
- [485] G.C. Ghirardi, R. Grassi, and A. Rimini, “Continuous-spontaneous-reduction model involving gravity”, *Phys. Rev. A* **42**, 1057–1064 (1990).
- [486] R. Penrose, *The Emperor’s New Mind*, Oxford University Press (1989); *Shadows of the Mind*, Oxford University Press (1994).
- [487] N. Gisin, “Stochastic quantum dynamics and relativity”, *Helv. Phys. Acta* **62**, 363–371 (1989).
- [488] G.C. Ghirardi, R. Grassi, and P. Pearle, “Relativistic dynamical reduction models: general framework and examples”, *Found. Phys.* **20**, 1271–1316 (1990).
- [489] P. Pearle, “Completely quantized collapse and consequences”, *Phys. Rev. A* **72**, 022112 (2005).
- [490] D.J. Bedingham, “Relativistic state reduction dynamics”, arXiv: 1003.2774v2 [quant-ph] (2010).

- [491] D.J. Bedingham, D. Dürr, G. Ghirardi, S. Goldstein, R. Tumulka, and N. Zanghi, “Matter density and relativistic models of wave function collapse”, arXiv: 1111.1425v2 [quant-ph] (2011).
- [492] N. Gisin, “Weinberg’s non-linear quantum mechanics and supraluminal communications”, *Phys. Lett. A* **143**, 1–2 (1990).
- [493] R. Tumulka, “On spontaneous wave function collapse and quantum field theory”, *Proc. Roy. Soc. A* **462**, 1897–1908 (2006); “A relativistic version of the Ghirardi–Rimini–Weber model”, *J. Stat. Phys.* **125**, 821–840 (2006); “Collapse and relativity”, arXiv: quant-ph/0602208 (2006).
- [494] A. Bassi and G. Ghirardi, “Dynamical reduction models”, *Phys. Rep.* **379**, 257–426 (2003).
- [495] S. Weinberg, “Precision tests of quantum mechanics”, *Phys. Rev. Lett.* **62**, 485–488 (1989).
- [496] K. Wódkiewicz and M.O. Scully, “Weinberg’s nonlinear wave mechanics”, *Phys. Rev. A* **42**, 5111–5116 (1990).
- [497] P. Pearle, J. Ring, J.I. Collar, and F.T. Avignone, “The CSL collapse model and spontaneous radiation: an update”, *Found. Phys.* **29**, 465–480 (1998).
- [498] H.S. Miley, F.T. Avignone, R.L. Brodzinski, J.I. Collar, and J.H. Reeves, “Suggestive evidence for the two-neutrino double- β decay of ^{76}Ge ”, *Phys. Rev. Lett.* **65**, 3092–3095 (1990).
- [499] K. Hornberger, S. Gerlich, P. Haslinger, S. Nimmrichter, and M. Arndt, “Quantum interference of clusters and molecules”, *Rev. Mod. Phys.* **84**, 157–173 (2012).
- [500] G.C. Ghirardi, “Quantum superpositions and definite perceptions: envisaging new feasible tests”, *Phys. Lett. A* **262**, 1–14 (1999).
- [501] J. Dalibard, Y. Castin, and K. Mølmer, “Wave function approach to dissipative processes in quantum optics”, *Phys. Rev. Lett.* **68**, 580–583 (1992).
- [502] K. Mølmer, Y. Castin, and J. Dalibard, “Monte Carlo wave-function method in quantum optics”, *Journ. Optical. Soc. Am. B* **10**, 524–538 (1993).
- [503] H.J. Carmichael, *An Open System Approach to Quantum Optics*, Lecture Notes in Physics, monograph 18, Springer-Verlag (1993).
- [504] M.B. Plenio and P.L. Knight, “The quantum-jump approach to dissipative dynamics in quantum optics”, *Rev. Mod. Phys.* **70**, 101–141 (1998).
- [505] N. Gisin and I.C. Percival, “The quantum-state diffusion model applied to open systems”, *J. Phys. A* **25**, 5677–5691 (1992); “Quantum state diffusion, localization and quantum dispersion entropy”, **26**, 2233–2243 (1993); “The quantum state diffusion picture of physical processes”, **26**, 2245–2260 (1993).
- [506] I.C. Percival, *Quantum State Diffusion*, Cambridge University Press (1998).
- [507] J.G. Cramer, “The transactional interpretation of quantum mechanics”, *Rev. Mod. Phys.* **58**, 647–687 (1986); in an appendix, this article contains a review of the various interpretations of quantum mechanics.
- [508] J.G. Cramer, “Generalized absorber theory and the Einstein–Podolsky–Rosen paradox”, *Phys. Rev. D* **22**, 362–376 (1980).
- [509] R. Omnès, “Logical reformulation of quantum mechanics”, *J. Stat. Phys.* **53**, “I: Foundations”, 893–932; “II: Interferences and the EPR experiments”, 933–955; “III: Classical limit and irreversibility”, 957–975 (1988).
- [510] M. Gell-Mann and J.B. Hartle, “Classical equations for quantum systems”, *Phys. Rev. D* **47**, 3345–3382 (1993).
- [511] R. Omnès *The Interpretation of Quantum Mechanics*, Princeton University Press (1994); *Understanding Quantum Mechanics*, Princeton University Press (1999).

- [512] R.B. Griffiths and R. Omnès, “Consistent histories and quantum measurements”, *Phys. Today* **52**, 26–31 (August 1999).
- [513] P.C. Hohenberg, “*Colloquium*: an introduction to consistent quantum theory”, *Rev. Mod. Phys.* **82**, 2835–2844 (2010).
- [514] Y. Aharonov, P.G. Bergmann, and J.L. Lebowitz, “Time symmetry in the quantum process of measurement”, *Phys. Rev.* **B 134**, 1410–1416 (1964).
- [515] R.B. Griffiths, “Consistent histories and quantum reasoning”, *Phys. Rev.* **A 54**, 2759–2774 (1996).
- [516] R.B. Griffiths, “Choice of consistent family, and quantum incompatibility”, *Phys. Rev.* **A 57**, 1604–1618 (1998).
- [517] “Observant readers take the measure of novel approaches to quantum theory: some get Bohmed”, *Phys. Today* **52**, 11–15 and 89–92 (February 1999).
- [518] R.B. Griffiths, “Correlations in separated quantum systems: a consistent history analysis of the EPR problem”, *Am. J. Phys.* **55**, 11–17 (1987).
- [519] F. Dowker and A. Kent, “Properties of consistent histories”, *Phys. Rev. Lett.* **75**, 3038–3041 (1995); “On the consistent histories approach to quantum mechanics”, *J. Stat. Phys.* **82**, 1575–1646 (1996).
- [520] A. Kent, “Quasiclassical dynamics in a closed quantum system”, *Phys. Rev.* **A 54**, 4670–4675 (1996).
- [521] T.A. Brun, “Continuous measurements, quantum trajectories, and decoherent histories”, *Phys. Rev.* **A 61**, 042107 (2000).
- [522] H. Everett III, “Relative state formulation of quantum mechanics”, *Rev. Mod. Phys.* **29**, 454–462 (1957); reprinted in *Quantum Theory and Measurement*, J.A. Wheeler and W.H. Zurek editors, Princeton University Press (1983), pp. 315–323.
- [523] H. Everett III, Letter to L.D. Raub dated April 7 (1983), <http://dspace.nacs.uci.edu/xmlui/handle/10575/1205>.
- [524] D. Deutsch, “The structure of the multiverse”, *Proc. Roy. Soc. London* **A 458**, 2911–2923 (2002).
- [525] D. Deutsch, “Quantum theory of probability and decisions”, *Proc. Roy. Soc. London* **A 455**, 3129–3137 (1999).
- [526] M.A. Rubin, “Relative frequency and probability in the Everett interpretation of Heisenberg–picture quantum mechanics”, *Found. Phys.* **33**, 379–405 (2002).
- [527] D. Wallace, “Everettian rationality: defending Deutsch’s approach to probability in the Everett interpretation”, *Stud. Hist. Phil. Mod. Phys.* **34**, 415–438 (2003).
- [528] S. Saunders, “Derivation of the Born rule from operational assumptions”, *Proc. Roy. Soc. London* **A 460**, 1771–1788 (2004).
- [529] W.H. Zurek, “Probabilities from entanglement, Born’s rule $p_k = |\Psi_k|^2$ from envariance”, *Phys. Rev.* **A 71**, 052105 (2005).
- [530] A. Kent, “Against many world interpretations”, *Int. Journ. Mod. Phys* **A5**, 1745–1762 (1990).
- [531] P. Van Esch, “On the Born rule and the Everett programme”, *Ann. Found. Louis de Broglie* **32**, 51–59 (2007).
- [532] D. Wallace, “Quantum probability from subjective likelihood: improving on Deutsch’s proof of the probability rule”, *Studies in History and Philosophy of Modern Physics* **38**, 311–332 (2007).
- [533] H. Price, “Probability in the Everett world: comments on Wallace and Greaves”, arXiv: quant-ph/0604191 (2006); “Decisions, decisions, decisions: can Savage salvage the Everettian probability?”, arXiv: quant-ph/0802.1390 (2008).
- [534] H.D. Zeh, “Roots and fruits of decoherence”, *Séminaire Poincaré* **1**, 115–129 (2005); available at <http://www.bourbaphy.fr/>.

- [535] D. Deutsch and P. Hayden, “Information flow in entangled quantum systems”, *Proc. Roy. Soc. London A* **456**, 1759–1774 (2000).
- [536] B. DeWitt, *The Global Approach to Quantum Field Theory*, vol. 1, Clarendon Press (2003), p. 144.
- [537] M. Tegmark, “Parallel universes”, in *Science and Ultimate Reality: From Quantum to Cosmos*, J.D. Barrow, P.C.W. Davies, and C.L. Harper editors, Cambridge University Press (2003); “Many worlds in context”, arXiv: 0905.2182v2 [quant-ph] (2009) in *Many Worlds? Everett, Quantum Theory and Reality*, S. Saunders, J. Barrett, A. Kent, and D. Wallace editors, Oxford University Press (2010).
- [538] T. Damour, “Einstein 1905–1955: son approche de la physique”, *Séminaire Poincaré* **1**, 1–25 (2005); available at <http://www.bourbaphy.fr/>.
- [539] J.S. Bell, “The measurement theory of Everett and de Broglie’s pilot wave”, in *Quantum Mechanics, Determinism, Causality, and Particles*, M. Flato *et al.* editors, Dordrecht-Holland, D. Reidel (1976), pp. 11–17; Chapter 11 of 2004 edition of [6].
- [540] A. Einstein, “Quantentheorie des einatomigen idealen Gases”, *Sitzungsberichte der Preussischen Akademie der Wissenschaften* **1**, 3–14 (1925).
- [541] M.H. Anderson, J.R. Ensher, M.R. Matthews, C.E. Wieman, and E.A. Cornell, “Observation of Bose–Einstein condensation in a dilute atomic vapor”, *Science* **269**, 198–201 (1995).
- [542] K.B. Davis, M.-O. Mewes, M.R. Andrews, N.J. van Druten, D.S. Durfee, D.M. Kurn, and W. Ketterle, “Bose–Einstein condensation in a gas of sodium atoms”, *Phys. Rev. Lett.* **75**, 3969–3973 (1995).
- [543] H. Hertz, *Miscellaneous Papers*, translated from first German edition (1895) by D.E. Jones and G.A. Schott, Macmillan (London, 1896), vol. 1, p. 318.

Index

- additional/hidden variables, 72, 207, 231
- Aharonov, 221
- algebraic theory, 227
- Aspect, 66, 95
- axiomatic, 227

- Bacchiagaluppi, 263
- Ballentine, 220
- BB84 protocol, 155
- BCHSH inequality, 58
- beable, 231
- Bell, 2, 18, 34, 51, 92, 96, 118, 231, 278
- Bell 1964 inequality, 67
- Bell inequalities, 57
- Bell theorem, 56
- Bell theorem (various forms), 66
- Bell-Kochen–Specker theorem, 114
- Berkowitz-Hemmo, 263
- biorthonormal decomposition, 128
- Birkhoff, 226
- Bohm, 233
- Bohm-Bub, 265
- Bohmian correlations, 353
- Bohmian measurement, 243
- Bohmian theory, 233
- Bohmian time correlations, 256
- Bohr, 12, 16, 31, 36, 49
- Born, 32, 232
- Born rule, 8, 309
- Born rule (generalized), 347
- Bose–Einstein condensates, 204
- Bose–Einstein condensation, 6, 54, 301
- branching of the state vector, 295
- Brownian motion, 188
- Bush theorem, 230

- Cabello, 108
- characterizing quantum mechanics, 86
- chiral molecules, 175
- Cirelson (theorem), 77
- Clauser, 65

- Commins, 65
- communication loophole, 94
- consciousness, 213
- consistent families, 362
- consistent histories, 281, 283
- conspiracy, 90, 94
- conspiracy loophole, 94
- contextuality, 73, 104, 117
- continuous measurements, 184
- continuous spontaneous localization, 270
- Copenhagen interpretation, 5, 15
- correlation interpretation, 216
- correlations, 42, 45, 125, 353
- counterfactuality, 73, 97
- Cramer, 280
- cryptographic key, 154
- cryptography, 154
- CSL, 270
- cyclotron, 200

- d’Espagnat, 230
- De Broglie, 4, 34, 232
- Debye, 4
- decoherence, 136, 212
- decoherent history, 281
- Dehmelt, 196
- detection loophole, 90
- deterministic boxes, 82
- Deutsch, 296
- Deutsch–Josza, 165
- DeWitt, 299
- Dickson, 263
- Dieks, 263
- Diosi, 273
- Dirac, 6, 10, 15, 33
- distillation, 141
- double solution (theory of the), 232

- efficiency loophole, 90
- eigenvectors, 306
- Einstein, 3, 22, 33, 220

- Einstein, Podolsky, and Rosen, 38
- electron in a trap, 200
- empty waves, 242, 243
- Englert, Scully, and Walther, 328
- entanglement, 120
- entanglement measures, 129
- entanglement swapping, 134
- entropy, 128, 314
- envariance, 296
- EPR, 38
- EPR protocol, 158
- EPR with macroscopic systems, 54
- EPRB, 45
- error correction codes, 166
- Everett interpretation, 140, 292
- Evolution operator, 311

- families of histories, 281
- fatalism, 95
- flash ontology, 275
- formal theory, 227
- forms of the Bell theorem, 66
- free will, 64, 72, 95, 298
- Freedman, 66
- Frenkel, 232

- Gabrielse, 200
- gambler's ruin game, 271
- Gell-Mann, 281
- Ghirardi, Rimini, and Weber, 267
- Ghirardi–Grassi–Rimini, 274
- Ghirardi–Pearle–Rimini, 272
- GHZ, 100
- Gleason theorem, 228
- Gottfried, 35
- gravity, 273
- Greenberger, Horne, and Zeilinger, 100
- Griffiths, 281
- Grover algorithm, 165
- guiding formula, 234

- Hadamard gate, 165
- Hardy, 111
- Hartle, 16
- Healey, 263
- Heisenberg, 6, 32
- Hertz, 301
- hidden/additional variables, 72, 207, 231
- histories, 290
- history interpretation, 281
- hits, 267
- Holt, 65
- Horne, 65
- Horodecki, 132
- Hund paradox, 175

- information point of view, 219
- instantaneous signaling, 80
- interaction-free measurement, 27
- ion in trap, 196

- Jammer, 36
- Jordan, 6, 33

- Kochen, 114, 262
- Kochen–Specker theorem, 114
- Kocher, 65
- Kraus sum, 146

- Landau and Lifshitz, 33
- Landau levels, 200
- Leggett, 35, 300
- Leggett–Garg, 75
- Leggett–Sols, 207
- Lindblad form, 147, 272
- local realism, 117
- locality, 47, 73, 97
- logical boxes, 82
- London and Bauer, 18
- London–Bauer, 213
- loopholes, 90

- Mackey, 227
- macroscopic decoherence, 212
- manipulating additional variables, 350
- maximal violation, 78
- measurement (Von Neumann), 168
- measurements at different times, 345
- Mercury ion, 198
- Mermin, 34, 70, 117, 218
- Mermin inequality, 70
- modal interpretation, 261
- modified Schrödinger, 264
- monogamy, 130
- multiverse, 294
- MWI (many-world interpretation), 292

- negative measurement, 27
- Nelson mechanics, 259
- Neumann (von), 1, 10, 19, 168, 212, 226, 314
- no-cloning theorem, 150
- no-crossing rule, 240
- no-determination theorem, 153
- no-signaling conditions, 81
- non-deterministic inequalities, 330
- non-locality, 97, 207
- non-locality in Bohmian theory, 241, 249, 257
- non-separability, 52

- observer, 18
- Omnès, 281
- open quantum system, 279

- pair selection, 336
- pair selection loophole, 90
- Parisi–Wu, 261
- part and the whole, 121
- partial trace, 138, 319

Pearle, 51, 92, 266, 270, 274
 Penrose, 274
 Peres, 15, 116, 132, 224, 300
 phase of Bose–Einstein condensates, 204
 pilot wave theory, 232
 Planck, 2
 pointer states, 173
 Popescu–Rohrlich boxes, 85
 Popper, 226
 POVM, 179, 230
 product rule, 104, 116
 protocol for key exchange, 155
 pure state, 312
 purification, 141

 QND, 172
 quantum computation, 163
 quantum cryptography, 154
 quantum equilibrium distribution, 235
 quantum gate, 164
 quantum information, 163
 quantum jumps, 196, 203
 quantum logic, 226
 quantum non-demolition, 172
 quantum non-demolition measurement, 202
 quantum reduction, 195
 quantum velocity, 234
 qubit, 163

 reality of Bohmian trajectories, 250
 Reichenbach, 227
 relational interpretation, 222
 relative state interpretation, 292
 resolution of unity, 229
 retrodictive (Bohmian theory), 250
 Rosenfeld, 34
 Rovelli, 222

 Saunders, 296
 Schmidt decomposition, 126
 Schrödinger, 4, 33, 199
 Schrödinger cat, 21, 139
 separability criterion, 131
 separable quantum theory, 332
 Shimony, 34, 65
 Shor algorithm, 165
 singlet state, 57
 SL, 267

Specker, 114
 splitting of the state vector, 295
 spontaneous localization, 267
 Stapp, 15, 35
 state vector, 7, 304
 statistical interpretation, 220
 status of the state vector, 13
 Stern–Gerlach, 322
 stochastic boxes, 82
 stochastic quantization, 261
 successive measurements, 214
 superdeterminism, 95, 298
 superluminal communication, 341

 teleportation, 160
 tensor product, 316
 trace, 308
 trajectories (Bohmian), 238
 transactional interpretation, 280
 Tumulka, 275
 two-particle interference, 242

 unitary operator, 307
 universal wave function, 292

 Van Fraassen, 261
 Van Kampen, 36, 328
 Von Neumann, 1, 10, 19, 168, 212, 226, 314
 Von Neumann regress, 19
 Von Weizsäcker, 36, 227

 W state, 136
 Wallace, 296
 wave function, 321
 weak measurements, 181
 Weinberg, 275
 Wiener process, 187, 191
 Wiener–Siegel, 233, 265
 Wigner formula, 216, 345
 Wigner inequalities, 68
 Wigner interpretation, 213
 Wigner’s friend, 26

 Zeh, 174
 Zeno effect, 176
 Zurek, 174, 296
 Zwicky, 226